SAHRDAYA AUTONOMOUS COLLEGE OF ENGINEERING & TECHNOLOGY

A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY | MANAGED BY IRINJALAKUDA DIOCESAN EDUCATION TRUST

Estd.2002

Approved by AICTE & Affiliated to APJ Abdul Kalam Technological University | Accredited by:

M. Tech
Curriculum (2024) and Syllabus - Semester I to IV
Computer Science and Engineering
Branch Code: CSE

(SHR/AC/Auto/Acad.Council/M.Tech/2/Curri./CSE)

Recommended by BoS on 30/08/2024
Approved by Academic Council on 31/08/2024

The M.Tech Computer Science and Engineering curriculum is meticulously drafted to cultivate industry-ready professionals endowed with creativity and innovative thinking. This comprehensive curriculum encompasses various components, including course work, miniproject, lab and dissertation work as specified for the programme. The curriculum is so drawn up that the minimum number of credits for successful completion of the M. Tech programme is 68. The curriculum ensures a holistic education that prepares students for the dynamic field of Computer Science and Engineering. Below is a detailed overview of the curriculum:

Core courses (Discipline core courses and Programme core courses)

Elective courses (Programme electives and Interdisciplinary electives)

Audit course

Research Methodology & IPR

Mini project

Laboratory work

Dissertation/Research work

This curriculum is designed to seamlessly blend theoretical knowledge with practical experience and enhance employability through hands-on projects and internships, thereby preparing students for successful careers in Computer Science and Engineering.

Table 1: *Distribution of credits among the Semesters*

Sem	Course work content	Total credits allotted	Credits allotted semester-wise
I	Core courses: 3 nos	3x3 = 9	18
	Programme electives: 2 nos	2x3 = 6	
	Laboratory: 1 no	1x1 = 1	
	Research Methodology & IPR: 1 no	1x2 = 2	
II	Core courses: 2 nos	2x3 = 6	18
	Industry/Interdisciplinary Elective 1 no	1 x3 =3	
	Programme electives: 2 nos	2x3 = 6	
	Laboratory: 1 no	1x1 = 1	
	Miniproject: 1 no	1x2 = 2	
III	MOOC: 1 no	1x2 = 2	16
	Internship: 1 no	1x3 = 3	
	Audit course: 1 no	No credit	
	Phase 1: Dissertation/Research Project: 1 no	1x11 = 11	
IV	Phase 2: Dissertation/Research Project: 1 no	1x16 = 16	16
Total credits in all four semesters			**68**

SEMESTER I

SLOT	COURSE CODE	COURSE NAME	MARKS		L-T-P	HOURS	CREDIT
			CIA	ESE			
A	241TCS100	ADVANCED MACHINE LEARNING	40	60	3-0-0	3	3
B	241TCS001	ADVANCED DATABASE MANAGEMENT	40	60	3-0-0	3	3
C	241TCS002	FOUNDATIONS OF COMPUTER SCIENCE	40	60	3-0-0	3	3
D	241ECSXXX	PROGRAM ELECTIVE 1	40	60	3-0-0	3	3
E	241ECSXXX	PROGRAM ELECTIVE 2	40	60	3-0-0	3	3
S	241RGE100	RESEARCH METHODOLOGY AND IPR	40	60	2-0-0	2	2
T	241LCS100	COMPUTING LAB I	100	--	0-0-2	2	1
Total			**340**	**360**		**19**	**18**

- L-T-P: Lecture-Tutorial-Practical
- CIA: Continuous Internal Assessment, ESE: End Semester Examination

PROGRAM ELECTIVE 1

PROGRAM ELECTIVE 1						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
D	1	241ECS100	OBJECT ORIENTED SOFTWARE ENGINEERING	3-0-0	3	3
	2	241ECS001	ADVANCED DATA MINING	3-0-0	3	3
	3	241ECS002	CLOUD COMPUTING	3-0-0	3	3
	4	241ECS003	WEB SERVICES	3-0-0	3	3
	5	241ECS004	COMPUTATIONAL INTELLIGENCE	3-0-0	3	3

	6	241ECS005	AUTOMATED VERIFICATION	3-0-0	3	3

PROGRAM ELECTIVE 2

PROGRAM ELECTIVE 2						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
E	1	241ECS006	ADVANCED COMPUTER NETWORKS	3-0-0	3	3
	2	241ECS007	PATTERN RECOGNITION	3-0-0	3	3
	3	241ECS008	ADVANCED COMPUTER ARCHITECTURE	3-0-0	3	3
	4	241ECS009	NATURAL LANGUAGE PROCESSING AND TEXT MINING	3-0-0	3	3
	5	241ECS010	ADVANCED COMPILER DESIGN	3-0-0	3	3
	6	241ECS011	BIOINFORMATICS	3-0-0	3	3

SEMESTER II

SLOT	**COURSE CODE**	**COURSE NAME**	**MARKS**		**L-T-P**	**HOURS**	**CREDIT**
			CIA	**ESE**			
A	242TCS0100	ADVANCED DATA STRUCTURES AND ALGORITHMS	40	60	3-0-0	3	3
B	242TCS001	ADVANCED OPERATING SYSTEMS	40	60	3-0-0	3	3
C	242ECSXXX	PROGRAM ELECTIVE 3	40	60	3-0-0	3	3
D	242ECSXXX	PROGRAM ELECTIVE 4	40	60	3-0-0	3	3

E	242EEXXXX / 242ECSXXX	INDUSTRY/ INTERDISCIPLINARY ELECTIVE	40	60	3-0-0	3	3
S	242PCS100	MINI PROJECT	100	--	0-0-4	4	2
T	242LCS100	COMPUTING LAB 2	100	--	0-0-2	2	1
Total			**400**	**300**		**21**	**18**

PROGRAM ELECTIVE 3

PROGRAM ELECTIVE 3						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
C	1	242ECS100	BIG DATA ANALYTICS	3-0-0	3	3
	2	242ECS001	WIRELESS SENSOR NETWORKS	3-0-0	3	3
	3	242ECS002	DEEP LEARNING	3-0-0	3	3
	4	242ECS003	COMPUTER VISION	3-0-0	3	3
	5	242ECS004	SEMANTIC WEB ARCHITECTURE	3-0-0	3	3
	6	242ECS005	PROGRAM ANALYSIS	3-0-0	3	3

PROGRAM ELECTIVE 4

PROGRAM ELECTIVE 4						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
D	1	242ECS006	BLOCKCHAIN TECHNOLOGY AND IOT	3-0-0	3	3
	2	242ECS007	SOCIAL NETWORK ANALYSIS	3-0-0	3	3
	3	242ECS008	MODERN DATABASE MANAGEMENT	3-0-0	3	3
	4	242ECS009	DISTRIBUTED ALGORITHMS	3-0-0	3	3

	5	242ECS010	CYBER FORENSICS AND INFORMATION SECURITY	3-0-0	3	3
	6	242ECS011	SOFTWARE TESTING	3-0-0	3	3

INTERDISCIPLINARY ELECTIVE

INTERDISCIPLINARY ELECTIVE						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
E	1	242ECS056	INTRODUCTION TO MACHINE LEARNING	3-0-0	3	3
	2	242ECS057	DATA STRUCTURES	3-0-0	3	3
	3	242ECS058	SOFTWARE PROJECT MANAGEMENT	3-0-0	3	3

INDUSTRY ELECTIVE

INDUSTRY ELECTIVE						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
E	1	242EEC086	INDUSTRIAL INTERNET OF THINGS	3-0-0	3	3
	2	242EEC087	INFORMATION SECURITY	3-0-0	3	3
	3	242EEC088	INDUSTRIAL AUTOMATION SYSTEM DESIGN	3-0-0	3	3

SEMESTER III

	COURSE CODE	COURSE NAME	L-T-P	HOURS	CREDIT

TRACK 1					
A	243MEC000	MOOC	--	--	2
B	243AGEXXX	AUDIT COURSE	3-0-0	3	-
C	243IEC000	INTERNSHIP	--	--	3
D	243PEC000	DISSERTATION PHASE1	0-0-17	17	11
TRACK 2					
A	243MEC000	MOOC	--	--	2
B	243AGEXXX	AUDIT COURSE	3-0-0	3	-
C	243IEC000	INTERNSHIP	---	--	3
D	243PEC001	RESEARCH PROJECT PHASE 1	0-0-17	17	11
Total				**20**	**16**

AUDIT COURSE

AUDIT COURSE						
SLOT	**SL NO**	**COURSE CODE**	**COURSE NAME**	**L-T-P**	**HOURS**	**CREDIT**
	1	243AGE000	ACADEMIC WRITING	3-0-0	3	3
	2	243AGE001	ADVANCED ENGINEERINGMATERIALS	3-0-0	3	3
	3	243AGE002	FORENSIC ENGINEERING	3-0-0	3	3
	4	243AGE003	DATA SCIENCE FOR ENGINEERS	3-0-0	3	3
	5	243AGE004	DESIGN THINKING	3-0-0	3	3

	6	243AGE005	FUNCTIONAL PROGRAMMING IN HASKELL	3-0-0	3	3
	7	243AGE006	FRENCH LANGUAGE (A1LEVEL)	3-0-0	3	3
	8	243AGE007	GERMAN LANGUAGE (A1LEVEL)	3-0-0	3	3
	9	243AGE008	JAPANESE LANGUAGE (N5LEVEL)	3-0-0	3	3
	10	243AGE009	PRINCIPLES OF AUTOMATION	3-0-0	3	3
	11	243AGE010	REUSE AND RECYCLE TECHNOLOGY	3-0-0	3	3
	12	243AGE011	SYSTEM MODELING	3-0-0	3	3
	13	243AGE012	EXPERT SYSTEMS	3-0-0	3	3

SEMESTER IV

SLOT	COURSE CODE	COURSE NAME	L-T-P	HOURS	CREDIT
TRACK 1					
A	244PEC000	DISSERTATION PHASE II	0-0-22	22	16
TRACK 2					
A	244PEC001	RESEARCH PROJECT PHASE II	0-0-22	22	16
Total				22	16

SEMESTER-I SYLLABUS

241TCS100	ADVANCED MACHINE LEARNING	CATEGORY	L	T	P	CREDIT
		CORE	3	0	0	3

Preamble: This course introduces machine learning concepts and popular machine learning algorithms. It will cover the standard and most popular supervised learning algorithms including linear regression, logistic regression, decision trees, k-nearest neighbour, an introduction to Bayesian learning and the naive Bayes algorithm, support vector machines and kernels and basic clustering algorithms. Dimensionality reduction methods and some applications to real world problems will also be discussed. It helps the learners to develop application machine learning based solutions for real world applications.

Course Outcomes:

After the completion of the course the student will be able to:*

CO 1	Analyse the Machine Learning concepts, classifications of Machine Learning algorithms and basic parameter estimation methods. **(Cognitive Knowledge Level: Analyse)**
CO 2	Illustrate the concepts of regression and classification techniques **(Cognitive Knowledge Level: Apply)**
CO 3	Describe unsupervised learning concepts and dimensionality reduction techniques. (**Cognitive Knowledge Level: Apply**)
CO 4	Explain Support Vector Machine concepts and graphical models.**(Cognitive Knowledge Level: Apply)**
CO 5	Choose suitable model parameters for different machine learning techniques and to evaluate a model performance. **(Cognitive Knowledge Level: Apply)**
CO6	Design, implement and analyse machine learning solution for a real world problem. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓		✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	60-80%
Analyse	20-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks

Course based task/Seminar/Quiz : 10 marks

Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted.
Test paper shall include minimum 80% of the syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.
Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.
Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Suppose that X is a discrete random variable with the following probability mass function: where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations were taken from such a distribution: ***(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)***. What is the maximum likelihood estimate of θ.

X	0	1	2	3
$P(X)$	$2\theta/3$	$\theta/3$	$2(1-\theta)/3$	$(1-\theta)/3$

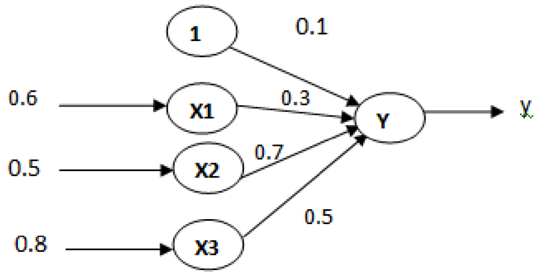
2. What is the difference between Maximum Likelihood estimation (MLE) and Maximum a Posteriori (MAP) estimation?
3. A gamma distribution with parameters α, β has the following density function, where $\Gamma(t)$ is the gamma function.

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

If the posterior distribution is in the same family as the prior distribution, then we say that the prior distribution is the conjugate prior for the likelihood function. Using the Gamma distribution as a prior, show that the Exponential distribution is a conjugate prior of the Gamma distribution. Also, find the maximum a posteriori estimator for the parameter of the Exponential distribution as a function of α and β.

Course Outcome 2 (CO2) :

1. How can we interpret the output of a two-class logistic regression classifier as a robability?
2. Calculate the output of the following neuron Y if the activation function is a binary sigmoid.



3. Suppose you have a 3-dimensional input x = (x1, x2, x3) = (2, 2, 1) fully connected with weights (0.5, 0.3, 0.2) to one neuron which is in the hidden layer with sigmoid activation function. Calculate the output of the hidden layer neuron.
4. Consider the case of the XOR function in which the two points {(0, 0),(1, 1)} belong to one class, and the other two points {(1, 0),(0, 1)} belong to the other class. Design a multilayer perceptron for this binary classification problem.
5. Why does a single perceptron cannot simulate simple XOR function? Explain how this limitation is overcome?
6. Consider a naive Bayes classifier with 3 boolean input variables, **X1**, **X2** and **X3**, and one boolean output, **Y**. How many parameters must be estimated to train such a naive Bayes classifier? How many parameters would have to be estimated to learn the above classifier if we do not make the naive Bayes conditional independence assumption?

Course Outcome 3(CO3):

1. Describe the basic operation of k-means clustering.
2. A Poisson distribution is used to model data that consists of non-negative integers. Suppose you observe m integers in your training set. Your model assumption is that each integer is sampled from one of two different Gaussian distributions. You would like to learn this model using the EM algorithm. List all the parameters of the model. Derive the E-step and M-step for this model.
3. A uni-variate Gaussian distribution is used to model data that consists of non-negative integers. Suppose you observe m integers in your training set. Your model assumption is that each integer is sampled from one of two different Gaussian distributions. You would like to learn this model using the EM algorithm. List all the parameters of the model. Derive the E-step and M-step for the model.
4. Suppose you want to cluster the eight points shown below using **k**-means

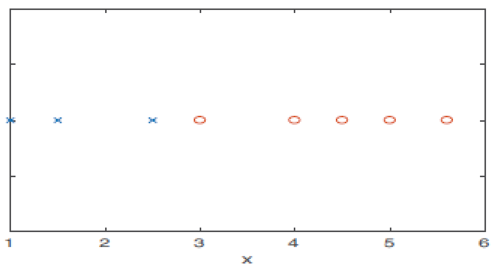
	A_1	A_2
x_1	2	10
x_2	2	5
x_3	8	4
x_4	5	8
x_5	7	5
x_6	6	4
x_7	1	2
x_8	4	9

Assume that **k = 3** and that initially the points are assigned to clusters as follows:

C1 = {x1, x2, x3}, C2 = {x4, x5, x6}, C3 = {x7, x8}. Apply the **k**-means algorithm until convergence, using the Manhattan distance.

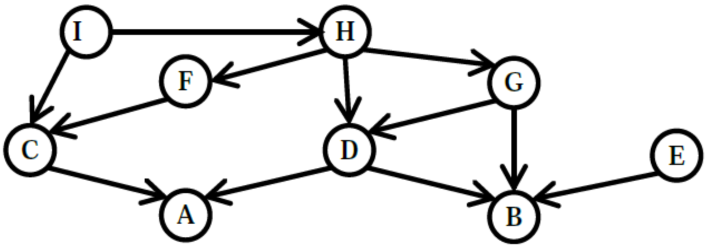
Course Outcome 4 (CO4):

1. Describe how Support Vector Machines can be extended to make use of kernels. Illustrate with reference to the Gaussian kernel $K(x, y) = e^{-y}$, where $y = (x-y)^2$.
2. Suppose that you have a linear support vector machine(SVM) binary classifier. Consider a point that is currently classified correctly, and is far away from the decision boundary. If you remove the point from the training set, and re-train the classifier, will the decision boundary change or stay the same? Justify your answer.
3. What is the primary motivation for using the kernel trick in machine learning algorithms?
4. Show that the Boolean function $(x_1 \wedge x_2) \vee (\neg x_1 \wedge \neg x_2)$ is not linearly separable (i.e. there is no linear classifier $sign(w_1\, x_1 + w_2\, x_2 + b)$ that classifies all 4 possible input points correctly). Assume that "true" is represented by 1 and "false" is represented by −1. Show that there is a linear separator for this Boolean function when we use the kernel $K(x, y) = (x \cdot y)^2$ (***x.y*** denotes the ordinary inner product) . Give the weights and the value of ***b*** for one such separator.
5. Consider the following one dimensional training data set, 'x' denotes negative examples and 'o' positive examples. The exact data points and their labels are given in the table. Suppose a SVM is used to classify this data. Indicate which are the support vectors and mark the decision boundary. Give the value of the cost function and of the model parameters after training.



x	1	1.5	2.5	3	4	4.5	5	5.6
y	-1	-1	-1	1	1	1	1	1

6. Write down the factored conditional probability expression that corresponds to the graphical Bayesian Network shown below.



7. How do we learn the conditional probability tables(CPT) in Bayesian networks if information about some variables is missing? How are these variables called?

Course Outcome 5 (CO5):

1. Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the accuracy, precision and recall for the data.
2. Given the following data, construct the ROC curve of the data. Compute the AUC.

Threshold	TP	TN	FP	FN
1	0	25	0	29
2	7	25	0	22
3	18	24	1	11
4	26	20	5	3
5	29	11	14	0
6	29	0	25	0
7	29	0	25	0

3. With an example classification problem, explain the following terms: a) Hyper parameters b) Training set c) Validation sets d) Bias e) Variance.
4. What is ensemble learning? Can ensemble learning using linear classifiers learn classification of linearly non-separable sets?
5. Describe boosting. What is the relation between boosting and ensemble learning?
6. Classifier A attains 100% accuracy on the training set and 70% accuracy on the test set. Classifier B attains 70% accuracy on the training set and 75% accuracy on the test set. Which one is a better

classifier. Justify your answer.

7. What are ROC space and ROC curve in machine learning? In ROC space, which points correspond to perfect prediction, always positive prediction and always negative prediction? Why?
8. Suppose there are three classifiers A,B and C. The (FPR, TPR) measures of the three classifiers are as follows – A (0, 1), B (1, 1) , C (1,0.5). Which can be considered as a perfect classifier? Justify your answer.
9. What does it mean for a classifier to have a high precision but low recall?

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241TCS100

Course Name: ADVANCED MACHINE LEARNING

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Explain the principle of the gradient descent algorithm.
2. In a two-class logistic regression model, the weight vector **w = [4, 3, 2, 1, 0]**. We apply it to some object that we would like to classify; the vectorized feature representation of this object is **x = [−2, 0,−3, 0.5, 3]**. What is the probability, according to the model, that this instance belongs to the positive class?
3. Expectation maximization (EM) is designed to find a maximum likelihood setting of the parameters of model when some of the data is missing. Does the algorithm converge? If so, do you obtain a locally or globally optimal set of parameters?
4. What is the basic idea of a Support Vector Machine?
5. What is the trade-off between bias and variance? **(5x5=25)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. Suppose $x_1, ..., x_n$ are independent and identically distributed(iid) samples from a distribution with density **(7)**

$$f_X(x \mid \theta) = \begin{cases} \frac{\theta x^{\theta-1}}{3^\theta}, & 0 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$$

Find the maximum likelihood estimate(MLE) for θ.

7. Derive the gradient descent training rule assuming for the target function $o_d = w_0 + w_1x_1 + ... + w_nx_n$. Define explicitly the squared cost/error function E, assuming that a set of training examples D is provided, where each training example $d \varepsilon D$ is associated with the target output t_d. **(7)**

8. Cluster the following eight points representing locations into three clusters: ***A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)***. **(7)**

Initial cluster centers are: ***A1(2, 10), A4(5, 8)*** and ***A7(1, 2)***.

The distance function between two points $a = (x1, y1)$ and $b = (x2, y2)$ is defined as $D(a, b) = |x2 - x1| + |y2 - y1|$

Use **k**-Means Algorithm to find the three cluster centers after the second iteration.

9. Describe Principal Component Analysis. What criterion does the method minimize? What is the objective of the method? Give a way to compute the solution from a matrix X encoding the features. **(7)**

10. Consider a support vector machine whose input space is 2-D, and the inner products are computed by means of the kernel $K(x, y) = (x.y + 1)^2 - 1$ ($x.y$ denotes the ordinary inner product). Show that the mapping to feature space that is implicitly defined by this kernel is the mapping to 5-D given by **(7)**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \phi(\mathbf{x}) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}\, x_1 x_2 \\ \sqrt{2}\, x_1 \\ \sqrt{2}\, x_2 \end{bmatrix}.$$

11. How does random forest classifier work? Why is a random forest better than a decision tree? **(7)**

12. Consider a two-class classification problem of predicting whether a photograph contains a man or a woman. Suppose we have a test dataset of 10 records with expected outcomes and a set of predictions from our classification algorithm. Compute the confusion matrix, accuracy, precision, recall, sensitivity and specificity on the following data. **(7)**

Sl.No.	Actual	Predicted
1	man	woman
2	man	man
3	woman	woman
4	man	man
5	man	woman
6	woman	woman
7	woman	man
8	man	man
9	man	woman
10	woman	woman

Syllabus

Module-1 (Parameter Estimation and Regression) 8 hours

Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning. Basics of parameter estimation: Maximum Likelihood Estimation(MLE), Maximum a Posteriori Estimation (MAP). Gradient Descent Algorithm, Batch Gradient Descent, Stochastic Gradient Descent. Regression algorithms: least squares linear regression, normal equations and closed form solution, Polynomial regression.

Module-2 (Regularization techniques and Classification algorithms) 9 hours

Overfitting, Regularization techniques - LASSO and RIDGE. Classification algorithms: linear and non-linear algorithms, Perceptrons, Logistic regression, Naive Bayes, Decision trees. Neural networks : Concept of Artificial neuron, Feed-Forward Neural Network, Back propagation algorithm.

Module-3 (Unsupervised learning) 8 hours

Unsupervised learning: clustering, k-means, Hierarchical clustering, Principal component analysis, Density-based spatial clustering of applications with noise (DBSCAN). Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model.

Module-4 (Support Vector Machine and Graphical Models) 7 hours

Support vector machines and kernels : Max margin classification, Nonlinear SVM and the kernel trick, nonlinear decision boundaries, Kernel functions. Basics of graphical models - Bayesian networks, Hidden Markov model - Inference and estimation.

Module-5 (Evaluation Metrics and Sampling Methods) 8 hours

Classification Performance Evaluation Metrics: Accuracy, Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC. Regression Performance Evaluation Metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination. Clustering Performance Evaluation Metrics: Purity, Jaccard index, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn's Index. Boosting: AdaBoost, gradient boosting machines. Resampling methods: cross-validation, bootstrap. Ensemble methods: bagging, boosting, random forests Practical aspects in machine learning: data preprocessing, overfitting, accuracy estimation, parameter and model selection Bias-Variance tradeoff

Course Plan

No	**Topics**	**No. of Lectures (40)**
1	**Module-1 (Parameter Estimation and Regression) 8 hours**	
1.1	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning.	1
1.2	Basics of parameter estimation: Maximum Likelihood Estimation(MLE)	1
1.3	Basics of parameter estimation: Maximum Likelihood Estimation(MLE) - Examples	1
1.4	Basics of parameter estimation: Maximum a Posteriori Estimation (MAP)	1
1.5	Basics of parameter estimation: Maximum a Posteriori Estimation (MAP) - Example	1
1.6	Gradient Descent Algorithm, Batch Gradient Descent, Stochastic Gradient Descent	1
1.7	Regression algorithms: least squares linear regression, normal equations and closed form solution	1

1.8	Polynomial regression	1
2	**Module-2 (Regularization techniques and Classification algorithms) 9 hours**	
2.1	Overfitting, Regularization techniques - LASSO and RIDGE	
2.2	Classification algorithms: linear and non-linear algorithms	
2.3	Perceptrons	
2.4	Logistic regression	
2.5	Naive Bayes	
2.6	Decision trees	
2.7	Neural networks : Concept of Artificial neuron	
2.8	Feed-Forward Neural Network	
2.9	Back propagation algorithm	
3	**Module-3 (Unsupervised learning) 8 hours**	
3.1	Unsupervised learning: clustering, k-means	
3.2	Hierarchical clustering	
3.3	Principal component analysis	
3.4	Density-based spatial clustering of applications with noise (DBSCAN)	
3.5	Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model	
3.6	Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model	
4	**Module-4 (Support Vector Machine and Graphical Models) 7 hours**	
4.1	Support vector machines and kernels : Max margin classification	
4.2	Support vector machines: Max margin classification	
4.3	Nonlinear SVM and the kernel trick, nonlinear decision boundaries	
4.3	Kernel functions	
4.5	Basics of graphical models - Bayesian networks	
4.6	Hidden Markov model - Inference and estimation	
4.7	Hidden Markov model - Inference and estimation	
4.8	Hidden Markov model - Inference and estimation	
5	**Module-5 (Evaluation Metrics and Sampling Methods) 8 hours**	

5.1	Classification Performance Evaluation Metrics: Accuracy, Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC	
5.2	Regression Performance Evaluation Metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination	
5.3	Clustering Performance Evaluation Metrics: Purity, Jaccard index, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn's Index	
5.4	Boosting: AdaBoost, gradient boosting machines.	
5.5	Resampling methods: cross-validation, bootstrap.	
5.6	Ensemble methods: bagging, boosting, random forests	
5.7	Practical aspects in machine learning: data preprocessing, overfitting, accuracy estimation, parameter and model selection	
5.8	Bias-Variance tradeoff	

Reference Books

1. Christopher Bishop. Neural Networks for Pattern Recognition, Oxford University Press, 1995.
2. Kevin P. Murphy. Machine Learning: A Probabilistic Perspective, MIT Press 2012.
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements Of Statistical Learning, Second edition Springer 2007.
4. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition, MIT Press 2010.
5. Tom Mitchell, Machine Learning, McGraw-Hill, 1997.

241TCS001	ADVANCED DATABASE MANAGEMENT	CATEGORY	L	T	P	CREDIT
		PCC	3	0	0	3

Preamble: This course provides an exposure to the concepts and techniques in advanced database management. Various strategies regarding query processing and optimization are discussed in this curriculum. An optimum insight of database security is provided. Different layouts of database system architecture and distributed system architecture, along with semi-structured data is included for better understanding of advanced data management. This course helps the learners to develop applications that manage data efficiently with the help of suitable data models and techniques.

Course Outcomes: After the completion of the course the student will be able to:

CO 1	Identify various measures of query processing and optimization. **(Cognitive Knowledge Level: Apply)**
CO 2	Analyze and implement security mechanisms to secure a database system. **(Cognitive Knowledge Level: Analyze)**
CO 3	Apply knowledge and awareness of the different database architectures in different scenarios. **(Cognitive Knowledge Level: Apply)**
CO 4	Analyze implementation aspects of distributed system on database architecture. **(Cognitive Knowledge Level: Analyze)**
CO 5	Make use of semi structured data, XML and XML queries for data management. **(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, and Implement innovative ideas on advanced database concepts and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓		✓	
CO 2	✓		✓	✓		✓	
CO 3	✓		✓	✓		✓	
CO 4	✓		✓	✓		✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks

Course based task/Seminar/Quiz : 10 marks

Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted.

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.
Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.
Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Suppose you want to get answers to r ⋈ s sorted on an attribute of r, and want only the top K answers for some relatively small K. Give a good way of evaluating the query:
   i. When the join is on a foreign key of r referencing s, where the foreign key attribute is declared to be not null.
   ii. When the join is not on a foreign key.
2. Why is it not desirable to force users to make an explicit choice of a query processing strategy? Are there cases in which it is desirable for users to be aware of the costs of competing query-processing strategies? Explain your answer.

Course Outcome 2 (CO2):

1. A database relation may have the values of certain attributes encrypted for security. Why do database systems not support indexing on encrypted attributes? Using your answer to this question, explain why database systems do not allow encryption of primary-key attributes.

Course Outcome 3(CO3):

1. Suppose relation r is stored partitioned and indexed on A, and s is stored partitioned and indexed on B. Consider the query: $_{r.C}\gamma_{count(s.D)}((\sigma_{A>5}(r)) \bowtie_{r.B=s.B} s)$
   i Give a parallel query plan using the exchange operator, for computing the subtree of the query involving only the select and join operators.
   ii Now extend the above to compute the aggregate. Make sure to use preaggregation to minimize the data transfer.
2. If a parallel data-store is used to store two relations *r* and *s* and we need to join *r* and *s*, it may be useful to maintain the join as a materialized view. What are the benefits and overheads in terms of overall throughput, use of space, and response time to user queries? Explain in detail.

Course Outcome 4 (CO4):

1. Insert and query on a Bloom filter of size m = 10 and number of hash functions k= 3. Let H(x) denote the result of the three hash functions which will write as a set of three values {h 1 (x), h 2 (x), h 3 (x)}.Has functions used: A= x mod 10, B= x mod 7, C= (sum of digits) mod 9.
2. Assume a relationship R_{AB} at site 1 and relationship S_{CD} at site 2 as follows: R= {(1,2), (3,4), (5, 6), (7, 8), (9, 10)}.S= {(1, 0), (8,1), (9, 2), (10, 3), (11, 4)}. Compute R⋈ S using bloom join with A=C and explain the intermediate steps. Show the tuples transferred with the hash function mod 4.

Course Outcome 5 (CO5):

1. Design an XML document for storing hostel mess food details (meals taken such as breakfast, lunch, dinner) with their charges for the month of June 2022. Charges may vary depending on the food taken. Students can opt not to take any meals on certain days.
   - i Write a sample XML for 2 students for 2 days.
   - ii Write a XQuery to return the lunch details of all.
   - iii Create an XSD for the same.

Course Outcome 6 (CO6):

1. Implement Student book finder application using XML.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241TCS001

Course Name: ADVANCED DATABASE MANAGEMENT

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Why is it not desirable to force users to make an explicit choice of a query-processing strategy? Are there cases in which it is desirable for users to be aware of the costs of competing query-processing strategies? Explain your answer.	**(5)**
2.	What are the relative merits of using Discretionary Access Control or Mandatory Access Control? What is role-based access control? In what ways is it superior to DAC and MAC?	**(5)**

3. Suppose relation r is stored partitioned and indexed on A, and s is stored partitioned and indexed on B. Consider the query: ${}_{r.C}\gamma_{count(s.D)}(\,(\sigma_{A>5}(r)) \bowtie_{r.B=s.B} s\,)$ **(5)**

i Give a parallel query plan using the exchange operator, for computing the subtree of the query involving only the select and join operators.

ii Now extend the above to compute the aggregate. Make sure to use preaggregation to minimize the data transfer.

4. Insert and query on a Bloom filter of size m = 10 and number of hash functions k= 3. Let H(x) denote the result of the three hash functions which will write as a set of three values {h 1 (x), h 2 (x), h 3 (x)} Has functions used: A= x mod 10, B= x mod 7, C= (sum of digits) mod 9. **(5)**

5. Design an XML document for storing hostel mess food details (meals taken such as breakfast, lunch, dinner) with their charges for the month of June 2022. Charges may vary depending on the food taken. Students can opt not to take any meals on certain days. **(5)**

i Write a sample XML for 2 students for 2 days.

ii Write a XQuery to return the lunch details of all.

iii Create an XSD for the same.

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) Consider the issue of interesting orders in optimization. Suppose you are given a query that computes the natural join of a set of relations S. Given a subset S1 of S, what are the interesting orders of S1? **(4)**

(b) Suppose you want to get answers to r ⋈ s sorted on an attribute of r, and want only the top K answers for some relatively small K. Give a good way of evaluating the query: **(3)**

i When the join is on a foreign key of r referencing s, where the foreign key attribute is declared to be not null.

ii When the join is not on a foreign key.

7. A database relation may have the values of certain attributes encrypted for security. Why do database systems not support indexing on encrypted attributes? Using your answer to this question, explain why database systems do not allow encryption of primary-key attributes. **(7)**

8. If a parallel data-store is used to store two relations *r* and *s* and we need to join *r* and *s*, it may be useful to maintain the join as a materialized view. What are the benefits and overheads in terms of overall throughput, use of space, and response time to user queries? Explain in detail. **(7)**

9. Consider the bitmap representation of the free-space map, where for each block in the file, two bits are maintained in the bitmap. If the block is between 0 and 30 percent full the bits are 00, between 30 and 60 percent the bits are 01, between 60 and 90 percent the bits are 10, and above 90 percent the bits are 11. Such bitmaps can be kept in memory even for quite large files. **(7)**

i Outline two benefits and one drawback to using two bits for a block, instead of one byte as described earlier in this chapter.

ii Describe how to keep the bitmap up to date on record insertions and deletions.

iii Outline the benefit of the bitmap technique over free lists in searching for free space and in updating free space information.

10. Assume a relationship R_{AB} at site 1 and relationship S_{CD} at site 2 as follows: R= {(1,2), (3,4), (5, 6), (7, 8), (9, 10)} **(7)**

S= {(1, 0), (8,1), (9, 2), (10, 3), (11, 4)}

Compute R⋈ S using bloom join with A=C and explain the intermediate steps. Show the tuples transferred with the hash function mod 4.

11. (a) Consider the country data. **(4)**

Write XPath for the following:

i Return the area of India.

ii Return the names of all countries with population greater than 100 million.

iii Return the names of all countries whose population is less than one thousandth that of some city (in any country).

iv Return the names of all cities that have the same name as the country in which they are located.

(b) Consider the country data. **(3)**

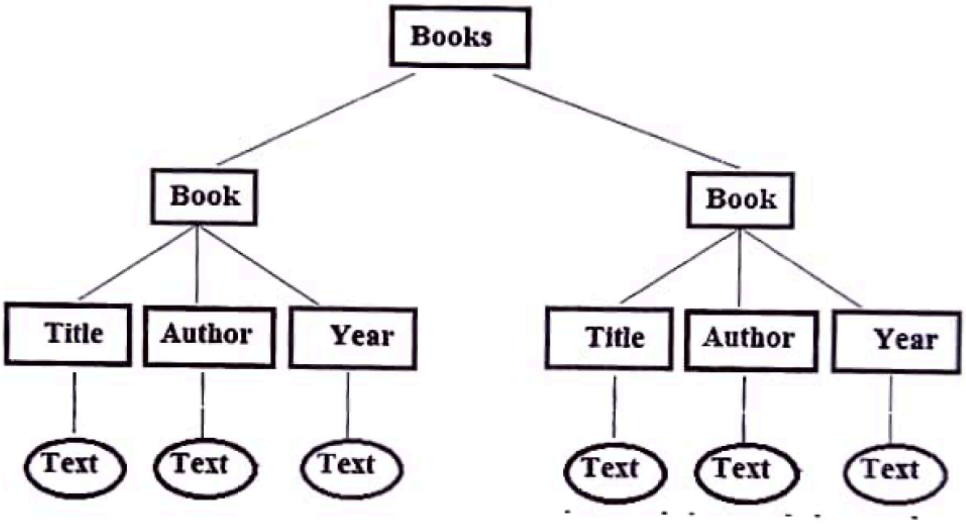
Write XQuery for the following:

i Return the name of the country with the highest population.

ii Return the name of the country that has the city with the highest population.

iii Return the average population of Russian-speaking countries.

12. Consider the following XML Tree **(7)**



Write an XML schema for the above, and also provide an XQuery expression to get the books published in the year 1992.

Syllabus

Module 1: Query Processing and Optimization

Review of indexing and Hashing - Overview- Measures of query cost- Algorithms for Selection and Join with cost analysis- Evaluation of expressions- Optimization of RA expressions.

Module 2: Database Security

Threats to databases, control measures, database security and DBA, Discretionary access control, Mandatory access control (role-based only), SQL injection.

Module 3: Database System Architectures

Centralized and Client-Server Architectures – Centralized server systems - Server System Architectures - Parallel Systems- - Parallel storage - Data partitioning, replication and indexing in Parallel Databases-Parallel query processing.

Module 4: Distributed System Architecture

Distributed System architecture- Distributed storage - Distributed file systems – Distributed RDB design-Transparency– Distributed Transactions - Commit Protocols – Concurrency Control - Distributed Query Processing
Advanced indexing Techniques: Bloom filter - Bitmap indices - Indexing spatial data - Hash indices.

Module 5: Semi-structured Data

Semi-structured Data and XML Databases: XML Data Model – XSD – XPath and XQuery – Example Queries. Native XML databases, Object Relational Systems

Course Plan

No	Topic	No. of Lectures (**40 Hours**)
1	**Module 1: Query Processing and Optimization**	**8**
1.1	Introduction to Query Processing and Optimization	1
1.2	Review of indexing	1
1.3	Hashing - Overview	1
1.4	Measures of query cost	1
1.5	Algorithms for Selection with cost analysis	1
1.6	Algorithms for Join with cost analysis	1
1.7	Evaluation of expressions	1
1.8	Optimization of RA expressions	1
2	**Module 2: Database Security**	**7**
2.1	Introduction to Database Security	1
2.2	Threats to databases	1
2.3	control measures	1
2.4	database security and DBA	1
2.5	Discretionary access control	1
2.6	Mandatory access control (role-based only)	1
2.7	SQL injection	1
3	**Module 3: Database System Architectures**	**9**
3.1	Introduction to Database System Architectures	1
3.2	Overview of Centralized and Client-Server Architectures	1
3.3	Centralized server systems	1
3.4	Server System Architectures	1
3.5	Parallel Systems	1
3.6	Parallel storage	1
3.7	Data partitioning, replication in Parallel Databases	1
3.8	Indexing in Parallel Databases	1
3.9	Parallel query processing.	1
4	**Module 4: Distributed System Architecture**	**10**
4.1	Introduction to Distributed System architecture	1
4.2	Distributed storage & Distributed file systems	1

4.3	Distributed RDB design & its Transparency	1
4.4	Distributed Transactions	1
4.5	Commit Protocols & Concurrency Control	1
4.6	Distributed Query Processing	1
4.7	Advanced indexing Techniques: Bloom filter	1
4.8	Bitmap indices	1
4.9	Indexing spatial data	1
4.10	Hash indices	1
5	**Module 5: Semi-structured Data**	**6**
5.1	Introduction to Semi-structured Data and XML Databases	1
5.2	XML Data Model – XSD	1
5.3	XPath and XQuery	1
5.4	Example Queries	1
5.5	Native XML databases	1
5.6	Object Relational Systems	1

References

1. R. Elmasri, S.B. Navathe, "Fundamentals of Database Systems", 7/e, Pearson Education/Addison Wesley, 2016
2. Thomas Cannolly and Carolyn Begg, "Database Systems, A Practical Approach to Design, Implementation and Management", 3/e, Pearson Education, 2010.
3. Henry F Korth, Abraham Silberschatz, S. Sudharshan, "Database System Concepts", 7/e, Tata McGraw Hill, 2019.
4. Joe Fawcett, Danny Ayers, Liam R. E. Quin, Beginning XML, 5/e, John Wiley & Sons, 2012
5. Grigoris Antoniou. Frank van Harmelen, "A Semantic Web Primer", The MIT Press,Cambridge, Massachusetts, 2003.

241TCS002	**FOUNDATIONS OF COMPUTER SCIENCE**	**CATEGORY**	**L**	**T**	**P**	**CREDIT**
		PCC	**3**	**0**	**0**	**3**

Preamble: The purpose of this course is to develop rigorous proof writing skills which can be used to prove different theorems and results in Computer Science and its applications. This course helps to understand and apply the elementary and advanced Counting Principles in solving various computational problems. Also the course helps the learners to solve problems on probability and also to understand a few classic probability problems.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply Direct proof technique, Indirect proof technique and Mathematical Induction to prove various theorems and results. (**(Cognitive Knowledge Level: Apply)**
CO 2	Solve counting problems using Pigeon hole principle, Principle of Inclusion exclusion, Permutations, Combinations, Cantor's Diagonalization argument and Derangements. (**(Cognitive Knowledge Level: Apply)**
CO 3	Solve Recurrence relations and counting problems using Generating Functions. **(Cognitive Knowledge Level: Apply)**
CO 4	Solve problems on probability using the fundamentals of Probability, Bayes theorem, and Probability Distributions. **(Cognitive Knowledge Level: Apply)**
CO 5	Solve problems using concepts in algebraic structures such as Groups, Cosets and Lagrange's Theorem. **(Cognitive Knowledge Level: Apply)**
CO 6	Design solutions for various computational problems using the mathematical concepts of computer science and prove the correctness of the solution developed. **(Cognitive Knowledge Level: Evaluate)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓		✓	
CO 2	✓		✓	✓		✓	

CO 3	✔		✔	✔		✔	
CO 4	✔		✔	✔		✔	
CO 5	✔		✔	✔		✔	
CO 6	✔	✔	✔	✔	✔	✔	

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks

Course based task/Seminar/Quiz : 10 marks

Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted.

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. For a $\in$ Z, if $a^2 - 2a + 7$ is even, then a is odd. Prove the statement using contradiction and contrapositive proof techniques.
2. Assume that in a group of 6 people, each pair of individuals consists of 2 friends or 2 enemies. Show that there are either 3 mutual friends or 3 mutual enemies in the group. Every pair of people at the party are either friends or enemies.
3. Using the principle of mathematical induction prove that $6^{n+2} + 7^{2n+1}$ is divisible by 43 for n $\in Z^+$.

Course Outcome 2 (CO2):

1. How many different words can be formed from the letters of the word 'EXTRA" so that the vowels are never together?

2. Suppose repetitions are not allowed:

   (a) How many 3 digit numbers can be formed from the 6 digits 2,3,5,6,7 and 9?
   (b) How many of these numbers are less than 400?
   (c) How many are even?

Course Outcome 3(CO3):

1. Solve the recurrence relation using generating function: $a_r - 7a_{r-1} + 10a_{r-2} = 0$ with initial conditions $a_0=3$ and $a_1=3$.
2. Find the coefficient of x^{17} in the expansion of $(1 + x^5 + x^7)^{20}$.
3. Find the number of solutions of $e_1 + e_2 + e_3 = 17$, where e_1 , e_2 , and e_3 are nonnegative integers with $2 \le e_1 \le 5$, $3 \le e_2 \le 6$ and $4 \le e_3 \le 7$.

Course Outcome 4 (CO4):

1. Two marbles are drawn successively from a box of 3 black and 4 white marbles.

(i) Find the probability that both are black if the first marble is not replaced before the second drawing?
(ii) Find the probability that both are black if the first marble is replaced before the second drawing?

2. State Bayes Theorem.
3. Find the probability distribution of number of green balls drawn when 3 balls are drawn

   one by one without replacement from a bag containing 3 green and 5 white balls.

Course Outcome 5 (CO5):

1. A group of n men enter a restaurant and check their hats. The hat-checker is absent-minded, and upon leaving, she redistributes the hats back to the men at random. Find the expected number of men who get their own hat.
2. Explain (i) Generators of a group (ii) Left coset and (iii) Homomorphism with an example each.

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 2**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241TCS002

Course Name: FOUNDATIONS OF COMPUTER SCIENCE

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Prove that $\sqrt{3}$ is irrational using proof by contradiction. **(5)**

2. Show that the set of real numbers is uncountable using Cantor's diagonalization principle. **(5)**

3. In how many different ways can eight identical cookies be distributed among three distinct children if each child receives at least two cookies and no more than four cookies? **(5)**

4. A woman has 11 close friends and she wants to invite 5 of them to dinner. In how many ways can she invite them if **(5)**

(i) there is no restriction on the choice.

(ii) two particular persons will not attend separately.

(iii) two particular persons will not attend together.

5. State and prove Birthday Paradox. **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) Determine which amounts of postage can be formed using four and seven rupees stamps. Prove your answer using principle of mathematical induction . **(4)**

(b) Prove your answer using strong induction. **(3)**

7. Every sequence of $(n^2 + 1)$ distinct real numbers contain a sub sequence of length $(n + 1)$ that is either strictly increasing or strictly decreasing. Prove the statement using Pigeonhole principle. **(7)**

8. (a) (i)Find the number of permutations that can be formed from the letters of the string 'ELEVEN'? **(4)**
(ii) How many of them begin and end with E?
(iii) How many of them have three Es together?
(iv) How many begin with E and end with N?

(b) Determine the number of integers between 1 and 10000 that are not divisible by 6, 7 or 8. **(3)**

9. Solve the recurrence relation $a_n - 4a_{n-1} = 6 \times 4^n$ with initial condition $a_0 = 1$, using generating function. **(7)**

10. Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university? Why? **(7)**

11. There are n distinct coupons placed in an urn. Coupons are randomly selected one at a time (with replacement) until at least one of each type of coupon has been selected. Find the expected number of selections made until all distinct n coupons are collected. **(7)**

12. State and prove Lagrange's Theorem. **(7)**

Syllabus

Module 1: Theorem Proving Techniques

Theorem proving techniques: Direct Proof, Indirect proof - Proof by Contrapositive, Proof by contradiction and Proof by exhausting cases, Principle of mathematical induction, Complete induction and Well-ordering principle. The Pigeonhole principle.

Module 2 : Fundamentals of Counting

The Basics of counting, Addition and multiplication principles, Permutations and Combinations. Countable and uncountable sets, Principle of inclusion and exclusion –applications, derangements.

Module 3 : Generating Functions

Recurrence Relations, Modeling problems with recurrence relations. Generating functions, Solving counting problems using Generating functions, Solving recurrence relations using Generating functions.

Module 4: Probability Theory

Probability theory – Properties of Probability, Conditional Probability, Independent Events, Bayes Theorem, Mathematical Expectation and Variance of Random variables.

Discrete Distributions and its mean and variance- Binomial Distribution, Bernoulli Distribution , Geometric Distribution, Poisson Distribution. Continuous Distributions and its mean and variance- Uniform and Exponential Distributions, Normal Distribution.

Module 5 : Classic Problems in Probability and Algebraic Structures

Classic Problems in Probability- Birthday Paradox, The Hat Problem, Coupon Collector Problem.
Groups and subgroups, generators for a group, Homomorphism theorems, cosets and normal subgroups, Lagrange's theorem.

Course Plan

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Theorem Proving Techniques**	**8**
1.1	Theorem proving techniques: Direct Proof	1
1.2	Indirect proof - Proof by Contrapositive,	1
1.3	Proof by contradiction	1
1.4	Principle of mathematical induction, Complete induction	1
1.5	Principle of mathematical induction, Complete induction	1
1.6	Well-ordering principle	1
1.7	The Pigeonhole principle	1
1.8	The Pigeonhole principle	1
2	**Module 2: Fundamentals of Counting**	**7**
2.1	The Basics of counting, Addition and multiplication principles	1

2.2	Permutations and Combinations.	1
2.3	Permutations and Combinations.	1
2.4	Countable and uncountable sets	1
2.5	Countable and uncountable sets	1
2.6	Principle of inclusion and exclusion –applications,	1
2.7	Derangements	1
3	**Module 3: Generating Functions**	**7**
3.1	Recurrence Relations, Modeling problems with recurrence relations.	1
3.2	Generating functions	1
3.3	Generating functions	1
3.4	Solving counting problems using Generating functions	1
3.5	Solving counting problems using Generating functions	1
3.6	Solving recurrence relations using Generating functions.	1
3.7	Solving recurrence relations using Generating functions.	1
4	**Module 4: Probability Theory**	**10**
4.1	Probability theory – Properties of Probability	1
4.2	Conditional Probability	1
4.3	Independent Events	1
4.4	Bayes Theorem	1
4.5	Mathematical Expectation and Variance of Random variables	1
4.6	Discrete Distributions and its mean and variance- Binomial Distribution, Bernoulli Distribution	1
4.7	Geometric Distribution, Poisson Distribution	1
4.8	Continuous Distributions and its mean and variance	1

4.9	Uniform and Exponential Distributions	1
4.10	Normal Distribution	1
5	**Module 5: Classic Problems in Probability and Algebraic Structures**	**8**
5.1	Classic Problems in Probability- Birthday Paradox	1
5.2	The Hat Problem	1
5.3	Coupon Collector Problem.	1
5.4	Coupon Collector Problem.	1
5.5	Groups and subgroups, generators for a group	1
5.6	Homomorphism theorems	1
5.7	Cosets and normal subgroups	1
5.8	Lagrange’s theorem	1

References

1. Kenneth H. Rosen, “Discrete Mathematics and its Applications” 7/e, McGraw Hill Inc, 2011.
2. J. P. Tremblay, R. Manohar, “Discrete Mathematical Structures with Application to Computer Science” Tata McGrawHill, 2000
3. Sheldon M. Ross, “Introduction to Probability Models”
4. Sheldon M. Ross, “A First Course in Probability”
5. William Feller, “An introduction to probability theory and its applications” Volume1.Wiley,1957.
6. Rajeev Motwani and Prabhakar Raghavan, “Randomized Algorithms” Cambridge University Press 1995

PROGRAM ELECTIVE 1

241ECS100	OBJECT ORIENTED SOFTWARE ENGINEERING	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: Study of this course provides the learners an exposure to the concepts and principles of object oriented software engineering. The course covers the various software lifecycle models, principles of design, coding, testing, maintenance and configuration management. The course helps the learners to analyse and design software using tools and will improve capability to efficiently develop, deploy and maintain software.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Make use of project organization and management concepts and analyse the various tasks carried out. **(Cognitive Level: Apply)**
CO 2	Identify and select suitable process model for a given problem. **(Cognitive Level: Apply)**
CO 3	Analyse the requirements of a given software project and produce requirement specification **(Cognitive Level: Analyse).**
CO 4	Examine the various designing principles and patterns of a software product**. (Cognitive Level: Analyse)).**
CO 5	Build the mapping of product design to code, its testing and maintenance. **(Cognitive Level: Apply).**
CO6	Design, analyse object models and dynamic models for a given problem statement. **(Cognitive Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓		
CO 3	✓		✓	✓	✓		
CO 4	✓		✓	✓	✓		
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyse	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Detail about the Project Organization Concepts.
2. Distinguish between planned and unplanned communication.
3. Can a role be shared between two or more participants? Why or why not?

Course Outcome 2 (CO2):

1. Analyse the various Life cycle models
2. Compare the Sequential Activity Centered model with the Iterative Activity Centered models.

Course Outcome 3(CO3):

1. Draw a class diagram representing a book defined by the following statement: “A book is composed of a number of parts, which in turn are composed of a number of chapters. Chapters are composed of sections.” Focus only on classes and relationships.
2. Conduct requirement elicitation for a social media application.

Course Outcome 4 (CO4):

1. Cohesion and coupling controls the complexity of a system during subsystem decomposition. Illustrate using example.
2. Compare and contrast Dynamic Object Modelling with Static Object modelling.

Course Outcome 5 (CO5):

1. Discuss the configuration management concepts of a system.
2. Explain why maintenance is unavoidable in the field of software.

Course Outcome 6 (CO6):

1. Consider the problem statement of an E-commerce application:
2. Analyse the requirements and design the different UML diagrams

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 2**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS100

Course Name: OBJECT ORIENTED SOFTWARE ENGINEERING

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Being the member of the design team to develop an interface for an online registration portal, you are not sure about the mandatory fields. People in what role can help you out. Whether planned or unplanned communication will be more beneficial in this situation. Why? **(5)**

2. For what type of project, the spiral model suit's best. Why? **(5)**

3. Describe the different Requirement Elicitation techniques. **(5)**

4. Discuss the design Principles of System Design. **(5)**

5. Demonstrate the different steps of software deployment. **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6(a) Explain Work Breakdown Structure. **(3)**

(b) What are the types of project organizations? **(4)**

(a) Explain Iterative Activity Centered Models. **(3)**

(b) Explain Agile process? **(4)**

	Problem statement: Student Attendance Monitoring System. Student Attendance Monitoring System is used to track the attendance of students in an Institute. Faculty advisor can add students into the system, which is verified and approved by HOD. Once the students list is approved, the teachers can mark attendance on the system. Students can apply for duty leaves to the faculty advisor. Faculty advisor forwards the application to the HOD for approval. Students, teachers, faculty advisor and HOD can view the attendance reports of every student. Draw the sequence diagram for this problem statement.	**(7)**
(a)	Identify any four functional and non-functional requirements of KTU website.	**(4)**
(b)	Compare the Dynamic Object Modelling with the Static Object Modelling	**(3)**
(a)	Design patterns speed up the development process quiet a lot. Illustrate with example.	**(3)**
(b)	Discuss about the Object Constraint Language?	**(4)**
(a)	Discuss System Documentation	**(3)**
(b)	What is skill matrix? Briefly describe the project management activities?	**(4)**
(a)	Consider a method that will return the fare of a transport bus, given the source, destination and number of passengers. The source and destination are specified as integers. 1 represents station A, 2 represents station B etc. The total number of seats is 30. Generate test cases for Unit testing the system.	**(4)**
(b)	List out the benefits of model transformation.	**(3)**

a. Create the necessary test cases so as to conduct functional testing on the application.

Syllabus

Module 1: Classical Paradigm

System Design Concepts – Project Organization Concepts : Project Organizations , Roles , Tasks and Work Products ,Schedule – Project Communication concepts : Planned Communication , Unplanned Communication ,Communication Mechanism – Project Management Concepts : Tasks and Activities ,Work Products , Work Packages and Roles , Work Breakdown Structure .

Module 2: Process Models

Life cycle models: Sequential Activity Centered Models, Iterative Activity Centered models, Entity Centered models – Unified Process – Iterative and Incremental – Workflow – Agile Processes

Module 3: Analysis

Requirements Elicitation Concepts – An Overview of Unified Modeling Language –Analysis Concepts : Analysis Object Model and Analysis Dynamic Models – Non-functional requirements – Analysis

Patterns – Executable specification

Module 4: Design

System Design, Architecture – Design Principles - Design Patterns – Dynamic Object Modeling Static Object Modeling – Model based approach vs Document based approach – Interface Specification – Object Constraint Language

Module 5: Implementation, Deployment And Maintenance

Mapping Design (Models) to Code – Testing - Usability – Deployment – Configuration Management – Maintenance

	Course Plan	
No	**Topic**	**No. of Lectures (40 Hours)**
1	**Module 1: Classical Paradigm**	**10**
1.1	System Design Concepts	1
1.2	Project Organization Concepts- Project Organizations	1
1.3	Roles , Tasks, Work Products and Schedule	1
1.4	Project Communication concepts	1
1.5	Planned Communication , Unplanned Communication .	1
1.6	Communication Mechanism	1
1.7	Project Management Concepts	1
1.8	Tasks and Activities ,Work Products ,	1
1.9	Work Packages and Roles	1
1.10	Work Breakdown Structure	1
2	**Module 2: Process Models**	**8**
2.1	Life cycle models	1
2.2	Sequential Activity Centered Models	1
2.3	Iterative Activity Centered models	1
2.4	Entity Centered models	1
2.5	Unified Process	1
2.6	Iterative and Incremental	1
2.7	Workflow	1
2.8	Agile Processes	1
3	**Module 3: Analysis**	**7**
3.1	Requirements Elicitation Concepts	1
3.2	An Overview of Unified Modeling Language	1
3.3	Analysis Concepts	1
3.4	Analysis Object Model and Analysis Dynamic Models	1
3.5	Non-functional requirements	1
3.6	Analysis Patterns	1
3.7	Executable specification	1

4	**Module 4: Design**	**8**
4.1	System Design, Architecture	1
4.2	Design Principles	1
4.3	Design Patterns	1
4.4	Dynamic Object Modeling s	1
4.5	Static Object Modeling	1
4.6	Model based approach vs Document based approach	1
4.7	Interface Specification	1
4.8	Object Constraint Language	1
5	**Module 5: Implementation, Deployment And Maintenance**	**7**
5.1	Mapping Design (Models) to Code	1
5.2	Mapping Design (Models) to Code(Continued)	1
5.3	Testing	1
5.4	Usability	1
5.5	Deployment	1
5.6	Configuration Management	1
5.7	Maintenance	1

References

1. Bernd Bruegge, Alan H Dutoit, Object-Oriented Software Engineering, 2nd edition, Pearson Education, 2004.
2. Craig Larman, Applying UML and Patterns 3rd edition, Pearson Education, 2005
3. Stephen Schach, Software Engineering 7th ed, McGraw-Hill, 2007.
4. Ivar Jacobson, Grady Booch, James Rumbaugh, The Unified Software Development Process, Pearson Education, 1999.
5. Alistair Cockburn, Agile Software Development 2nd ed, Pearson Education.

241ECS001	ADVANCED DATA MINING	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course provides exposure to the concepts, principles and techniques of data mining. This course will enable the learners to identify the key process of Data mining and Warehousing, apply appropriate techniques to convert raw data into suitable format for practical data mining tasks, apply various data mining algorithms in appropriate domain, analyze the performance using performance metrics and extend data mining methods to the new domains of data. This course also helps to develop Data Mining systems which can analyze data efficiently and rigorously with suitable data models and techniques for respective applications.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Summarise basic concepts of Data mining and Illustrate feature vector representation for a given data collection **(Cognitive Knowledge Level: Understand)**
CO 2	Design Data Warehouse for problems in various domains. **(Cognitive Knowledge Level: Apply)**
CO 3	Implement Association Rules for analysing Transactional databases **(Cognitive Knowledge Level: Apply)**
CO4	Implement major Classification And Clustering Algorithms to a given problem. **(Cognitive Knowledge Level: Analyze)**
CO 5	To develop Data Mining system and analyze the performance **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✔		✔			✔	
CO 2	✔		✔			✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly. For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Implement an intelligent disease prediction system using feature selection techniques
2. Discuss different data reduction techniques with example
3. Implement an intrusion detection system using feature selection techniques
4. Discuss **how** mapping is done for different types of raw data to ML features with example
5. How we can eliminate noise using clustering? Discuss with example
6. Distinguish between cluster sampling and stratified sampling techniques with example

Course Outcome 2 (CO2):

1. Differentiate between Stars, Snowflakes and Fact constellation schemas
2. Suppose that a data warehouse consists of the dimensions time, branch, dealer, location and product, and the two measures unit-sold and revenue. Draw a star schema diagram and snowflake schema diagram for the data warehouse. Provide DMQL representation of star schema diagram and snowflake schema
3. List different schemas for a Data Warehouse Suppose that a Data Warehouse for Big University consists of the following four dimensions: student, course, semester and instructor and two measures count, avg_grade. When at the lowest conceptual level (e.g for a given student, course, semester and instructor combination), the avg_grade measure stores the actual course grade of the student. At higher conceptual levels avg_grade stores the average grade for the given combination. a) Draw the Snowflake schema diagram for the data warehouse b) Starting with the base cuboid [student, course, semester, instructor) what specific OLAP operations should one perform in order to list the average grade of CS courses for each Big University student.
4. What is the difference between BigQuery and Snowflake? What are the different ways to access the BigQuery Cloud Data warehouse ?
5. What are the data security features in Bigquery ?

Course Outcome 3(CO3):

1. Discuss Bayesian Networks and Data Modeling with an example
2. Implement spam filtering, Image enhancement using Bayesian Networks
3. Compare the R-tree to the R*-tree Discuss different spatial datamining primitives with example
4. Investigate and describe two techniques which have been used to predict future stock prices.
5. Apply the Apriori algorithm for discovering frequent itemsets from the following data set minimum support of 50% and minimum confidence of 75%.

Transaction ID	Items
100	Bread, Cheese
200	Bread, Cheese, Juice
300	Bread, Milk
400	Cheese, Juice. Milk

Course Outcome 4 (CO4):

1. Suppose a data collection consists of customer data of a bank. Implement customer fraud detection system
2. Suppose a corpus consists of data from medical domain. Implement a disease prediction system
3. Implement a Data Mining system to detect intrusions that may harm the database to offer greater security to the entire system.

Course Outcome 5 (CO5):

1. Implement a Data Mining system to assist Mobile service providers to design their marketing campaigns and to retain customers from moving to other vendors.
   Data collection consists of billing information, email, text messages, web data transmissions, and customer service and so on. The data mining system has to predict "churn" that tells the customers who are looking to change the vendors. The mobile service providers are then able to provide incentives, offers to customers who are at higher risk of churning.

Model Question Paper

QP CODE:

Reg No: ________________

Name: __________________ **PAGES: 3**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

CODE 241ECS001

Course Name: ADVANCED DATA MINING

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Differentiate between classification and regression with example	**(5)**
2.	Explain concept hierarchy generation .With a suitable example show how is it done for categorical data.	
3.	How can you generate association rules from frequent item sets?	**(5)**
4.	Why are nearest neighbor algorithms called lazy learners? What are the disadvantages of a lazy learner?	**(5)**
5.	How do we relate text mining and web mining? Differentiate between spatial and non spatial data with example	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.	(a)	Why feature engineering is important? What is the output of feature engineering in machine learning?	**(3)**
	(b)	Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. i) Use smoothing by bin means to smooth the above data, using a bin depth of 3. Illustrate your steps. ii) How might you determine outliers in the data	**(4)**
7.	(a)	How do data warehousing relate to data mining? Discuss	**(3)**
	(b)	Suppose that a data warehouse consists of the three dimensions time, doctor, and patient, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit. a) List three classes of schemas that are popularly used for modeling data warehouses. b) Draw a schema diagram for the above data warehouse using one of the schema classes listed in(a). Starting with the base cuboid [day,doctor,patient],what a specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2022?	**(4)**
8.	(a)	Why is the FP growth algorithm so efficient?	**(3)**

(b) Discuss FP growth algorithm. Using Apriori and FP growth algorithm find the frequent itemsets from the following transactional database? (min_sup= 2, confidence 70%). Compare the two processes **(4)**

TID	List of item-IDs
T100	I1,I2,I3
T200	I2,I4
T300	I2,I3
T400	I1,I2,I4
T500	I1,I3
T600	I2,I3
T700	I1,I3
T800	I1,I2,I3,I5
T900	I1,I2,I3

9. (a) What is the metric for classification tasks in CART? How to use the CART algorithm for classification **(3)**

(b) Differentiate between different types of ensemble methods for classification with example **(4)**

10. (a) How is the parameter “Distance-function” estimated in the DBSCAN Algorithm? What are the advantages and disadvantages of DBSCAN algorithm **(3)**

(b) What is the purpose of cluster ensemble? How do you create a cluster ensemble**?** Discuss with example **(4)**

Syllabus

Module 1: Data Mining and Knowledge Discovery

Desirable Properties of Discovered Knowledge – Knowledge representation, Data Mining Functionalities, Motivation and Importance of Data Mining, Classification of Data Mining Systems, Integration of a Data Mining System with a Database or Data Warehouse System, Classification, Clustering, Regression, Data Pre-processing: Data Cleaning, Data Integration and Transformation, normalization, standardization, Data Reduction, Feature vector representation. importance of feature engineering in machine learning; forward selection and backward selection for feature selection; curse of dimensionality; data imputation techniques; No Free Lunch theorem in the context of machine learning, Data Discretization and Concept Hierarchy Generation

Module 2: Data Warehouse and OLAP Technology for Data Mining

Data warehouses and its Characteristics - Data warehouse Architecture and its Components, Data

Warehouse Design Process, Data Warehouse and DBMS, Data marts, Metadata, Data Cube and OLAP, Extraction - Transformation – Loading - Schemas for Multidimensional Database: Stars, Snowflakes and Fact constellations, OLAP Cube - OLAP Operations - OLAP Server Architecture - Data Warehouse Implementation - From Data Warehousing to Data Mining, Trends in data warehousing

Module 3: Association Pattern Mining

Mining Frequent Patterns, Associations and Correlations –Mining Methods – Mining Various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining, Single Dimensional Boolean Association Rules From Transaction Databases, Multilevel Association Rules from transaction databases – Multi dimension Association Rules from Relational Database and Data Warehouses, Frequent Item Set Generation, Apriori Algorithm, Improved Apriori Algorithm for Association Rules Mining, Methods to improve Apriori, FP Growth Algorithm - Generating association rules from frequent itemset, Compact Representation of Frequent Item set - Maximal Frequent Item Set - Closed Frequent Item Sets. Pattern Evaluation Methods- Relationship Between FP-Growth and Enumeration-Tree Methods From Association Analysis to Correlation Analysis, Lift

Module 4: Classification and Prediction

Classification Techniques, Decision Tree - Decision tree Construction, Measures for Selecting the Best Split - Algorithm for Decision tree Induction - CART, Bayesian Belief Networks, Instance-Based Learning, K-Nearest neighbor classification, Accuracy and Error measures, Multiclass Classification, Semi-Supervised Classification, Multi class Learning, Rare class learning, Active Learning, Transfer Learning, Fuzzy Set Approaches for Classification, Rough Set Approaches, Techniques to improve classification accuracy-Ensemble methods, Bias-Variance Trade-off, Improving classification accuracy of class imbalanced data

Module 5: Cluster Analysis

Desired features of cluster Analysis, Types of data in cluster analysis, Categorization of Major Clustering Methods, Density-Based Methods, Clustering High Dimensional Data, Constraint Based Cluster Analysis, GA based clustering, Dealing with Large Databases, Probabilistic Model Based Clustering, Clustering with Constraints, Semi supervised clustering, Cluster Ensembles, Quality and validity of cluster analysis methods, Outlier Analysis-Statistical Approaches, Proximity Based Approaches. Advanced Mining: Multimedia Data Mining - Text Mining, Graph Mining and Social Network Analytics - Geospatial Data Mining, Temporal Mining, Data Mining Applications - Social Impacts of Data Mining.

	Course Plan	
No	**Topic**	**No. of Lectures (40 Hours)**
1	**Module 1: Data Mining and Knowledge Discovery**	**6**
1.1	Data Mining Functionalities, Motivation and Importance of Data Mining	1

1.2	Integration of a Data Mining System with a Database or Data Warehouse System, Major Issues in Data Mining. Classification, Clustering, Regression	1
1.3	Data Pre-processing: Data Cleaning, Data Integration and Transformation, normalization	1
1.4	Data Reduction, Different techniques	1
1.5	Feature vector representation. importance of feature engineering in machine learning;	1
1.6	Forward selection and backward selection for feature selection;	1
2	**Module 2: Data Warehouse and OLAP Technology for Data Mining**	**7**
2.1	Data warehouses and its Characteristics - Data warehouse Architecture and its Components	1
2.2	Data Warehouse and DBMS, Data marts, Metadata Extraction - Transformation – Loading in DW,	1
2.3	Multidimensional model	1
2.4	Schemas for Multidimensional Database: Stars, Snowflakes Fact constellations	1
2.5	Design Data Warehouse for problems in different domains	1
2.6	OLAP Cube - OLAP Operations	1
2.7	OLAP Server Architecture - Data Warehouse Implementation	1
3	**Module 3: Association Rule Mining**	**7**
3.1	Mining Frequent Patterns, Associations and Correlations	1
3.2	Mining Various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining	1
3.3	Multilevel Association Rules from transaction databases – Multi dimension Association Rules from Relational Database and Data Warehouses	1
3.4	Frequent Item Set Generation, Apriori Algorithm, Apriori Algorithm-illustration with example	1
3.5	Methods to improve Apriori, FP Growth Algorithm	1
3.6	FP Growth Algorithm- illustration with example, Compact Representation of Frequent Item set	1
3.7	Pattern Evaluation Methods, Association Analysis to Correlation Analysis, Lift	1
4	**Module 4: Classification and Prediction**	**10**
4.1	Classification Techniques, Decision Tree - Decision tree Construction Measures for Selecting the Best Split	1
4.2	Decision tree Induction - illustration with example Algorithm for Decision tree Induction - CART	1
4.3	Bayesian Belief Networks	1
4.4	Bayesian Belief Networks- Training	1

4.5	K-Nearest neighbor classification, Accuracy and Error measures	1
4.6	Multiclass Classification, Semi-Supervised Classification	1
4.7	Active Learning, Transfer Learning	1
4.8	Fuzzy Set Approaches for Classification	1
4.9	Rough Set Approaches	1
4.10	Ensemble methods. Improving classification accuracy of class imbalanced data	1
5	**Module 5: Cluster Analysis**	**10**
5.1	Desired features of cluster Analysis, Types of data in cluster analysis,	1
5.2	Categorization of Major Clustering Methods, Density-Based Methods,	1
5.3	Semi supervised clustering, Clustering High Dimensional Data, Constraint Based Cluster Analysis,	1
5.4	GA based clustering	1
5.5	Probabilistic Model Based Clustering	1
5.6	Quality and validity of cluster analysis methods, Outlier Analysis-Statistical Approaches, Proximity Based Approaches	1
5.7	Multimedia Data Mining	1
5.8	Text Mining	1
5.9	Graph Mining and Social Network Analytics	1
5.10	Geospatial Data Mining, Temporal Mining	1

References

1. Kevin Murphy, Machine Learning: A Probabilistic Perspective (MLAPP), MIT Press, 2012
2. Christopher Bishop, Pattern Recognition and Machine Learning (PRML), Springer, 2007.
3. Jiawei Han, Micheline Kamber, Jian Pei, “Data Mining: Concepts and Techniques”, Morgan Kaufmann, 2nd Ed., 2005
4. Charu C. Aggarwal ,Data Mining, Springer, ISBN 978-3-319-14141-1,2015
5. Data Mining Techniques,Arun K Puari, Universities Press,2001
6. Margaret H. Dunham, “Data Mining: Introductory and Advanced Topics”, Prentice Hall, 1st Ed., 2002.
7. David G. Stork, Peter E. Hart, and Richard O. Duda. Pattern Classification (PC), Wiley-Blackwell, 2000
8. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning (ESL), Springer, 2009
9. G. K. Gupta “Introduction to Data Mining with Case Studies”, Eastern Economy Edition, Prentice Hall of India, 2006.
10. Soumen Chakrabarti, "Mining the Web: Discovering Knowledge from Hypertext Data", Morghan Kaufmann, 1st Ed., 2005.
11. Da Ruan, Guoqing Chen, Etienne E. Kerre, Geert Wets, “Intelligent Data Mining: Techniques and Applications (Studies in Computational Intelligence)”, Springer, 1st Ed., 2010.

12. Masoud Mohammadian, "Intelligent Agents for Data Mining and Information Retrieval", Idea Group Publishing, 2004.
13. I. H. Witten and E. Frank. Data Mining: Practical Machine Learning Tools and Techniques Morgan Kaufmann, 2000.
14. D. Hand, H. Mannila and P. Smyth. Principles of Data Mining. Prentice-Hall. 2001

241ECS002	CLOUD COMPUTING	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: Study of cloud computing is an essential to understand the overall concepts of virtualization and virtual machines This course helps to gain expertise in server, network, storage virtualization, deploy practical virtualization solutions, enterprise solutions etc. They will be able to set up a private cloud by understand the security issues in the grid and the cloud environment.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Employ the concepts of storage virtualization, network virtualization and its management. **(Cognitive Knowledge Level: Apply)**
CO 2	Apply the concept of virtualization in the cloud computing. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply domain knowledge in architecture, infrastructure and delivery models of cloud computing in designing and developing cloud applications. **(Cognitive Knowledge Level: Apply)**
CO 4	Develop services using Cloud computing. **(Cognitive Knowledge Level: Apply)**
CO 5	Analyse and choose security models appropriate to the cloud environment. **(Cognitive Knowledge Level: Analyse)**
CO 6	Design, develop and implement cloud-based applications. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✔		✔	✔	✔	✔	
CO 2	✔		✔	✔	✔	✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔		✔	✔	✔	✔	
CO 6	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyse	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii.Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. A project for 2 months requires 1000TB of memory during the development phase. Predict the cloud service that can be used and list the advantages.
2. Illustrate different types of hypervisors with examples. Also enlist the advantages and disadvantages of each.

Course Outcome 2 (CO2):

1. Virtualization can be applied into different levels, "ranging from hardware to application". Comment your opinion with explanation.
2. An American e-commerce web site "Nordstrom" was experiencing a high increase in their customers before New Year. What type of resource provisioning can be done here? Explain.
3. In a virtual environment, guest OS cannot directly access Host machine memory. How can this be achieved.

Course Outcome 3(CO3):

1. How hybrid cloud helps in the growth of your business.
2. Your company runs a virtualized web application server in-house. You decide to make the web applications available over the Internet through a cloud provider. Which method is the quickest way to accomplish this?
3. If 2 teams from US and India are collaboratively working on a project, discuss a means by which they can access data. Explain with 2 examples.

4. Imagine you are conducting Arts Festival of your college. Explain the different steps that you will take to make the event successful using cloud.

Course Outcome 4 (CO4):

1. Write the steps to configure Hadoop Map Reduce environment in Linux for developing a Map Reduce program.
2. Write a word count Map Reduce program in Java.
3. Identify the storage system used by Google Earth software. Explain how to locate a data in such a data store.
4. Identify the cloud service model used in Netflix. Justify your answer.

Course Outcome 5 (CO5):

1. A company XYZ wishes to lease resources in the cloud. List and explain security issues that must be discussed with Technology Analyst to ensure secure cloud usage.
2. Identify the cloud service offered by Gmail & Google drive and explain key features of each service?
3. Why it is harder to establish security in the cloud?

Course Outcome 6 (CO6):

1. Design, develop and implement an efficient cloud based parallel programming model to count distinct place names in kerala.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES:2**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS002

Course Name: CLOUD COMPUTING

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Sketch the core the differences between a traditional computer and a virtual machine.	**(5)**
2.	Explain your understanding about virtualization. What is the role of VMM in virtualization?	**(5)**
3.	Illustrate PaaS model for cloud computing.	**(5)**
4.	Summarize the concept of Map Reduce? Explain the logical data flow of Map Reduce function using suitable example.	**(5)**
5.	illustrate the major security challenges in clouds?	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.	How Memory virtualization is implemented? Provide necessary examples and diagrams wherever necessary	**(7)**
7.	Investigate the functional modules of Google App Engine ?	**(7)**
8.	Sketch the core idea about virtualization. What is the role of VMM in virtualization?	**(7)**
9.	Is it harder to establish security in the cloud? Justify	**(7)**
10.	With a neat diagram explain the Generic Cloud architecture and components.	**(7)**
11.	With neat diagram explain your understanding Security Architecture Design in cloud.	**(7)**
12.	Demonstrate Private Cloud Design using Open Nebula.	**(7)**

Syllabus

Module 1: Virtualization

Basics of Virtual Machines - Process Virtual Machines – System Virtual Machines –Emulation – Interpretation – Binary Translation - Taxonomy of Virtual Machines. Virtualization –Management — Hardware Maximization – Architectures – Virtualization Management – Storage Virtualization – Network Virtualization

Module 2: Virtualization Infrastructure

Comprehensive Analysis – Resource Pool – Testing Environment –Server Virtualization – Virtual Workloads – Provision -Virtual Machines – Desktop Virtualization – Application Virtualization - Implementation levels of virtualization – virtualization structure – virtualization of CPU-Memory and I/O devices – virtual clusters and Resource Management – Virtualization for data centre automation.

Module 3: Cloud Platform Architecture

Understanding cloud computing-Cloud Computing – History of Cloud Computing- Advantages and Disadvantages of Cloud Computing- Cloud deployment models-public-private- hybrid- Categories of cloud computing-Everything as a service-Infrastructure-platform-software- A Generic Cloud Architecture Design – Layered cloud Architectural Development – Virtualization Support and Disaster Recovery – Architectural Design Challenges - Public Cloud Platforms –GAE-AWS – Inter-cloud Resource Management

Module 4: Programming Mode

Introduction to Hadoop Framework – Map Reduce-Input splitting-map and reduce functions-specifying input and output parameters-configuring and running a job –Developing Map Reduce Applications - Design of Hadoop file system –Setting up Hadoop Cluster - Cloud Software Environments –Eucalyptus-Open Nebula-Open Stack-Nimbus

Module 5: Cloud Security

Cloud Infrastructure security- network, host and application level – aspects of data security-provider data and its security-Identity and access management architecture-IAM practices in the cloud-SaaS-PaaS-IaaS availability in the cloud - Key privacy issues in the cloud –Cloud Security and Trust Management

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Virtualization**	**8**
1.1	Basics of Virtual Machines	1
1.2	Process Virtual Machines, System Virtual Machines	1
1.3	Emulation, Interpretation	1
1.4	Binary Translation	1
1.5	Taxonomy of Virtual Machines	1
1.6	Virtualization –Management, Hardware Maximization	1
1.7	Architectures, Virtualization Management	1
1.8	Storage Virtualization, Network Virtualization	1
2	**Module 2: Virtualization Infrastructure**	**8**
2.1	Comprehensive Analysis, Resource Pool	1

2.2	Testing Environment, Server Virtualization	1
2.3	Virtual Workloads	1
2.4	Provision, Virtual Machines	1
2.5	Desktop Virtualization, Application Virtualization	1
2.6	Implementation levels of virtualization, virtualization structure, virtualization of CPU	1
2.7	Memory and I/O devices	1
2.8	virtual clusters and Resource Management, Virtualization for data centre automation	1
3	**Module 3: Cloud Platform Architecture**	**9**
3.1	Understanding cloud computing-Cloud Computing – History of Cloud Computing- Advantages and Disadvantages of Cloud Computing	1
3.2	Cloud deployment models, Public-private- hybrid, Categories of cloud computing	1
3.3	Everything as a service, Infrastructure	1
3.4	Platform, Software	1
3.5	A Generic Cloud Architecture Design, Layered cloud Architectural Development	1
3.6	Virtualization Support and Disaster Recovery, Architectural Design Challenges	1
3.7	Public Cloud Platforms	1
3.8	GAE, AWS	1
3.9	Inter-cloud Resource Management	1
4	**Module 4: Programming Mode**	**8**
4.1	Introduction to Hadoop Framework, Map Reduce	1
4.2	Input splitting	1
4.3	map and reduce functions, specifying input and output parameters	1
4.4	configuring and running a job, Developing Map Reduce Applications	1
4.5	Design of Hadoop file system, Setting up Hadoop Cluster	1
4.6	Cloud Software Environments, Eucalyptus	1
4.7	Open Nebula, Open Stack	1
4.8	Nimbus	1
5	**Module 5: Cloud Security**	**7**
5.1	Cloud Infrastructure security	1
5.2	network, host and application level	1
5.3	aspects of data security, provider data and its security	1
5.4	Identity and access management architecture	1
5.5	IAM practices in the cloud	1
5.6	SaaS, PaaS, IaaS availability in the cloud	1
5.7	Key privacy issues in the cloud, Cloud Security and Trust Management	1

References

1. Greg Schulz, “Cloud and Virtual Data Storage Networking”, Auerbach Publications [ISBN: 978-1439851739], 2011.
2. Michael Miller, Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online, Que Publishing, August 2008.
3. GauthamShroff, “Enterprise Cloud Computing: Technology, Architecture, Applications”, Cambridge press, 2010.
4. EMC, “Information Storage and Management” Wiley; 2 edition [ISBN: 978-0470294215],2012.
5. Kai Hwang , Geoffrey C Fox, Jack J Dongarra : “Distributed and Cloud Computing – From Parallel Processing to the Internet of Things” , Morgan Kaufmann Publishers – 2012.

241ECS003	WEB SERVICES	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course provides an understanding of the purpose of using web services in web development. The topics covered in this course are the basic concepts and types of web services, server side and client-side web services and an introduction to development of dynamic web pages. The learners will be able to develop a web-based service application.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Illustrate the need of web services in developing applications. **(Cognitive Knowledge Level: Understand)**
CO 2	Make use of Server-side and Client-side RESTful web services. **(Cognitive Knowledge Level: Analyze)**
CO 3	Analyze how web services can be published in standalone web servers. **(Cognitive Knowledge Level: Analyze)**
CO 4	Employ techniques on creating dynamic web pages. **(Cognitive Knowledge Level: Apply)**
CO 5	Utilize emerging technologies in web services. **(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, Implement and Present innovative ideas on modern web services concepts and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓		✓	
CO 2	✓		✓	✓		✓	
CO 3	✓		✓	✓		✓	
CO 4	✓		✓	✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Illustrate how HTTP requests and responses are handled.
2. Discuss the different types of Web services
3. Identify the commonly used HTTP methods in REST Architecture
4. Explain the advantages of Service Oriented Architecture

Course Outcome 2 (CO2):

1. Explain RESTful web services.
2. Examine the purpose of Client-Side and Server side API
3. Elaborate on the advanced features of Client API.

Course Outcome 3(CO3):

1. Explain the need for JSON
2. Illustrate the steps for publishing JAX-RS Resources to Apache Tomcat.

Course Outcome 4 (CO4):

1. How can we interpret the architecture of SOAP based service?
2. Demonstrate how the WSDL document, can be put to good use in writing a Java client against the RandService.
3. Implement dynamic content on web page using Web method in form of JSON object.

Course Outcome 5 (CO5):

1. Demonstrate dynamic webpages using React.js and Angular.js
2. Use ReactJS to build a simple search filter functionality to display a filtered list based on the search query entered by the user.

Course Outcome 6 (CO6):

1. Develop a UI based application and create client/server side web services to perform relevant tasks related to project requirements.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 2**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS003

Course Name: WEB SERVICES

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Examine the possible reasons for the choice of Web services over Web applications.	**(5)**
2.	Why are HTTP servlets a convenient way for implementing RESTful web services?	**(5)**
3.	Explain how a basic client request can be created using the Client API	**(5)**
4.	Describe the architecture of a SOAP based Web Service.	**(5)**
5.	Discuss the facts how React is different from Angular.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.	How are SOAP and RESTful Web services architecturally different from each other?	**(7)**
7.	Illustrate the process of publishing JAX-RS resources with Tomcat.	**(7)**
8.	'JSONP brings an event-driven API to client-side processing'- Justify the statement using examples.	**(7)**
9.	Explain how a documents client service *REST Client* can be created from the DOCS DWADL.	**(7)**
10.	Describe the structure of a WSDL document with example.	**(7)**
11.	Analyse the purpose of render in React. How would you prevent unnecessary component re-render in ReactJS?	**(7)**
12.	Demonstrate with examples, the two approaches that AngularJS takes, to build forms.	**(7)**

Syllabus

Module 1: Introduction to Web Services
Web Services Standards Organizations, Service oriented architecture, Advantages of web services over distributed object architecture, SOAP-based web services, RESTful web services.
Review of HTTP requests and responses- HTTP as an API, A RESTful example, Use of Servlets for RESTful Web Services.

Module 2: RESTful Web Services: The Service Side
A RESTful service as an HttpServlet, Implementation details.
A RESTful Web Service as a JAX-RS Resource- JAX-RS Web Service Using Jersey, Publishing JAX-RS Resources with a Java Application, Publishing JAX-RS Resources with Tomcat, JAX-RS Generation of XML and JSON Responses, Porting the Predictions Web Service to JAX-RS.
A RESTful Web Service as Restlet Resources.
GraphQL- Introduction to GraphQL, GraphQL Architecture, Basic Queries.

Module 3: RESTful Web Services: The Client Side- A Perl Client Against a Java RESTful Web Service
RESTful Clients and WADL Documents- The JAX-RS Client API, JSON for JavaScript Clients-JSONP and Web Services.

Module 4: SOAP-Based Web Services- Introduction and Evolution of SOAP, Architecture of a typical SOAP-based service, Publishing a SOAP-Based Service with a Standalone Web Server
RandService- JavaClient Against the RandService, C# Client Against the RandService, A Perl Client Against the RandService. WSDL – WSDL document structure.

Module 5: Introduction to React.js, Node.js and Angular.js
Introduction to React and Node-Basic Concepts and Applications, Rendering Elements and Components, Comparison and Purpose of Node.js and React.js, Angular JS Basics-Modules, Creating Components, Directives, Filters, Angular Forms-Services, Single page application and Multipage application, Use case of a real-time single page chat application.

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Introduction to Web Services**	**8**
1.1	Introduction to web services- Web Services Standards Organizations, Service oriented architecture	1
1.2	Advantages of web services over distributed object architecture	1
1.3	SOAP-based web services	1
1.4	RESTful web services	1

1.5	Review of HTTP requests and responses	1
1.6	HTTP as an API	1
1.7	A RESTful example	1
1.8	Use of Servlets for RESTful Web Services	1
2	**Module 2: RESTful Web Services: The Service Side**	**8**
2.1	A RESTful service as an HttpServlet, Implementation details	1
2.2	A RESTful Web Service as a JAX-RS Resource	1
2.3	JAX-RS Web Service Using Jersey	1
2.4	Publishing JAX-RS Resources with a Java Application	1
2.5	Publishing JAX-RS Resources with Tomcat	1
2.6	JAX-RS Generation of XML and JSON Responses	1
2.7	Porting the Predictions Web Service to JAX-RS	1
2.8	GraphQL- Introduction to GraphQL, GraphQL Architecture, Basic Queries	1
3	**Module 3: RESTful Web Services: The Client Side**	**8**
3.1	A Perl Client Against a Java RESTful Web Service (Lecture 1)	1
3.2	A Perl Client Against a Java RESTful Web Service (Lecture 2)	1
3.3	RESTful Clients and WADL Documents (Lecture 1)	1
3.4	RESTful Clients and WADL Documents (Lecture 2)	1
3.5	The JAX-RS Client API	1
3.6	JSON for JavaScript Clients	1
3.7	JSONP and Web Services (Lecture 1)	1
3.8	JSONP and Web Services (Lecture 2)	1
4	**Module 4: SOAP-Based Web Services**	**8**
4.1	Introduction and Evolution of SOAP	1
4.2	Architecture of a typical SOAP-based service	1
4.3	Publishing a SOAP-Based Service with a Standalone Web Server (Lecture 1)	1
4.4	Publishing a SOAP-Based Service with a Standalone Web Server (Lecture 2)	1
4.5	RandService- JavaClient Against the RandService	1
4.6	C# Client Against the RandService	1
4.7	A Perl Client Against the RandService	1
4.8	WSDL – WSDL document structure	1
5	**Module 5: Introduction to React.js, Node.js and Angular.js**	**8**
5.1	Introduction to React and Node, Basic Concepts and Applications	1
5.2	Rendering Elements and Components	1
5.3	Comparison and Purpose of Node.js and React.js	1
5.4	Angular JS Basics-Modules	1
5.5	Creating Components, Directives,Filters	1
5.6	Angular Forms-Services	1

5.7	Single page application and Multipage application	1
5.8	Use case of a real-time single page chat application.	1

References

1. Martin Kalin, Java Web Services: Up and Running, Second Edition, O’Reilly, 2013
2. Robin Wieruch, The Road to Learn React, https://www.roadtoreact.com, 2022 Edition
3. Andrew Grantt, Beginning Angular JS, Apress, 2014
4. Vipul A M and Prathamesh Sonpatki, ReactJS by Example - Building Modern Web Applications with React, PACKT Publishing, 2016
5. Node.js Web Development: Server-side web development made easy with Node 14 using practical examples, 5th Edition, PACKT Publishing, 2020
6. Alex Banks, Eve Porcello, Learning GraphQL, O'Reilly Media, 2018

241ECS004	COMPUTATIONAL INTELLIGENCE	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: The aim of this course is to provide the students with the knowledge and skills required to design and implement effective and efficient Computational Intelligence solutions to problems for which a direct solution is impractical or unknown. This course covers concepts of fuzzy logic, genetic algorithms, and swarm optimization techniques. The learners will be able to provide Fuzzy and AI –based solutions to real world problems.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply fuzzy logic to handle uncertainty and solve engineering problems. **(Cognitive Knowledge Level : Apply)**
CO 2	Apply Fuzzy Logic Inference methods in building intelligent machines. **(Cognitive Knowledge Level : Apply)**
CO 3	Design genetic algorithms for optimized solutions in engineering problems. **(Cognitive Knowledge Level : Analyze)**
CO 4	Analyze the problem scenarios and apply Ant colony system to solve real optimization problems. **(Cognitive Knowledge Level : Analyze)**
CO 5	Apply PSO algorithm to solve real world problems. **(Cognitive Knowledge Level : Apply)**
CO6	Design, develop and implement solutions based on computational intelligence concepts and techniques. **(Cognitive Knowledge Level : Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1				✓		✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Let V = {A, B, C, D} be the set of four kinds of vitamins, F = {f_1, f_2, f_3} be three kinds of fruits containing the vitamins to various extents, and D = {d_1, d_2, d_3} be the set of three diseases that are caused by deficiency of these vitamins. Vitamin contents of the fruits are expressed with the help of the fuzzy relation R over F x V, and the extent of which diseases are caused the deficiency of these vitamins is given by the fuzzy relation S over V x D. Relations R and S are given below

$R = [0.5\ 0.2 \quad 0.2\ 0.7 \quad 0.4\ 0.4 \quad 0.1\ 0.1\ 0.4\ 0.3\ 0.8\ 0.1\ ] S = [0.3\ 0.5\ 0.1\ 0.8 \ \ 0.7\ 0.4 \ \ 0.9\ 0.1\ 0.5\ 0.$

Find the correlation between the amount of certain fruit that should be taken while suffering from a disease.

Course Outcome 2 (CO2):

1. In mechanics, the energy of a moving body is called kinetic energy. Suppose we model mass and velocity as inputs to a moving body and energy as output. Observe the system for a while and the following rule is deduced.

IF x is small and y is high
THEN z is medium
The graphical representation of rule is given below. Let the inputs given are 0.35kg and 55m/s. What will the output using Mamdani inference? Any defuzzification method can be used to obtain the crisp single output.

Course Outcome 3(CO3):

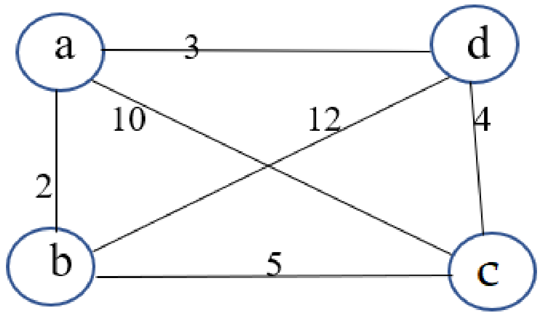
1. Describe how Roulette wheel is used for selection. Draw the Roulette wheel for six chromosomes corresponding to the table given below.

Chromosome #	***Fitness***
1	10
2	5
3	25
4	15
5	30
6	20

Course Outcome 4 (CO4):

1. Consider an Ant Colony System based on Ant Quantity model for solving the following Travelling Salesman Problem. Compute the pheromone content at each of the edges after 4 steps (1 iteration). Assume pheromone decay factor ρ=0.1, Q = 120. Assume initial pheromone of 50 units at each of the edges and that three ants k1,k2 and k3 follow the paths given below in the first iteration.

   k1= a b c d a; k2=a c b d a; k3=a d c b a



2. Six jobs go first on machine A, then on machine B, and finally on machine C. The order of the completion of the jobs in the three machines is given in Table

Jobs	Processing time(hr)		
	Machine A	Machine B	Machine C
1	8	3	8
2	3	4	7
3	7	5	6
4	2	2	9
5	5	1	10
6	1	6	9

Find the sequence of jobs that minimizes the time required to complete the jobs using the ACS model.

Course Outcome 5 (CO5):

1. Consider a particle swarm optimization system composed of three particles and maximum velocity 10. Assume that both the random numbers r1 and r2 used for computing the movement of the

particle towards the individual best position and social best position are 0.5. Also assume that the space of solutions is the two-dimensional real valued space and the current state of swarm is as follows:

Position of particles: x1 = (4,4); x2 = (8,3); x3 = (6,7)
Individual best positions : x14,4) = *); x27,3) = *); x35,6) = *)
Velocities: v1 = (2,2); v2 = (3,3); v3 = (4,4)
What would be the next position of each particle after one iteration of the PSO algorithm if the inertia parameter ω that is used along with current velocity update formula is 0.8 ?

Course Outcome 6 (CO6):

1. Implement travelling salesman problem using appropriate optimization technique.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 5**

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS004

Course Name: Computational Intelligence

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Consider the set of Colours A= {Blue, Red, Orange, Yellow, Green}, Attributes B = {Bright, Warmth, Dullness}, Feelings C= {Unpleasant, happiness, Angry}. Given R and S where R is the relationship between colours and their attributes and S is the relationship between colour attributes and feelings created. Find the relationship Q between colours and feelings created **(5)**

R	Bright	Warmth	Dullness
Blue	0.8	0.6	0.4
Red	0.8	0.8	0.2
Orange	0.5	0.7	0.2
Yellow	0.3	0.6	0.5
Green	0.8	0.6	0.4

S	Unpleasant	Happiness	Angry
Bright	0.2	0.8	0.6
Warmth	0.4	0.7	0.8
Dullness	0.8	0.3	0.6

2. Develop a membership function for "Tall". Based on that devise membership function for "Very Tall". Explain how it is done **(5)**

3. Mention the importance of objective (fitness) function in genetic algorithm **(5)**

4. Describe how pheromone is updated. What is elitist / elastic ants ? Are they useful in this scenario? **(5)**

5. What is the significance of pbest and gbest particles in solving problems with particle swarm optimization? **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) Consider the set of fruits F = {Apple , Orange, Lemon, Strawberry, Pineapple}. **(3)**

Let sweet fruits B= $\{\frac{0.8}{Apple} + \frac{0.6}{Orange} + \frac{0.2}{Lemon} + \frac{0.4}{Strawberry} + \frac{0.7}{Pineapple}\}$ and

Sour Fruits F= $\{\frac{0.6}{Apple} + \frac{08}{Orange} + \frac{0.9}{Lemon} + \frac{0.7}{Strawberry} + \frac{0.5}{Pineapple}\}$

Find Fruits that are Sweet or Sour, Sweet but not Sour, Sweet and Sour

(b) Consider two fuzzy Sets given by **(4)**

$$P = \left\{\frac{0.9}{short} + \frac{0.3}{medium} + \frac{0.5}{tall}\right\}$$

$$Q = \left\{\frac{0.7}{positive} + \frac{0.4}{zero} + \frac{0.8}{negative}\right\}$$

Find the fuzzy relation for the Cartesian product of P and Q i.e, R=P x Q .
Introduce a fuzzy set T given by

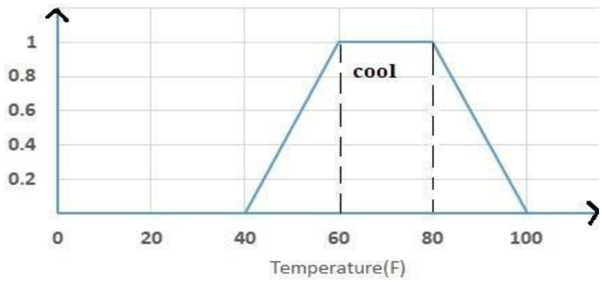
$$T = \left\{\frac{0.9}{short} + \frac{0.3}{medium} + \frac{0.6}{tall}\right\}$$

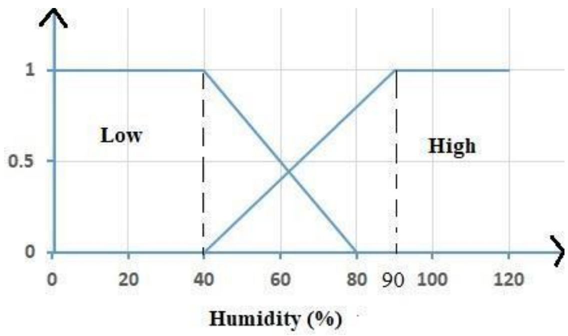
and Find T o R using max-min composition

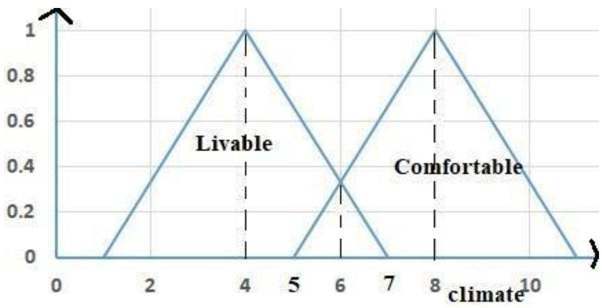
7. Consider a Fuzzy Inference System for checking climate comfortability of human beings for long time living. The system accepts two inputs – temperature and humidity. The rules and membership functions of FIS is given below. Using Mamdani inference and center of sum, calculate output when the temperature is 50Fahrenheit and humidity is 50%. **(7)**

Rule 1: IF temperature is cool and humidity is low, THEN climate is comfortable.

Rule 2: IF temperature is cool and humidity is high, THEN climate is livable







The fuzzy sets “Easy Question Paper” and their corresponding “Student Performance” are given below

$$\text{Easy\_QP} = \left\{ \frac{0.8}{1} + \frac{0.2}{2} + \frac{0.6}{3} + \frac{0.7}{4} \right\}$$

$$\text{Stud\_Perf} = \left\{ \frac{0.3}{a} + \frac{0.4}{b} + \frac{0.8}{c} + \frac{0.9}{d} \frac{0.8}{1} + \frac{0.2}{2} + \frac{0.6}{3} \right\}$$

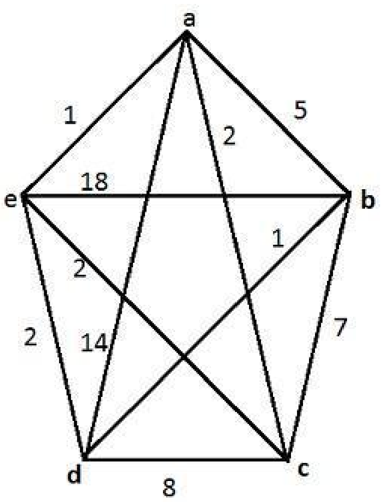
Find the performance of students c and d for the question paper “Somewhat Easy”

$$\text{Somewhat\_Easy} = \left\{ \frac{0.7}{1} + \frac{0.3}{2} + \frac{0.5}{3} + \frac{0.6}{4} \right\}$$

8. Explain any procedure to map a solution to the corresponding chromosome and vice versa in genetic algorithms. Also illustrate it with an example **(7)**

9. Describe two methods used to select individuals from a population for the mating pool in Genetic Algorithms **(7)**

10. (a) Consider the TSP with the following edge costs. Given the evaporation factor ρ =0.02 and initial pheromone at all edges T_{ij}=100 **(1)**



What is the cost of best tour?

(b) Using the equation $T_{ij}(t+1)=(1-\rho)T_{ij}(t) + \Delta T_{ij}(t,t+1)$, compute the T_{ij} of the edge<a,c> when 10 ants uses the edges <a, c>, using the following models: **(6)**

i. Ant Density Model (Constant Q=10)

ii. Ant Quantity Model(Constant Q=100)

where Q is the constant related to the pheromone updation.

11. Describe Ant Colony System. What are the different types of Ant systems? **(7)**

12.	Consider a particle swarm optimization system composed of three particles and maximum velocity 10. Assume that both the random numbers r1 and r2 used for computing the movement of the particle towards the individual best position and social best position are 0.5. Also assume that the space of solutions is the two dimensional real valued space and the current state of swarm is as follows: Position of particles: x1 = (4,4); x2 = (8,3); x3 = (6,7) Individual best positions : x14,4) = *); x27,3) = *); x35,6) = *) Velocities: v1 = (2,2); v2 = (3,3); v3 = (4,4) What would be the next position of each particle after one iteration of the PSO algorithm if the inertia parameter ω that is used along with current velocity update formula is 0.8 ?	**(7)**

Syllabus

Module 1 : Fuzzy Logic

Crisp sets vs fuzzy sets- Operations and properties of Fuzzy sets. Membership functions - Linguistic variables. Operations on fuzzy sets- Fuzzy laws- Operations on fuzzy relations, Fuzzy composition- Max- min **,** Max – product. Alpha-cut representation.

Module 2: Fuzzy Systems

Fuzzy Reasoning – GMP and GMT. Fuzzy Inference System: Defuzzification methods - Fuzzy Controllers -Mamdani FIS, Larsen Model

Module 3: Genetic Algorithms

Introduction to Genetic Algorithms – Theoretical foundation - GA encoding , decoding - GA operations – Elitism – GA parameters – Convergence. Multi-objective Genetic Algorithm – Pareto Ranking.

Module 4: Ant Colony Systems

Swarm intelligent systems - Background Ant colony systems – Biological systems- Development of the ant colony system- - Working - Pheromone updating- Types of ant systems- ACO algorithms for TSP

Module 5: Particle Swarm Optimization

Basic Model - Global Best PSO- Local Best PSO- Comparison of 'gbest' to 'lbest'- PSO Algorithm Parameters- Problem Formulation of PSO algorithm- Working. Rate of convergence improvements -Velocity clamping- Inertia weight- Constriction Coefficient- Boundary Conditions- Guaranteed Convergence PSO- Initialization, Stopping Criteria, Iteration Terms and Function Evaluation.

	Course Plan	
No	**Topic**	**No. of Lectures (40)**
1	**Module 1: Fuzzy Logic**	**9**
1.1	Crisp sets vs fuzzy sets, Operations and properties of Fuzzy sets	1
1.2	Membership functions	1
1.3	Linguistic Variables	1
1.4	Operations on fuzzy sets	1
1.5	Fuzzy laws	1
1.6	Operations on fuzzy relations	1

1.7	Fuzzy Composition- Max- min	1
1.8	Fuzzy Composition – Max- Product	1
1.9	Alpha-cut representation	1
2	**Module 2: Fuzzy Systems**	**7**
2.1	Fuzzy Reasoning – GMP	1
2.2	Fuzzy Reasoning –GMT	1
2.3	Fuzzy Inference System	1
2.4	Defuzzification methods	1
2.5	Fuzzy Controllers	1
2.6	Mamdani Model	1
2.7	Larsen Model	1
3	**Module 3: Genetic Algorithms**	**7**
3.1	Introduction to Genetic algorithm	1
3.2	Theoretical foundation	1
3.3	GA encoding - decoding	1
3.4	GA operations	1
3.5	Elitism, GA parameters ,Convergence of GA	1
3.6	Multi – objective Genetic Algorithm	1
3.7	Pareto Ranking	1
4	**Module 4: Ant Colony Systems**	**8**
4.1	Swarm intelligent systems	1
4.2	Background	1
4.3	Ant colony systems – biological systems	1
4.4	Development of the ant colony system	1
4.5	Working	1
4.6	Pheromone updating	1
4.7	Types of ant systems	1
4.8	ACO algorithms for TSP	1
5	**Module 5: Particle Swarm Optimization**	**9**
5.1	Basic Model	1
5.2	Global Best PSO	1
5.3	Local Best PSO, Comparison of ‘gbest’ to ‘lbest’	1
5.4	PSO Algorithm Parameters	1
5.5	Problem Formulation	1
5.6	Working	1
5.7	Rate of convergence improvements – velocity clamping	1
5.8	Inertia-weight - Constriction Coefficient- Boundary Conditions	1
5.9	Initialization, Stopping Criteria, Iteration Terms and Function Evaluation	1

References

1. Samir Roy, Udit Chakraborty, Introduction to Soft Computing Neuro- Fuzzy Genetic Algorithms, Pearson, 2013
2. N.P. Padhy, Artificial Intelligence and Intelligent systems, Oxford Press, New Delhi, 2005.
3. Xin-She Yang School of Science and Technology, Middlesex University London, Nature-Inspired Optimization Algorithms, Elsevier, First edition, 2014
4. Satyobroto Talukder, Blekinge Institute of Technology, Mathematical Modelling and Applications of Particle Swarm Optimization, February 2011
5. Mitchell Melanie, An Introduction to Genetic Algorithm, Prentice Hall,1998
6. Andries Engelbrecht, Computational Intelligence: An Introduction, Wiley, 2007
7. Marco Dorigo and Thomas Stutzle, "Ant Colony optimization", Prentice Hall of India, New Delhi 2005

241ECS005	AUTOMATED VERIFICATION	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: The course presents an algorithmic approach to the development of formal verification systems. It concentrates on decision procedures for decidable first-order theories that are useful in the context of automated verification and reasoning, theorem proving, compiler optimization, and so forth. Since the ability of these techniques to cope with problems arising in industry depends critically on decision procedures, this is a vibrant and prospering research subject for many researchers around the world, both in academia and in industry.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Use the model-theoretic and proof-theoretic approaches towards formal reasoning. (**Cognitive Knowledge Level: Apply**)
CO 2	Demonstrate how decision procedures can be developed for propositional logic using SAT solvers and Binary Decision Diagrams. (**Cognitive Knowledge Level: Apply**)
CO 3	Develop methods to prove the validity and satisfiability of formulas using Equality Logic and Uninterpreted Functions. (**Cognitive Knowledge Level: Apply**)
CO 4	Illustrate decision procedures using linear arithmetic (**Cognitive Knowledge Level: Analyze**)
CO 5	Design, develop and implement solutions based on the concepts of automated verification. (**Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✔		✔	✔		✔	
CO 2	✔		✔	✔		✔	
CO 3	✔		✔	✔		✔	
CO 4	✔		✔	✔		✔	
CO 5	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Using Tseitin's encoding, transform the following formula to Conjunctive Normal Form (CNF).

$$\emptyset : = \neg\left(x_1 \wedge (x_2 \vee ... \vee x_n)\right)$$

2. Let T_1 and T_2 be two theories whose satisfiability problem is decidable and in the same complexity class. Is the satisfiability problem of a T_1 formula reducible to a satisfiability problem of a T_2 formula? Why?

Course Outcome 2 (CO2):

1. Construct the Binary Decision Diagram (BDD) for $\neg(x_1 \vee (x_2 \wedge \neg x_3))$ with the variable order x_1, x_2, x_3 starting from a decision tree.
2. Show a formulation in propositional logic of the following problem: given a directed graph, does it contain a Hamiltonian cycle?

Course Outcome 3(CO3):

1. Given the formula $F\left(F\left(x_1\right)\right) \neq F\left(x_1\right) \wedge F\left(F\left(x_1\right)\right) \neq F\left(x_2\right) \wedge x_2 = F(x_1)$, reduce its validity problem to a validity problem of an equality logic formula through Ackermann's reduction and Bryant's reduction.
2. Prove the following Lemma.

Lemma: If a domain D is adequate for $\emptyset(e)$and $e^{'} \subseteq e$, then D is adequate for $\emptyset(e')$.

Course Outcome 4 (CO4):

1. Find the worst case run time of the general simplex algorithm if applied to a conjunction of difference logic constraints.
2. Prove that the feasibility problem for integer linear programming is NP-hard.

Course Outcome 5 (CO5):

1. Develop a SAT solver that verifies a compilation process with Translation Validation.

Model Question paper

First Semester M. Tech. Degree Examination, Month, Year

241ECS005 – Automated Verification

Time: 2.5 Hours **Max. Marks: 60**

Part A

(Answer all questions. Each question carries 5 marks) **5x5 = 25 Marks**

1. Use Tseitin's encoding to convert the formula $x_1 \Rightarrow (x_2 \wedge x_3)$ to Conjunctive Normal Form (CNF).
2. Consider a formula that contains the following set of clauses.

$$C_1 = (\neg x_1 \vee x_2),$$
$$C_2 = (\neg x_1 \vee x_3 \vee x_5),$$
$$C_3 = (\neg x_2 \vee x_4),$$
$$C_4 = (\neg x_3 \vee \neg x_4),$$
$$C_5 = (x_1 \vee x_5 \vee \neg x_2),$$
$$C_6 = (x_2 \vee x_3),$$
$$C_7 = (x_2 \vee \neg x_3),$$
$$C_8 = (x_6 \vee \neg x_5)$$

   a. Draw a partial implication graph for decision level 6, after a decision $x_1 = 1$.
   b. Draw a partial implication graph after learning a conflict clause $C_9 = (x_5 \vee \neg x_1)$ and backtracking to level 3.
3. Prove the equivalence between the following two programs by replacing the multiplications with uninterrupted functions.

```
int power3(int in) {
int i, out_a;
    out_a = in;
    for (i = 0; i < 2; i++)
        out_a = out_a * in;
    return out_a; }
```

```
int power3_new(int in) {
int out_b;

    out_b = (in * in) * in;
    return out_b; }
```

4. Give the algorithm to convert a formula in Equality Logic to an equisatisfiable formula in Propositional Logic.
5. Consider the following system of constraints.

$$\begin{array}{rrrrr} x_1 & -x_2 & & \leq & 0 \\ x_1 & & -x_3 & \leq & 0 \\ -x_1 & +x_2 & +2x_3 & \leq & 0 \\ & & -x_3 & \leq & -1 \end{array}$$

Use Fourier-Motzkin variable elimination method to check whether the system is satisfiable or not.

Part B

(Answer any 5 questions. Each question carries 7 marks) 5x7 = 35 Marks

6.
   a. Let T_1 and T_2 be two theories whose satisfiability problem is decidable and in the same complexity class. Is the satisfiability problem in a T_1 formula reducible to a satisfiability problem in T_2 formula? Why?
   b. Let T_1 T_2 be two theories whose satisfiability problems are reducible to one another. Are they in the same complexity class? Why?
7. Explain any three decision heuristics used in SAT solvers.
8. Given BDDs for $B = (x_1 \Leftrightarrow x_2)$and $B' = \neg x_2$, compute the BDD for $B \vee B'$.
9. Illustrate how uninterpreted functions can be used for verifying a compilation process with translation validation.
10. Give the algorithm to simplify an equality logic formula. Illustrate with an example.
11. Use general simplex method to check satisfiability of the following set of constraints:

$$\begin{array}{rrl} x & +y & \geq 2 \wedge \\ 2x & -y & \geq 0 \wedge \\ -x & +2y & \geq 1 \ . \end{array}$$

12. A 0-1 integer linear system is an integer linear system in which all variables are constrained to be either 0 or 1. Show how a 0-1 integer linear system can be translated to a Boolean formula. What is the complexity of the translation?

Syllabus

Module 1: Introduction and Basic Concepts

Two approaches to Formal Reasoning, Basic Definitions, Normal forms and their properties, The theoretical point of view, Expressiveness vs. Decidability, Boolean structure in Decision Problems.

Module 2 : Decision Procedures for Propositional Logic

Propositional Logic, SAT Solvers, Binary Decision Diagrams

Module 3 : Equality Logic and Uninterpreted Functions

Introduction, Uninterpreted Functions, From Uninterpreted Functions to Equality Logic, Functional Consistency is not enough, Two examples of the use of Uninterpreted Functions.

Module 4 : Decision Procedures for Equality Logic and Uninterpreted Functions

Congruence Closure, Basic Concepts, Simplification of the formula, A Graph-Based Reduction to Propositional Logic, Equalities and Small Domain Instantiations, Ackermann's vs. Bryant's Reduction.

Module 5 : Linear Arithmetic

Introduction, The Simplex Algorithm, The Branch and Bound Method, Fourier-Motzkin Variable Elimination, The Omega Test, Preprocessing, Difference Logic.

	Course Plan	
No	**Topic**	**No. of Lectures (40 hrs)**
1	**Module 1: Introduction and Basic Concepts**	**6**
1.1	Two approaches to formal reasoning – Proof by deduction, Proof by enumeration, Deduction and enumeration	1
1.2	Basic definitions	1
1.3	Normal forms and their properties (Lecture 1)	1
1.4	Normal forms and their properties (Lecture 1)	1
1.5	The theoretical point of view	1
1.6	Expressiveness vs. Decidability, Boolean structure in decision problems	1
2	**Module 2: Decision Procedures for Propositional Logic**	**8**
2.1	Propositional Logic – Introduction	1
2.2	SAT solvers – Introduction, The *Davis-Putnam-Loveland-Logemann (DPLL)* framework	1
2.3	Boolean Constraints Propagation (BCP) and Implication Graph	1
2.4	Conflict Clauses and Resolution	1
2.5	Decision Heuristics	1
2.6	The Resolution Graph and the Unsatisfiable Core	1
2.7	Binary Decision Diagrams (Lecture 1)	1
2.8	Binary Decision Diagrams (Lecture 2)	1
3	**Module 3: Equality Logic and Uninterpreted Functions**	**7**
3.1	Introduction – Complexity and Expressiveness, Boolean Variables, Removing the Constraints	1
3.2	Uninterpreted Functions – How they are used	1
3.3	Proving equivalence of programs	1
3.4	From Uninterpreted Functions to Equality Logic – Ackermann's Reduction	1
3.5	From Uninterpreted Functions to Equality Logic – Bryant's Reduction	1
3.6	Functional Consistency is not enough	1
3.7	Two examples of the use of Uninterpreted Functions – Proving equivalence of circuits, Verifying a compilation process with Translation Validation	1

4	**Module 4: Decision Procedures for Equality Logic and Uninterpreted Functions**	**8**
4.1	Deciding a conjunction of Equalities and Uninterpreted Functions with Congruence Closure	1
4.2	Basic Concepts	1
4.3	Simplifications of the formula	1
4.4	A Graph-Based Reduction to Propositional Logic	1
4.5	Equalities and Small-Domain Instantiations – Some Simple Bounds, Graph-Based Domain Allocation	1
4.6	The Domain Allocation Algorithm	1
4.7	A Proof of Soundness	1
4.8	Ackermann's vs. Bryant's Reduction: Where does it matter?	1
5	**Module 5: Linear Arithmetic**	**11**
5.1	Basic Definitions, Solvers for Linear Arithmetic	1
5.2	The Simplex Algorithm – The basics	1
5.3	Simplex with Upper and Lower Bounds, Incremental Problems	1
5.4	The Branch and Bound Method	1
5.5	Cutting-Planes	1
5.6	Fourier-Motzkin Variable Elimination	1
5.7	The Omega Test (Lecture 1)	1
5.8	The Omega Test (Lecture 2)	1
5.9	The Omega Test (Lecture 3)	1
5.10	Preprocessing	1
5.11	Difference Logic	1

References

1. Daniel Kroening andOferStrichman. Decision Procedures – An Algorithmic Point of View, Springer, 2008.
2. ChristelBaier and Joost-Pieter Katoen. Principles of Model Checking, The MIT Press.
3. Michael Huth and Mark Ryan. Logic in Computer Science – Modelling and Reasoning about Systems, Cambridge University Press.

PROGRAM ELECTIVE 2

241ECS006	ADVANCED COMPUTER NETWORKS	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course enables the learners to get a good grasp of emerging technologies in the field of computer networks. The syllabus dwells at length on wireless networking, as well as solutions for problems faced while efficiently routing data. Newer networking applications and protocols particularly in multimedia are introduced. The learners are given a glimpse of recent trends in networking like software defined networking. The course enables the learners to analyze network protocols and develop network based applications.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Examine the problem of scalability for routing and also identify the challenges in mobile and multicast routing. **(Cognitive knowledge Level: Analyze)**
CO 2	Choose the technique that provides the Quality-of-Service needs of a particular application. (**Cognitive knowledge Level: Apply**)
CO 3	Survey various wired and wireless networking technologies including wireless cellular technologies. (**Cognitive knowledge Level: Analyze**)
CO 4	Classify the multimedia applications in the Internet and compile the various protocols handling these applications. (**Cognitive knowledge Level: Analyze**)
CO 5	Describe examples of current networking trends and identify the technological gaps. (**Cognitive knowledge Level: Evaluate**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Examine how IPV6 deals with the scalability problem in routing.
2. Distinguish the various approaches in multicast routing.
3. How is the problem of mobility solved in mobile routing?

Course Outcome 2 (CO2)

1. List the categories of service offered by ISA.
2. Examine the role of MPLS in Internet traffic management.
3. Examine the issues affecting network performance and suggest solutions for the same.

Course Outcome 3(CO3):

1. Choose the network technology that can be used to cover areas that cannot support sufficient infrastructure.
2. Show the evolution of cellular technologies from 3G to 5G.
3. Compare the media access techniques of Bluetooth and Zigbee.

Course Outcome 4 (CO4):

1. Categorize the multimedia applications on the Internet and briefly explain their characteristics.
2. Illustrate how real time protocols support interactive applications like VoIP.
3. Justify the need for compressing audio and video before sending it over the Internet.

Course Outcome 5 (CO5):

1. How do overlay networks introduce new functionality into the Internet?
2. Point out the concept behind software defined networking.
3. A new routing protocol is to be implemented in the SDN control plane. Choose the appropriate layer where it should be implemented giving reasons for the same.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES :** 2

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS006

Course Name: Advanced Computer Networks

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Illustrate with an example how standard TCP can be enhanced to support mobile users.	**(5)**
2.	Explain the architectural framework for supporting Quality of Service in packet networks.	**(5)**
3.	Examine the role of core network in 3G cellular data network.	**(5)**
4.	There is one sender and eight receivers in a real time multimedia communication system. If the sender is sending multimedia data at 2 Mbps, how many RTCP packets can be sent by the sender and each receiver in a second? The system allocates 75 percent of the RTCP bandwidth to the receivers and 25 percent to the sender. The average size of each RTCP packet is 125 bytes.	**(5)**
5.	Define OpenFlow specification used in SDN.	**(5)**

		Part B **(Answer any five questions. Each question carries 7 marks)**	
6.	(a)	X, Y, Z are three ASs. X and Z are connected through Y. X has a peering agreement with Y and Y with Z. Z moves all traffic from Y but does not forward traffic from X. Can Z use BGP to implement this policy?	**(4)**
	(b)	How does PIM solve the scalability problem of existing multicast protocols.	**(3)**
7.	(a)	Derive the hexadecimal form of representation of the following link local multicast address: (i) a permanently-assigned multicast group address of 66 (ii) a transient multicast group address of 316	**(4)**
	(b)	A foreign network has a foreign agent. Explain if it is possible for two mobile nodes in the foreign network to use the same care-of address in mobile IP.	**(3)**
8.	(a)	Justify the need for Resource Reservation in multicast transmission.	**(4)**
	(b)	How is VPN implemented using MPLS?	**(3)**
9.	(a)	Elaborate on the various elements of 4G LTE network and the interaction between them.	**(5)**
	(b)	Calculate the minimum time required to download $2*10^6$ bytes using ADSL modem with minimum rate.	**(2)**
10	(a)	Sketch the superframe format of Zigbee 802.15.4 standard.	**(3)**
	(b)	Name some applications which use Zigbee standard and justify its use.	**(4)**
11		Describe H323 architectural model for Internet Telephony.	**(7)**
12	(a)	Comment on the statement “Distributed Hash Tables are said to build structured P2P networks”.	**(4)**
	(b)	Explain Data Center Networking.	**(3)**

Syllabus

Module 1:Advanced Internetworking

The Global Internet, Routing Areas, Interdomain Routing -BGP, IP Version 6, Multicast, Multicast Addresses, Multicast Routing -DVMRP-PIM-MSDP, Routing to a mobile node, Mobile IP, TCP and Mobility, Mobile TCP

Module 2: Internetwork Quality of Service

QoS Architectural Framework - Integrated Services Architecture – RSVP - Differentiated Services, Multiprotocol Label Switching- Destination-Based Forwarding - Explicit Routing

Virtual Private Networks and Tunnels, Performance issues in networks, Delay Tolerant Networking

Module 3: Networking Technologies

Wired: DSL, Cable Networks, SONET, ATM, VLAN, Wireless: Satellite Networks, WiMAX.Cellular Networks: Introduction-Wireless links and Network characteristics -CDMA, Cellular Internet access -An overview of cellular network architecture, 3G cellular data networks, 4G LTE Cellular networks - LTE Protocol Stacks -LTE Radio Access Network -Additional LTE functions, 5G Cellular networks, Managing mobility in cellular networks, Wireless and Mobility-Impact on higher level protocols, Personal Area Networks: Bluetooth, Zigbee

Module 4: Networking Applications

Multimedia in the Internet: Streaming stored audio/video, Streaming live audio/video, Real time interactive audio/video, Real time Interactive Protocols: RTP- RTCP-SIP-H.323, SCTP Compression: Audio Compression, Image compression- JPEG, Video Compression- MPEG

Module 5: Current Topics in Networking

Overlay Networks: Routing overlays -Resilient overlay networks, Peer-Peer Networks – Bit Torrent-Distributed Hash Tables, Content Distribution networks, Software Defined Networks: Architecture – Control and Data Planes – Open Flow – SDN Controllers, Network Function Virtualization, Data Center Networking

Course Plan

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Advanced Internetworking**	**8**
1.1	The Global Internet, Routing Areas	1
1.2	Inter-domain Routing -BGP	1
1.3	IP Version 6	1
1.4	Multicast, Multicast Addresses	1
1.5	Multicast Routing – DVMRP	1
1.6	PIM, MSDP	1
1.7	Routing to a mobile node, Mobile IP	1
1.8	TCP and Mobility, Mobile TCP	1
2	**Module 2: Internetwork Quality of Service**	**8**
2.1	QoS Architectural Framework	1
2.2	Integrated Services Architecture	1
2.3	RSVP - Differentiated Services	1
2.4	Multiprotocol Label Switching,	1

2.5	Virtual Private Networks and Tunnels	1
2.6	Destination-Based Forwarding - Explicit Routing	1
2.7	Performance issues in networks	1
2.8	Delay Tolerant Networking	1
3	**Module 3: Networking Technologies**	**9**
3.1	Wired: DSL, Cable Networks, SONET,	1
3.2	ATM, VLAN	1
3.3	Wireless: Satellite Networks, WiMAX	1
3.4	Cellular Networks: Introduction-Wireless links and Network characteristics -CDMA,	1
3.5	Cellular Internet access-An overview of cellular network architecture, 3G cellular data networks,	1
3.6	4G LTE Cellular networks - LTE Protocol Stacks -LTE Radio Access Network -Additional LTE functions	1
3.7	5G Cellular networks	1
3.8	Managing mobility in cellular networks, Wireless and Mobility-Impact on higher level protocols	1
3.9	Personal Area Networks: Bluetooth, Zigbee	1
4	**Module 4: Networking Applications**	**7**
4.1	Multimedia in the Internet: Streaming stored audio/video, Streaming live audio/video,	1
4.2	Real time interactive audio/video	1
4.3	Real time Interactive Protocols: RTP- RTCP	1
4.4	H-323	1
4.5	SIP, SCTP	1
4.6	Compression: Audio Compression, Image compression- JPEG,	1
4.7	Video Compression- MPEG	1
5	**Module 5: Current Topics in Networking**	**8**
5.1	Overlay Networks: Routing overlays	1
5.2	-Resilient overlay networks,	1
5.3	Peer-Peer Networks – Bit Torrent – Distributed Hash Tables,	1
5.4	Content Distribution networks	1
5.5	Software Defined Networks: Architecture – Control and Data Planes	1
5.6	Open Flow, SDN Controllers	1
5.7	Network Function Virtualization	1
5.8	Data Center Networking	1

References

1. Larry Peterson and Bruce Davie, Computer Networks - A Systems Approach, Morgan Kaufmann, 6th edition, 2022
2. James F. Kurose and Keith W. Ross, Computer Networking A Top-Down Approach, Pearson, 8th edition, 2022
3. Jochen Schiller, Mobile Communications, Addison-Wesley, 2nd edition, 2003
4. William Stallings, Data and Computer Communications, Pearson, 5th edition, 2017
5. Andrew Tanenbaum and David Wetherall, Computer Networks, Pearson, 5th edition, 2010
6. Behrouz A Forouzan, Data Communications and Networking, McGraw Hill, 5th edition, 2017
7. Thomas D. Nadeau and Ken Gray, SDN – Software Defined Networks, O’Reilly, 2013

241ECS007	PATTERN RECOGNITION	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course aims to impart the fundamentals of statistical pattern recognition and neural network techniques. It introduces to the learner the various pattern recognition algorithms, feature selection, classification, clustering and the use of neural networks in feature extraction. This helps the learner to apply the algorithms in applications that works on pattern recognition and machine intelligence.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply probability and numerical methods in statistical pattern recognition. **(Cognitive Knowledge Level: Apply)**
CO 2	Apply statistical methods in feature selection.**(Cognitive Knowledge Level: Apply)**
CO 3	Apply linear algebra and statistical methods in parameter and non-parameter estimation.**(Cognitive Knowledge Level: Apply)**
CO 4	Apply the technique of decision trees in pattern recognition. **(Cognitive Knowledge Level: Apply)**
CO 5	Analyze the use of deep learning networks andartificial neural networks in pattern recognition. **(Cognitive Knowledge Level: Analyze)**
CO 6	Design, Develop, Implement and Present innovative ideas in problem solving with various pattern recognition techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	

CO 2	✔		✔	✔	✔	✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔		✔	✔	✔	✔	
CO 6	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii.Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (suchquestions shall be useful in the testing of overall achievement and maturity of thestudents in a course, through long answer questions relating totheoretical/practical knowledge, derivations, problem solving and quantitativeevaluation), with minimum one question from each module of which studentshould answer

any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Illustrate the design cycle of a pattern recognition system with the help of an example.
2. Suppose that we have three coloured bottles r (red), b(blue) and g(green).Box r contains 3 apples, 4 oranges and 3 limes. Box B contains 1 apple, 1 orange and 0 limes and box g contains 3 apples, 3 oranges and 4 limes. If a box is chosen at random with probability p(g)=0.2, p(b)=0.2 and p(g)=0.6 and piece of fruit is removed from the box(with equal probability of selecting items from the box), then what is the probability of selecting an apple? If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

Course Outcome 2 (CO2):

1. Illustrate feature selection using t-Test with the help of an example.

Course Outcome 3(CO3):

1. Derive the fuzzy C spherical shells (FCSS) algorithm for the case that spherical clusters are to be identified.

Course Outcome 4 (CO4):

1. Illustrate decision tree with the help of an example. How does it enable pattern classification?

2. Construct a decision tree using the following data.

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Cloudy	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No

Cloudy	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes
Cloudy	Mild	High	True	Yes
Cloudy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No

Course Outcome 5 (CO5):

1. How do artificial neural networks play a significant role in pattern recognition? Also discuss about its parameter optimisation techniques.

Course Outcome 6 (CO6):

1. Suppose an accident prone area is under surveillance and real time CCTV visuals are available to you. Design a solution to automatically detect accidents on the road from those real time CCTV visuals. Explain about any one pattern recognition algorithm you will make use here and how?

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES: 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS007

Course Name: Pattern Recognition

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. In a town it was estimated that 3% of people have a particular disease. Adiagnosis test was conducted for all the people, which yielded 8% false positiveand 92% true positive results. A person is found as positive after the test. What isthe probability that this person is truly having the disease? **(5)**
2. How does morphological operations play a role in pattern recognition? **(5)**
3. How can visual imagery be analysed using convolutional neural networks? **(5)**
4. How does a decision tree handle continuous attributes? **(5)**
5. Define the terms: weights, bias, activations with respect to neural networks **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. Illustrate the design principles of pattern recognition system with an example. **(7)**

7. Derive the fuzzy C spherical shells (FCSS) algorithm for the case that spherical clusters are to be identified. **(7)**

8. Show that in the case of Gaussian distributions the Chernoff bound becomes **(7)**

$$\epsilon_{CB} = \exp(-b(s))$$

where

$$b(s) = \frac{s(1-s)}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T [s\Sigma_j + (1-s)\Sigma_i]^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \frac{1}{2}\ln \frac{|s\Sigma_j + (1-s)\Sigma_i|}{|\Sigma_j|^s |\Sigma_i|^{1-s}}$$

Then take the derivative with respect to s and show that for equal covariance matrices the optimum is achieved for $s = 1/2$. Thus, in this case $b(s)$ equals the Bhattacharyya distance.

9. Let N_1, N_2 be the available values of a feature in two classes, respectively. The feature is assumed to follow a Gaussian distribution with the same variance in each class. Define the test statistic **(7)**

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s_z\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

where

$$s_z^2 = \frac{1}{N_1 + N_2 - 2}\left(\sum_{i=1}^{N_1}(x_i - \bar{x})^2 + \sum_{i=1}^{N_2}(y_i - \bar{y})^2\right)$$

and μ_1, μ_2 are the respective true mean values. Show that q follows the t-distribution with $N_1 + N_2 - 2$ degrees of freedom.

10. Discuss the significance of pre-processing in feature selection. Illustrate any two methods used for pre-processing. **(7)**

11. How can artificial neural networks be applied in Pattern recognition? Also illustrate the features of recurrent neural networks. **(7)**

12. Construct a decision tree using the following data. **(7)**

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Cloudy	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Cloudy	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes
Cloudy	Mild	High	True	Yes
Cloudy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No

Syllabus

Module 1: Introduction to Pattern Recognition

Basics of pattern recognition systems, various applications, Machine Perception, classification of pattern recognition systems.Design of Pattern recognition system, Pattern recognition Life Cycle.

Statistical Pattern Recognition: Review of probability theory, Gaussian distribution. Normal density and discriminant functions.

Module 2: Feature Selection

Feature selection – Outlier removal – Data normalization – Missing data, The Peaking phenomenon, Feature selection using statistical hypothesis testing- Hypothesis testing basics – Application of t-Test in feature selection. Class separability measures-Divergence-Chernoff bound and Bhattacharya distance-Scatter matrices, Feature subset selection –Scalar feature selection, Feature vector selection.

Module 3: Clustering Algorithms

Unsupervised learning and clustering - Criterion functions for clustering. Cluster validation.

Fuzzy clustering algorithms- Point representatives- quadratic surfaces and representatives – hyper plane representatives. Binary morphology clustering algorithms (BMCAs) – Discretization – Morphological operations - Determination of clusters in a discrete binary set- Assignment of feature vectors to clusters – The algorithmic scheme, Boundary detection algorithms.

Module 4: Dimensionality reduction

Dimensionality reduction: Principal component analysis - its relationship to Eigen analysis. Fisher discriminant analysis - Generalised Eigen analysis. Eigen vectors/Singular vectors as dictionaries. Factor Analysis, Total variability space - a dictionary learning method.Non negative matrix factorisation - a dictionary learning method.

Linear discriminant functions: Gradient descent procedures, Perceptron.

Module 5: Artificial neural networks and Pattern Classification

Artificial neural networks: Review of Artificial neural network concepts, convolutional neural networks, recurrent neural networks.

Non-metric methods for pattern classification: Non-numeric data or nominal data. Decision trees: Classification and Regression Trees (CART).

Course Plan

No.	Topic	No. of Lectures (**40 Hours**)
1	**Module 1: Introduction to Pattern Recognition**	**7**
1.1	Basics of pattern recognition systems, applications	1
1.2	Machine Perception, Classification of pattern recognition systems	1
1.3	Design of Pattern recognition system	1
1.4	Pattern recognition Life Cycle	1
1.5	Statistical Pattern Recognition	1
1.6	Review of probability theory	1
1.7	Normal density and discriminant functions	1
2	**Module 2: Feature Selection**	**10**
2.1	Feature selection – Outlier removal	1

2.2	Data normalization – Missing data	1
2.3	The peaking phenomenon	1
2.4	Feature selection using statistical hypothesis testing	1
2.5	Hypothesis testing basics – Application of tTest in feature selection	1
2.6	Class separability measures-Divergence	1
2.7	Chernoff bound and Bhattacharya distance	1
2.8	Scatter matrices	1
2.9	Feature subset selection –Scalar feature selection	1
2.10	Feature vector selection	1
3	**Module 3: Clustering Algorithms**	**9**
3.1	Unsupervised learning and clustering	1
3.2	Criterion functions for clustering. Cluster validation.	1
3.3	Fuzzy clustering algorithms- Point representatives	1
3.4	Quadratic surfaces and representatives – hyper plane representatives.	1
3.5	Binary morphology clustering algorithms (BMCAs)	1
3.6	Discretization	1
3.7	Morphological operations - Determination of clusters in a discrete binary set	1
3.8	Assignment of feature vectors to clusters	1
3.9	The algorithmic scheme, Boundary detection algorithms.	1
4	**Module 4: Dimensionality reduction**	**8**
4.1	Principal component analysis - its relationship to Eigen analysis	1
4.2	Fisher discriminant analysis	1
4.3	Generalised Eigen analysis	1
4.4	Eigen vectors/Singular vectors as dictionaries	1
4.5	Total variability space - a dictionary learning method	1
4.6	Non negative matrix factorisation - a dictionary learning method	1
4.7	Linear discriminant functions: Gradient descent procedures	1
4.8	Perceptron	1
5	**Module 5: Artificial neural networks and Pattern Classification**	**6**
5.1	Review of Artificial neural networks, Introduction to deep neural networks	1
5.2	Convolutional neural networks	1
5.3	Recurrent neural networks	1
5.4	Non-metric methods for pattern classification: Non-numeric data or nominal data	1
5.5	Decision trees: Classification and Regression Trees (CART) lecture 1	1
5.6	Decision trees: Classification and Regression Trees (CART) lecture 2	1

References

1. S.Theodoridis and K.Koutroumbas, "Pattern Recognition", 4th Ed., Academic Press, 2009
2. C.M.Bishop, "Pattern Recognition and Machine Learning", Springer, 2006
3. R.O.Duda, P.E.Hart and D.G.Stork, "Pattern Classification", John Wiley, 2001
4. Hastie, T., Tibshirani, R. and Friedman, J. "The Elements of Statistical Learning". Springer. 2001.

241ECS008	ADVANCED COMPUTER ARCHITECTURE	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This purpose of this course is to provide a solid foundation that furnishes the learner with in-depth knowledge of current and emerging trends in computer architectures, focusing on performance and the hardware/software interface. This course covers design and analysis, memory hierarchy, pipelining, operation of multiprocessors, thread level parallelism, and data level parallelism. This course helps the learner to develop software/hardware applications based on architectural framework.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Identify and solve the advanced issues in design of computer processors, caches and memory**(Cognitive Knowledge Level: Apply)**
CO 2	Analyze the memory hierarchy design, performance improvement techniques and cache optimization techniques**(Cognitive Knowledge Level: Analyze)**
CO 3	Analyze the working and features of branching and exception handling in pipeline architecture**(Cognitive Knowledge Level: Analyze)**
CO 4	Analyze the operation of multiprocessors and thread level parallelism**(Cognitive Knowledge Level: Evaluate)**
CO 5	Demonstrate the concepts of data level parallelism including SIMD and GPU processors**(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, Implement and Present innovative ideas on advanced computer architecture and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1			✓	✓	✓	✓	
CO 2			✓	✓	✓	✓	
CO 3			✓	✓	✓	✓	
CO 4			✓	✓		✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly. For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Identify the various techniques for instruction encoding. Illustrate with examples.

Course Outcome 2 (CO2):

1. Consider an in-order execution computer. Assume that the cache miss penalty is 200 clock cycles, and all instructions normally take 1.0 clock cycles. Assume that the average miss rate is 2%, there is an average of 1.5 memory references per instruction, and the average number of cache misses per 1000 instructions is 30. What is the impact on performance when behaviour of the cache is included? Calculate the impact using both misses per instruction & miss rate.

Course Outcome 3(CO3):

1. Consider the execution of following instructions, on our pipelined example processor:

   ```
   ADD R1, R2, R3
   SUB R4, R1, R5
   AND R6, R1, R7
   OR R8, R1, R9
   XOR R10, R1, R11
   ```

   Analyze type of hazards may occur in the above code? If hazard exists, explain how we can solve it.

Course Outcome 4 (CO4):

1. Determine the limitations in symmetric shared memory multiprocessors.

Course Outcome 5 (CO5):

1. Identify the data dependencies between the statements S1 and S2 in the loop.

   ```
   for (i=1; i<=100; i=i+1) {
   A[i+1] = A[i] + C[i]; /* S1 */
   B[i+1] = B[i] + A[i+1]; /* S2 */
   ```

}

Course Outcome 6 (CO6):

1. Implement pipelining and various level parallelisms using tools like OPENMP.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES: 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS008

Course Name: Advanced Computer Architecture

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Impact of optimization improves the performance of complier. Justify your answer. Also mention types of optimization. **(5)**

2. Determine whether a 32 KB four-way set associative L1 cache has a faster memory access time than a 32 KB two-way set associative L1 cache. Assume the miss penalty to L2 is 15 times the access time for the faster L1 cache. Ignore misses beyond L2. Which has the faster average memory access time? (Miss rate for two-way set associative cache is 0.038 and four-way set associative cache is 0.037) **(5)**

3. Analyze the type of hazards may occur in the following code. **(5)**

```
LW R1, 0(R2)
SUB R4, R1
AND R6, R1, R7
OR R8, R1, R9
```

4. Suppose we have an application running on a 32-processor multiprocessor, which has a 200ns time to handle reference to a remote memory. For this application, assume that all the references except those involving communication hit in the local memory hierarchy, which is slightly optimistic. Processors are stalled on a remote request, and the processor clock rate is 2GHz. If the base CPI (assuming that all references hit in the cache) is 0.5, evaluate how much faster is the multiprocessor if there is no communication versus if 0.2% of the instructions involve a remote communication reference? **(5)**

5. Consider the following loops, identify the true dependencies, output dependences and anti-dependences and eliminate the output dependences and anti-dependences. **(5)**

```
for(i=0;i<100;i++){
Y[i]=X[i]/C; /*S1*/
X[i]= X[i]+C; /*S2*/
Z[i]=Y[i]+C; /*S3*/
Y[i]=C-Y[I]; /*S4*/
}
```

Part B

(Answer any five questions. Each question carries 7 marks)

6. A benchmark program is executed on a 40MHZ processor. The benchmark program has the following statistics. Determine the effective CPI, MIPS rate and execution time of this program. **(7)**

Instruction Type	Instruction Count	Clock Cycle Count
Arithmetic	45000	1
Branch	32000	2
Load/Store	15000	2
Floating Point	8000	2

7. "Fully associative caches do not have conflict misses". Examine the statement. **(7)**

8. Consider an in-order execution computer. Assume that the cache miss penalty is 200 clock cycles, and all instructions normally take 1.0 clock cycles. Assume that the average miss rate is 2%, there is an average of 1.5 memory references per instruction, and the average number of cache misses per 1000 instructions is 30. What is the impact on performance when behaviour of the cache is included? Calculate the impact using both misses per instruction and miss rate. **(7)**

9. Explain the methodologies used for reducing the branch cost with prediction in instruction level parallelism. **(7)**

10. Illustrate and explain extending the MIPS pipeline to handle multicycle operations. **(7)**

11. Explain multiprocessor cache coherence. **(7)**

12. Compare and contrast multimedia SIMD computers and GPUs. **(7)**

Syllabus

Module 1: Design and Analysis

Principles of computer design, Fallacies and Pitfalls, Instruction Set Principles- Classifying instruction set architecture, Memory addressing, Type and size of operands, Operations in the instruction set, Instruction for control flow, Encoding an instruction set, Role of compiler.

Module 2: Memory Hierarchy

Introduction, Cache performance, Basic cache optimizations, Virtual memory–Techniques for fast address translation, Protection via virtual memory, Fallacies and Pitfalls, Case study of Pentium/Linux memory system-Pentium address translation.

Module 3: Pipelining

Introduction, Pipeline hazards, Static branch prediction and dynamic branch prediction, Implementation of MITS, Basic pipeline of MITS, Implementing the control in MITS pipeline, Dealing with branches in pipeline, Dealing with exceptions, Handling of multi-cycle operations, Maintaining precise exceptions, Case study of MITS R4000 pipeline.

Module 4: Thread Level Parallelism

Introduction, Centralized Shared-Memory Architectures, Performance of Symmetric Shared-Memory Multiprocessors, Distributed Shared-Memory and Directory-Based Coherence, Synchronization: The Basics, Models of Memory Consistency: An Introduction, Crosscutting Issues, Case study of Sun T1 Multiprocessor.

Module 5: Data Level Parallelism

Vector architecture, SIMD instruction set, Extension for multimedia, Graphic Processing Units, Case study Envida GPU instruction set architecture, GPU memory structure, Innovations in GPU architecture, Comparisons between vector architecture and GPUs, Comparisons between multimedia SIMD computers and GPUs, Loop level parallelism, Finding dependencies, Eliminating dependencies.

Course Plan

No	Topic	No. of Lectures (**40 Hours**)
1	**Module 1: Design and Analysis**	**8**
1.1	Principles of computer design	1
1.2	Fallacies and Pitfalls	1
1.3	Instruction Set Principles- Classifying instruction set architecture	1
1.4	Memory addressing, Type and size of operands	1
1.5	Operations in the instruction set	1
1.6	Instruction for control flow	1
1.7	Encoding an instruction set	1
1.8	Role of compiler	1
2	**Module 2: Memory Hierarchy**	**8**
2.1	Introduction	1
2.2	Pipeline hazards	1
2.3	Static branch prediction and dynamic branch prediction	1
2.4	Implementation of MITS, Basic pipeline of MITS	1
2.5	Implementing the control in MITS pipeline	1

2.6	Dealing with branches in pipeline, Dealing with exceptions	1
2.7	Handling of multi-cycle operations, Maintaining precise exceptions	1
2.8	Case study of MITS R4000 pipeline	1
3	**Module 3: Multiprocessors and Thread level Parallelism**	**8**
3.1	Introduction	1
3.2	Centralized Shared-Memory Architectures	1
3.3	Performance of Symmetric Shared-Memory Multiprocessors	1
3.4	Distributed Shared-Memory and Directory-Based Coherence	1
3.5	Synchronization: The Basics	1
3.6	Models of Memory Consistency: An Introduction	1
3.7	Crosscutting Issues	1
3.8	Case study Sun T1 Multiprocessor	1
4	**Module 4: Multiprocessors and Thread level Parallelism**	**8**
4.1	Introduction	1
4.2	Centralized Shared-Memory Architectures	1
4.3	Performance of Symmetric Shared-Memory Multiprocessors	1
4.4	Distributed Shared-Memory and Directory-Based Coherence	1
4.5	Synchronization: The Basics	1
4.6	Models of Memory Consistency: An Introduction	1
4.7	Crosscutting Issues	1
4.8	Case study Sun T1 Multiprocessor	1
5	**Module 5: Data Level Parallelism**	**8**
5.1	Vector architecture, SIMD instruction set	1
5.2	Extension for multimedia, Graphic Processing Units	1
5.3	Case study Envida GPU instruction set architecture	1
5.4	GPU memory structure	1
5.5	Innovations in GPU architecture, Comparisons between vector architecture and GPUs	1
5.6	Comparisons between multimedia SIMD computers and GPUs	1
5.7	Loop level parallelism	1
5.8	Finding dependencies, Eliminating Dependencies	1

References

1. Hennessy J.L and David A. Patterson “Computer Architecture- A Quantitative Approach” Morgan Kaufmann Publication, Fifth edition, 2002.
2. Randal E Bryant and David O'Hallaron “Computer Systems A programmer's perspective” Pearson Education, 2nd edition 2010.
3. Kaihwang and Naresh Jotwani, “Advanced Computer Architecture” 2nd edition Tata Mcgraw-Hill, 2010.
4. Sima D,Fountain T and Kacsuk P “Advanced Computer Architecture: A Design Space Approach” Pearson Education, 1st edition 1997.

241ECS009	NATURAL LANGUAGE PROCESSING AND TEXT MINING	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course provides an exposure to the concepts and techniques in Natural language processing and Text mining. Fundamental concepts and practical applications of Natural Language Processing (NLP) are covered in this course. This helps the learners to analyze and interpret textual data.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply different approaches of syntax and semantics in NLP. **(Cognitive Knowledge Level: Apply)**
CO 2	Employ approaches to generate dialogue and summarisation within NLP. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply different statistical approaches to machine translation. **(Cognitive Knowledge Level: Apply)**
CO 4	Research, analyze and deploy appropriate machine learning techniques in NLP including hidden Markov models and unsupervised methods. **(Cognitive Knowledge Level: Analyze)**
CO 5	Use text mining concepts and methods to model real-world problems and develop technical solutions. **(Cognitive Knowledge Level: Analyze)**
CO6	Design, develop and implement NLP and text mining methods to solve real world problems. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓				✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyse	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Consider the following Corpus

   - Mid priced food is what I am looking for.
   - Tell me about Thai food.
   - Can you give me a list of Thai food?
   - I am looking for a good restaurant for Thai food.

   Draw the Raw Bigram count table for the corpus and then Find the bigram probability of the sentence " I am looking for Thai food."
2. Write the Regular Expression for the following

   a. To find the all occurrence of the word **hood** in a text

   b. To find any line in which a particular word **country** appears twice

   c. To find similar pattern like ***The bigger they were, the bigger they will be*** or ***The smaller they were, the smaller they will be.***
3. Find the minimum edit distance between the string INTENTION and EXECUTION using dynamic programming.

Course Outcome 2 (CO2)

1. Consider a document collection having 37 documents. Find the cosine similarity of the words "Fool and Wit"

Words	Document Frequency
Romeo	1
Salad	2
Falstaff	4
Forest	12
Battle	21
Wit	34
Fool	36
Good	37
Sweet	37

Raw Word Count	Doc1	Doc2	Doc3	Doc4
Battle	1	0	7	13
Good	114	80	62	89
Fool	36	58	1	4
Wit	20	15	2	3

2. Explain the SND method of word embeddings generation? What are its Limitations?

Course Outcome 3(CO3):

1. Specify the roles of Named Entity Recognition in Question Answering systems?
2. Using Transition probabilities and Emission probabilities given below, POS tag the sentence **Files like a Flower** with HMM tagger

   <s> is he start symbol and wherever the transition probability is not explicitly given in the table assume the value 0.0001

Category	Bigram	Estimate
<s>	PROB(ART|<s>)	71
<s>	PROB(N|<s>)	29
ART	PROB(N|ART)	1
N	PROB(V|N)	43
N	PROB(N|N)	13

N	PROB(P|N)	44
V	PROB(N|V)	35
V	PROB(ART|V)	65
P	PROB(ART|P)	74
P	PROB(N|P)	26

PROB(the |ART)	0.54	PROB(a|ART)	0.360
PROB(files|N)	0.025	PROB(a|N)	0.001
PROB(files|V)	0.076	PROB(flowers|N)	0.063
PROB(like|V)	1	PROB(flowers|V)	0.05
PROB(like|N)	0.012		

3. Explain how a feed forward network can be used for sentiment analysis.

Course Outcome 4 (CO4):

1. Whether Bidirectional RNN can be used for sentiment classification? Justify.
2. Draw the architecture of a RNN that can be used for sequence classification.

Course Outcome 5 (CO5):

1. Consider the following sentence:
   "You are very kind and beautiful "
   Explain the procedures to convert the above sentence into Malayalam language using sequence to sequence model with attention mechanism. Is there any better method we can opt for? Analyse different methods you can adopt in this scenario and justify your better option.
2. You are developing a text mining application utilized for assisting organizations with finding their contacts, reports, business correspondence, emails, and so on. Which Text mining algorithm will be best suitable in your application? Justify.

Course Outcome 6 (CO6):

1. Implement an automatic question answering system for farmers.

2. Implement extractive summary preparation from medical reports.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS009

Course Name: NATURAL LANGUAGE PROCESSING AND TEXT MINING

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. For each sentence, identify whether the different meanings arise from structural ambiguity, semantic ambiguity or pragmatic ambiguity? **(5)**
1. Time flies like an arrow
2. He crushed the key to my heart

2. The confusion matrix for a binary classifier is given below. Compute Precision, Recall, F- score, Specificity and Accuracy of the model **(5)**

Actual	Predicated	
	Class 1	Class 2
Class 1	14	2
Class 2	8	40

3. Justify whether a model/ program designed to perform sentimental analysis can be adapted for document grouping **(5)**

4. Describe the structure of GRU cell. How the output is computed at each gate? **(5)**

5. Produce the extractive summary by creating a document of your own with the title as "Kerala- at a glance". Evaluate the performance of your summarizer by using the appropriate evaluation measures. **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) Define minimum edit distance. **(2)**

(b) Compute the minimum edit distance between the strings INTENTION and EXECUTION using dynamic programming. **(5)**

7. (a) Consider a document collection having 37 documents **(5)**

Find the TF-IDF embedding of the following words, "Battle, Good, Fool, Wit" using the tables given below

Words	Document Frequency
Romeo	1
Salad	2
Falstaff	4
Forest	12
Battle	21
Wit	34
Fool	36
Good	37
Sweet	37

Raw Word Count	Doc1	Doc2	Doc3	Doc4
Battle	1	0	7	13
Good	114	80	62	89
Fool	36	58	1	4
Wit	20	15	2	3

(b) Find the cosine similarity off the words "Fool & Wit" **(2)**

8. Explain the Skip-gram and CBOW methods of word embedding generation. **(7)**

9. (a) Identify the Named Entities in the text given below and tag them with appropriate NE tags **(4)**

Citing high fuel prices, Indian Airlines said Friday it has increased by 56 per round trip on flights to some cities also served by lower-cost carriers. Spice Jet, a unit of AMR Corp, immediately matched the move, Spokesman Ram Goel said Indian Airlines, a unit of TATA Corp, said the increase took effect Thursday and applies to most routes where it competes against discount carriers, such as cochin to Delhi and Chennai to Calcutta.

(b) Why NE tagging is considered hard when compared to POS tagging? Give one method to overcome it. **(3)**

10. Describe the architecture of a RNN that can perform sequence classification. **(7)**

11. Explain in detail how LSTM can be used for language translation. **(7)**

12. (a) What application of text mining is coming under the Social Media Platform YouTube? Justify your answer. **(3)**

(b) Consider that you are running a business and you are having a social media page to promote your business. Is it possible to design a sentiment analyser by using the raw data extracted from the promotion page? Construct an architecture for the same. **(4)**

Syllabus

Module 1: Introduction

Natural Language Processing (NLP) - Syntax, semantics, pragmatics, and ambiguity in NLP, Regular Expressions, Text Normalisation, Edit Distance.

N-gram Language Models-N-Grams, Evaluating Language Models, Generalisation and Zeros, Smoothing, Kneser-Ney Smoothing, The Web and Stupid Backoff, Perplexity's Relation to Entropy.

Module 2: Neural Language Models, Vector Semantics and Embeddings

Neural Networks and Neural Language Models-Units, Feed-Forward Neural Networks, Training Neural Nets, Neural Language Models.

Vector Semantics and Embeddings-Lexical Semantics, Vector Semantics, Words and Vectors, Cosine for measuring similarity, TF-IDF: Weighing terms in the vector, Applications of the tf-idf vector model, Word2vec, Visualizing Embeddings, Semantic properties of embeddings, Bias and Embeddings, Evaluating Vector Models.

Module 3: Sentiment Classification and Part-of-Speech Tagging

Sentiment Classification –Sentiment classification. Machine Learning for Sentiment Classification - Training the Classifier (Naive Bayes, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest), Optimising for Sentiment Analysis - Other text classification tasks – Evaluation of classification models: Precision, Recall, F-measure, Test sets and Cross-validation, Statistical Significance Testing.

Part-of-Speech Tagging-English Word Classes, The Penn Treebank Part-of-Speech Tagset, Part-of-Speech Tagging, HMM Part-of-Speech Tagging, Maximum Entropy Markov Models, Bi-directionality, Part-of Speech Tagging for Morphological Rich Languages. Information Extraction-Named Entity Recognition, Relation Extraction, Extracting Times, Extracting Events and their Times, Template Filling.

Module 4: Sequence Processing with Recurrent Networks

Sequence Processing with Recurrent Networks-Simple Recurrent Neural Networks, Applications of Recurrent Neural Networks, Deep Networks: Stacked and Bidirectional RNNs, Managing Context in RNNs: LSTMs and GRUs, Words, Subwords and Characters Neural Language Models and Generation Revisited, Encoder-Decoder Networks, Attention, Applications of Encoder-Decoder Networks. Case study: Machine translation, Question Answering

Module 5: Text Mining

Document representation - representing unstructured text documents with appropriate format and structure, automated text mining algorithms.

Text Mining: Text categorization, Text clustering, Topic modeling, Applications - classification, image annotation, collaborative filtering, and hierarchical topical structure modeling. Document summarization - Extraction- based summarization methods

Sentiment analysis - concept, sentiment polarity prediction, review mining, aspect identification. Text visualization - introduction to mathematical and programming tools.

Course Plan

No	Topic	No. of Lectures (40)
1	**Module 1: Introduction**	**6**
1.1	Natural Language Processing (NLP) - Syntax, semantics, pragmatics, and ambiguity in NLP	1
1.2	Regular Expressions	1
1.3	Text Normalisation, Edit Distance	1
1.4	N-gram Language Models-N-Grams, Evaluating Language Models	1
1.5	Generalisation and Zeros, Smoothing	1
1.6	Kneser-Ney Smoothing, The Web and Stupid Backoff, Perplexity’s Relation to Entropy	1
2	**Module 2: Neural Language Models, Vector Semantics and Embeddings**	**8**
2.1	Units, Feed-Forward Neural Networks	1
2.2	Training Neural Nets	1
2.3	Neural Language Models	1
2.4	Lexical Semantics, Vector Semantics	1
2.5	Words and Vectors, Cosine for measuring similarity	1
2.6	TF-IDF: Weighing terms in the vector, Applications of the tf-idf vector model	1

2.7	Word2vec, Visualizing Embeddings	1
2.8	Semantic properties of embeddings, Bias and Embeddings, Evaluating Vector Models	1
3	**Module 3: Sentiment Classification and Part-of-Speech Tagging**	**10**
3.1	sentiment classification. Machine Learning for Sentiment Classification	1
3.2	Training the Classifier (Naive Bayes, Logistic Regression)	1
3.3	Training the Classifier (Support Vector Machine, Decision Tree, Random Forest)	1
3.4	Optimising for Sentiment Analysis - Other text classification tasks	1
3.5	Evaluation of classification models: Precision, Recall, F-measure, Test sets and Cross-validation, Statistical Significance Testing.	1
3.6	English Word Classes, The Penn Treebank Part-of-Speech Tagset	1
3.7	Part-of-Speech Tagging, HMM Part-of-Speech Tagging,	1
3.8	Maximum Entropy Markov Models, Bi-directionality, Part-of Speech Tagging for Morphological Rich Languages.	1
3.9	Information Extraction-Named Entity Recognition, Relation Extraction	1
3.10	Extracting Times, Extracting Events and their Times, Template Filling	1
4	**Module 4: Sequence Processing with Recurrent Networks**	**7**
4.1	Simple Recurrent Neural Networks, Applications of Recurrent Neural Networks	1
4.2	Deep Networks: Stacked and Bidirectional RNNs	1
4.3	Managing Context in RNNs: LSTMs	1
4.4	GRUs, Words, Subwords and Characters	1
4.5	Neural Language Models and Generation Revisited	1
4.6	Encoder-Decoder Networks, Attention, Applications of Encoder-Decoder Networks.	1
4.7	Case study: Machine translation, Question Answering	1

5	**Module 5: Text Mining**	**9**
5.1	Representing unstructured text documents with appropriate format and structure	1
5.2	automated text mining algorithms	1
5.3	Text categorization	1
	Text clustering	1
5.5	Topic modelling, Applications - classification, imagine annotation, collaborative filtering, and hierarchical topical structure modeling	1
5.6	Document summarization - Extraction- based summarization methods	1
5.7	Sentiment analysis - concept, sentiment polarity prediction	1
5.8	Review mining, aspect identification.	1
5.9	Text visualization - introduction to mathematical and programming tools.	1

References

1. Daniel Jurafsky and James H. Martin. Speech and Language Processing (2nd ed), Pearson International edition, 2008
2. Manning C, Schuetze H. Foundations of Statistical Natural Language Processing, MIT Press, 1999
3. James Allen, "Natural Language Understanding", 2/E, Addison-Wesley, 1994
4. Charu C. Aggarwal and Cheng Xiang Zhai, “Mining Text Data”, Springer, 2012.
5. Text Mining Classification, Clustering, and Applications - Ashok N. Srivastava, Mehran Sahami, CRC Press.

241ECS010	ADVANCED COMPILER DESIGN	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course enables the students to analyze the different phases of compiler / techniques for designing a compiler. This course introduces students to the advanced concepts of compilation phases such as lexical analysis, syntax analysis, semantic analysis, intermediate code generation, code optimization and code generation. This course helps the learners to design and develop compilers for programming languages.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Illustrate lexical rules and grammars for a representative programming language **(Cognitive Level: Understand)**
CO 2	Construct intermediate code representations and code optimization techniques **(Cognitive Level: Apply)**
CO 3	Experiment with register allocation strategies and code scheduling **(Cognitive Level: Apply)**
CO 4	Inspect programming language design, target machine language design and run time environment of compilers**(Cognitive Level: Analyze)**
CO 5	Assess recent trends in compiler design and build a compiler for a hypothetical language**(Cognitive Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✔		✔	✔	✔	✔	
CO 2	✔		✔	✔	✔	✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔		✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly. For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain context free grammar representation methods.

Course Outcome 2 (CO2):

1. Construct the translation of Boolean expressions into three-address code.

Course Outcome 3(CO3):

1. Build graph coloring global register allocator**.**

Course Outcome 4 (CO4):

1. Discover challenges in instruction level parallelism.

Course Outcome 5 (CO5):

1. Discuss loop optimization techniques in IBM XLCompilers.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES: 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS010

Course Name: ADVANCED COMPILER DESIGN

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Explain different implementation methods of Symbol table. **(5)**

2. With an example explain the following loop optimization techniques: **(5)**

a) Code motion.

b) Induction variable elimination and

c) Strength reduction

3. Explain graph coloring global register allocator. **(5)**

4. Explain reference counting garbage collector. **(5)**

5. Illustrate the role of peep hole optimization in the compilation process **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. Translate the following statements into three address statements and construct the flow graph. **(7)**

```
for(i=0;i<n;i++)
for(j=0;j<n;j++)
for(k=0;k<n;k++)
        C[i][j]=c[i][j]+a[i][k]*b[k][j]
```

7. Explain the translation of Boolean expressions into three-address code. **(7)**

8. Illustrate the liveness analysis on the flow graph with suitable example? **(7)**

9. Explain Dominators and Algorithm for finding Dominators. **(7)**

10. Explain static allocation and heap allocation strategies. Construct the Directed Acyclic Graph for the basic block given below and simplify the three address code **(7)**

```
d = b *
ce=a+b
b =b *c
a = e –d
```

11. Illustrate loop invariant code motion with the following code segment. **(7)**

```
b= 2
i = 1
L1: if i>100 gotoL3
      a = b +1
      c =2
If i mode 2 == 0 goto L2
     d = c
      f= a + i
i = i + 1
goto  L1
L2: d = a + d
       e = i + d
L3 : end
```

12. Generate a code sequence for the assignment **d=(a-b)+(a-c)+(a-c).** Also define the use of algebraic identities in optimization of basic blocks? **(7)**

Syllabus

Module 1: Overview of Compiler Design

Introduction - The phases of Compiler - Lexical Analysis - Role of Lexical Analyzer - Specification and Recognition of Tokens - Context Free Grammar – Symbol - Table Structure, Symbol Attributes and Symbol - Table Entries, Local Symbol - Table Management, Global Symbol - Table Structure, Storage Binding and Symbolic Registers, Approaches to Generating Loads and Stores. Intermediate representation – Issues – High level, medium level, low level intermediate languages –MIR, HIR, LIR – ICANfor Intermediate code.

Module 2: Intermediate Representations

The value - number method for constructing DAGs - Addresses and instructions – Quadruples - Triples - Storage organization – Static versus dynamic storage allocation - stack allocation of space -Activation trees - Activation records - Garbage collection - Design goals for Garbage collectors -Reference counting garbage collectors - Introduction to trace - based collection - A basic mark - and-sweep collector. Translation of expressions Translation of expressions – Operations within expressions – Incremental translation – Addressing array elements – Translation of array references Control flow – Boolean expressions – Short - circuit code – flow – of - control statements – Control flow translation of Boolean expressions – Avoiding redundant Gotos – Boolean values and jumping code – Backpatching – One - passcode generation using back patching – Back patching for Boolean expressions – Flow-of-control statements – Break, continue and Goto statements Translation of switch statements – syntax directed translation of switch statements - intermediate code for procedures

Module 3: Code Optimization

Principal sources of optimization - causes of redundancy preserving transformations - Global common subexpressions - Copy Propagation - Dead code elimination – Code motion –Upward code motion – Downward code motion – Induction variables and reduction in strength -Introduction to data flow analysis - Loops in Flow graphs – Dominators - Introduction to global data flow analysis - Points and Paths - Reaching definitions - Live variable analysis - Data flow analysis of structured program.

Module 4: Register Allocation and Code Scheduling

Register allocation and assignment – graph coloring – control flow and low level optimizations - Inter - procedural analysis and optimization – call graph — register allocation – global References: – Optimization for memory hierarchy. Code Scheduling – Instruction scheduling – peculative scheduling – Software pipelining – trace scheduling –percolation scheduling.

Module 5: Parallelism and Case study

Instruction - level parallelism - Instruction pipelines and branch delays – pipelined execution –data dependence – dependencies among memory accesses. CaseStudies – Sun Compilers for SPARC–IBMXL Compilers – Alpha compilers – PA – RISC assembly language – COOL – (Classroom Object oriented language) – Compiler testing tools – SPIM

Course Plan

No	Topic	No. of Lectures (**40 Hours**)
1	**Module 1: Overview of Compiler Design**	**8**
1.1	Introduction - The phases of Compiler. Lexical Analysis – Role of Lexical Analyzer	1
1.2	Specification and Recognition of Tokens – Context Free Grammar	1
1.3	Symbol – Table Structure - Symbol Attributes and Symbol – Table Entries	1
1.4	Local Symbol – Table Management, Global Symbol – Table Structure,	1
1.5	Storage Binding and Symbolic Registers, Approaches to Generating Loads and Stores.	1
1.6	Intermediate representation – Issues	1
1.7	High level, medium level, low level intermediate languages	1
1.8	MIR, HIR, LIR – ICAN for Intermediate code.	1
2	**Module 2: Intermediate Representations**	**9**
2.1	The value – number method for constructing DAGs - Addresses and Instructions – Quadruples - Triples.	1

2.2	Storage organization – Static versus dynamic storage allocation - stack Allocation of space. Activation trees – Activation records.	1
2.3	Garbage collection – Design goals for Garbage collectors - Reference Counting garbage collectors. Introduction to trace – based collection - Abasic mark – and – sweep collector.	1
2.4	Translation of expressions Translation of expressions –Operations Within expressions–Incremental translation–Addressing array elements.	1
2.5	Translation of array references.	1
2.6	Control flow –Boolean expressions – Short – circuit code - flow- of - Control statements – Control flow translation of Boolean expressions.	1
2.7	Avoiding redundant Go to – Boolean values and jumping code.	1
2.8	Back patching - One-pass code generation using back patching –Backpatching for Boolean expressions.	1
2.9	Flow-of-control statements – Break, continue and Goto statements, Translation of switch statements – syntax directed translation of Switch statements - intermediate code for procedures.	1
3	**Module 3: Code Optimization**	**8**
3.1	Principal sources of optimization – causes of redundancy - Semantics Preserving transformations	1
3.2	Global common subexpressions – Copy Propagation – Dead code elimination.	1
3.3	Code motion – Upward code motion - Downward code motion.	1
3.4	Induction variables and reduction in strength.	1
3.5	Introduction to data flow analysis.	1
3.6	Loops in Flow graphs - Dominators	1
3.7	Introduction to global data flow analysis – Points and Paths - Reaching Definitions.	1
3.8	Live variable analysis-Data flow analysis of structured program	1
4	**Module 4: Register Allocation and Code Scheduling**	**7**
4.1	Register allocation and assignment – graph coloring .	1
4.2	Control flow and low level optimizations.	1
4.3	Inter-procedural analysis and optimization .	1
4.4	Call graph—register allocation.	1
4.5	Global References:– Optimization for memory hierarchy.	1
4.6	Code Scheduling –Instructions scheduling – Speculative scheduling	1

4.7	Software pipelining . Trace scheduling – percolation scheduling	1
5	**Module 5: Parallelism and Case study**	**8**
5.1	Optimization of basic blocks.	1
5.2	The DAG representation of basic blocks.	1
5.3	Peep hole optimization.	1
5.4	Instruction- level parallelism-Instruction pipelines and branch Delays – pipelined execution.	1
5.5	Data dependence – dependencies among memory accesses.	1
5.6	Case Studies – Sun Compilers for SPARC.	1
5.7	IBMXL Compilers –Alpha compilers.	1
5.8	PA–RISC assembly language, COOL– (Classroom Object oriented language), Compiler testing tools –SPIM.	1

References

1. Steven S Muchnik, "Advanced Compiler Design and Implementation", Morgan Kaufmann publishers, Elsevier Science, India, Indian Reprint 2003.
2. Compilers: Principles, Techniques and Tools (2nd edition), Alfred V. Aho, Monica S. Lam, Ravi Sethi, and Jeffery D. Ullman., Addison Wesley, Boston, MA, 2006.
3. Compilers: Principles, Techniques and Tools, Aho, A. V, Sethi, R. and Ullman, J. D. Pearson Education, 1986.
4. D. M. Dhamdhere, "Compiler Construction" (2/e),Macmillan.
5. Cooper &Torczon, "Engineering a Compiler"Elsevier.
6. K C. Louden, "Compiler Construction: Principles and Practice"Cengage.

241ECS011	BIOINFORMATICS	CATEGORY	L	T	P	CREDIT
		PEC	3	0	0	3

Preamble: This course provides an exposure to the fundamental concepts and techniques in molecular biology and Bioinformatics. Accessing, retrieval and analysis of data from different types of biological databases are examined in this course. Sequence Alignment and Phylogenetic analysis helps to identify the relationship between species. Protein structure prediction and application of machine learning to bioinformatics is also discussed. This course helps the learners to identify computational problems in molecular biology and apply efficient algorithms to solve them.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Make use of fundamental concepts of molecular biology to provide computational solutions **(Cognitive Knowledge Level: Understand)**
CO 2	Utilize bioinformatics tools and databases for retrieving, analysing and understanding biological data **(Cognitive Knowledge Level: Apply)**
CO 3	Analyze multiple sequences and find conserved regions.**(Cognitive Knowledge Level: Analyze)**
CO 4	Find the relationships between species by constructing phylogenetic tree. **(Cognitive Knowledge Level: Apply)**
CO 5	Predict unknown protein structures and apply concepts of Machine learning and their applications in Bioinformatics.**(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, Implement and Present innovative ideas on Bioinformatics techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Total duration of the examination will be 150 minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain the bio-sequences associated with central dogma of molecular biology.
2. Define translation. List two essential roles of ribosome during translation.

Course Outcome 2 (CO2)

1. Discuss the applications of Bioinformatics.
2. Examine the importance of primary biological databases.

Course Outcome 3(CO3):

1. Using Needleman and Wunsch dynamic programming method, construct the partial alignment score table for the following two sequences, using the following scoring parameters: match score: +5, mismatch score: -1, gap penalty: -2.

   GCATGCU
   GATTACA

   Write down the optimal global alignment between these sequences along with optimal score.

Course Outcome 4 (CO4):

1. Use UPGMA to reconstruct a phylogenetic tree using the following distance matrix.

Species	A	B	C	D
B	9	-	-	-
C	8	11	-	-
D	12	15	10	-
E	15	18	13	5

Course Outcome 5 (CO5):

1. Examine different algorithms for protein folding.
2. Elaborate Chou-Fasman and GOR methods for predicting secondary structure.

Course Outcome 6 (CO6):

1. Perform Phylogenetic Analysis on 2 gene Sequences.
2. Examine Microarrays and their applications in Bioinformatics.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES:2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIRST SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 241ECS011

Course Name: Bioinformatics

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Explain the concept of base pairing. If the amount of thymine in genome is 30%, calculate the percentage of cytosine? **(5)**

2. Compare and contrast DDBJ, Genbank. **(5)**

3. Differentiate PAM and BLOSUM series. **(5)**

4. Differentiate between rooted and unrooted phylogenetic tees. How many rooted and unrooted trees are possible for n species? **(5)**

5. Explain the significance of protein folding. **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) If the sequence of one strand of DNA is written as follows: **(2)**

5'-ATGCATGCATGCATGCATGCATGCATGC-3'

Construct the sequence of complementary strand.

(b) With the help of a diagram explain the concept of central dogma of molecular Biology. **(5)**

7. Describe the importance of biological databases in bioinformatics. Explain Protein sequence databases. **(7)**

8. Define scoring matrices? Explain how PAM is derived? **(7)**

9. Using Smith Waterman method, construct the partial alignment scoring table and obtain the optimal local alignment of the following two sequences: **(7)**

_ ACGTATCGCGTATA

GATGCTCTCGGAJAA

10. Illustrate the concept of a phylogenetic tree? Explain the steps of UPGMA method for phylogenetic tree construction with an example. **(7)**

11. Explain the significance of Hidden Markov Model in bioinformatics. Discuss the advantages and disadvantages of using HMMs **(7)**

12. Explain any two methods for Protein Secondary Structure Prediction. **(7)**

Syllabus

Module 1: Molecular Biology

Biomolecules-DNA, RNA and proteins-Components and structure. Genome organization. Letter codes for amino acids. Central dogma of molecular Biology. Genetic code. Interatomic forces in proteins, different levels of protein structure, protein domains, motifs.

Module 2: Bioinformatics

Definition and brief history. Bioinformatics vs Computational Biology. Scope and research areas of Bioinformatics. Data archives: Biological Databases-classification and importance; Nucleic acid databases: GenBank, DDBJ, EMB. Protein Sequence Databases: SwissProt, PIR. Derived databases: InterPro, Prosite, Pfam. Structure Databases: RCSB PDB, CATH, SCOP. Bibliographic Databases: PubMed, MEDLINE. Specialized databases. Gateways to archives: Entrez, SRS, ExPASy.

Module 3: Sequence Alignment

Concept of sequence alignment, Gaps in alignment, Scoring matrices: PAM and BLOSUM, Alignment of pairs of sequences: Dot Plot, Dynamic Programming, Alignment algorithms: The Needleman and Wunsch algorithm, Smith-Waterman algorithm. Search for homologous sequences using BLAST and FASTA programs. Statistical significance of database searches. Multiple sequence Alignment: Concept and Algorithms in MSA. Tools: Clustal, Mega.

Module 4: Molecular Phylogenetics

Concept and its relation to Multiple Sequence Alignment. Representation of phylogeny. Concept of Outgroup. Gene and Species phylogeny. Phylogenetic tree construction methods: Distance based and Character based. Phylogenetic software's: PHYLIP, MrBayes.

Module 5: Protein Structure Prediction and Advanced Bioinformatics

Protein Stability and folding, Ramachandran plot, Homology modelling, Energy minimization, CASP; Protein structure prediction software's: ESyPred3D, Rosetta; PSI-BLAST; Introduction to machine learning techniques: Hidden Markov models, Genetic algorithms and artificial neural networks. Applications of machine learning techniques in multiple sequence alignment, structure prediction and phylogenetic analysis

Course Plan

No	Topic	No. of Lectures (**40 Hours**)
1	**Module 1: Molecular Biology**	**5**
1.1	Biomolecules-DNA, RNA and proteins	1
1.2	Components and structure. Genome organization. Letter codes for amino acids.	1
1.3	Central dogma of molecular Biology	1
1.4	Genetic code. Interatomic forces in proteins	1
1.5	Different levels of protein structure, protein domains, motifs.	1
2	**Module 2: Bioinformatics**	**7**
2.1	Definition and brief history. Bioinformatics vs Computational Biology, Scope and research areas of Bioinformatics	1
2.2	Data archives: Biological Databases-classification and importance, Nucleic acid databases: GenBank, DDBJ, EMB	1
2.3	Protein Sequence Databases: SwissProt, PIR	1
2.4	Derived databases: InterPro, Prosite, Pfam	1
2.5	Structure Databases: RCSB PDB, CATH, SCOP	1
2.6	Bibliographic Databases: PubMed, MEDLINE	1
2.7	Specialized databases. Gateways to archives: Entrez, SRS, ExPASy.	1
3	**Module 3: Sequence Alignment**	**10**
3.1	Concept of sequence alignment, Gaps in alignment	1
3.2	Scoring matrices: PAM and BLOSUM	1
3.3	Alignment of pairs of sequences: Dot Plot, Dynamic Programming	1
3.4	Alignment algorithms: The Needleman and Wunsch algorithm	1
3.5	Smith-Waterman algorithm	1
3.6	Search for homologous sequences using BLAST and FASTA programs	1
3.7	Statistical significance of database searches	1

3.8	Multiple sequence Alignment: Concept	1
3.9	Algorithms in MSA	1
3.10	Tools: Clustal, Mega.	1
4	**Module 4: Molecular Phylogenetics**	**8**
4.1	Concept and its relation to Multiple Sequence Alignment	1
4.2	Representation of phylogeny	1
4.3	Concept of Outgroup	1
4.4	Gene and Species phylogeny	1
4.5	Phylogenetic tree construction methods: Distance based (Lecture 1)	1
4.6	Phylogenetic tree construction methods: Distance based (Lecture 2)	1
4.7	Phylogenetic tree construction methods: Character based	1
4.8	Phylogenetic software's: PHYLIP, MrBayes	1
5	**Module 5: Protein Structure Prediction and Advanced Bioinformatics**	**10**
5.1	Protein Stability and folding	1
5.2	Ramachandran plot, Homology modelling	1
5.3	Energy minimization	1
5.4	CASP; Protein structure prediction software's: ESyPred3D, Rosetta	1
5.5	PSI-BLAST	1
5.6	Introduction to machine learning techniques: Hidden Markov models	1
5.7	Genetic algorithms	1
5.8	Artificial neural networks	1
5.9	Applications of machine learning techniques in multiple sequence alignment	1
5.10	Structure prediction and phylogenetic analysis	1

References

1. Jin Xiong, Essential Bioinformatics, Cambridge University Press, 2006.
2. Brown T. A. Genomes 3, 3rd edition, BIOS Scientific Publishers Limited, 2007.
3. Jeremy M. Berg, John L. Tymoczko, Lubert Stryer, Biochemistry, 5th Edition, W. H. Freeman, 2002.
4. Arthur Lesk, Introduction to Bioinformatics, 5th Edition, Oxford University Press,2019.
5. Dan E. Krane, Fundamental concepts of Bioinformatics, 3rd Edition, Pearson Education India, 2002.
6. David W. Mount, Bioinformatics Sequence and Genome Analysis, 2nd Edition, Cold Spring Press.
7. Andreas D. Baxevanis, B. F. Francis Ouellette, Bioinformatics: A practical Guide to the Analysis of Genes and Proteins, John Wiley & Sons, 2004.
8. Dov Stekel, Microarray Bioinformatics, Cambridge University Press, 2003.

241LCS100	COMPUTING LAB 1	CATEGORY	L	T	P	Credit
		PCC	0	0	3	2

Preamble: Study of the course enables the learners to make use of the machine learning concepts and algorithms to derive data insights. The course provides exposure to the design and implementation aspects of machine learning algorithms such as decision trees, regression, naive bayes algorithm, clustering algorithms and artificial neural network. This helps the students to develop machine learning based solutions to real world problems.

Course Outcomes: After the completion of the course the student will be able to

CO#	Course Outcomes
CO1	Apply modern machine learning notions in predictive data analysis(**Cognitive Knowledge Level: Apply**)
CO2	Analyze the range of machine learning algorithms along with their strengths and weaknesses (**Cognitive Knowledge Level: Analyze**)
CO3	Design and develop appropriate machine learning models to solve real world problems. (**Cognitive Knowledge Level: Analyze**)
CO4	Build predictive models from data and analyze their performance(**Cognitive Knowledge Level: Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓	✓	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	✓	✓	

Continuous Internal Evaluation Pattern:

The laboratory courses will be having only Continuous Internal Evaluation and carries 100 marks. Final assessment shall be done by two examiners; one examiner will be a senior faculty from the same department.

Continuous Evaluation : 60 marks

Final internal assessment : 40 marks

Lab Report:

All the students attending the Lab should have a Fair Report. The report should contains details of experiment such as Objective, Algorithm/Design, Description, Implementation, Analysis, Results, and Outcome. The report should contain a print out of the respective code with inputs addressing all the aspects of the algorithm described and corresponding outputs. All the experiments noted in the fair report should be verified by the faculty regularly. The fair report, properly certified by the faculty, should be produced during the time of the final assessment.

Syllabus

Decision tree (ID3), Naïve bayesian classifier , Bayesian network, Expectation Maximization (EM) algorithm,K-means algorithm, K-nearest neighbor, Regression, Cross validation, Support Vector Machine (SVM), Artificial neural network, Backpropagation algorithm, Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), Google colab.

Practice Questions

1. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
2. Write a program to implement the naïve bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
3. Assuming a set of documents that need to be classified, use the naïve bayesian Classifier model to perform this task. Calculate the accuracy, precision, and recall for your data set.
4. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Python ML library classes/API.
5. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Python ML library classes/API in the program.
6. Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions. Python ML library classes can be used for this problem.
7. Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.
8. Write a program to implement 5-fold cross validation on a given dataset. Compare the accuracy, precision, recall, and F-score for your data set for different folds.
9. Implement SVM/Softmax classifier for CIFAR-10 dataset: (i) using KNN, (ii) using 3 layer neural network.
10. Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.
11. Image Captioning with Vanilla RNNs .
12. Image Captioning with LSTMs.
13. Familiarisation of cloud based computing like Google colab.

References:

1. Jiawei Han, Micheline Kamber, Jian Pei. Data Mining Concepts and Techniques, Third Edition. Morgan Kaufmann.
2. Christopher M. Bishop. Pattern recognition and machine learning. Springer 2006.
3. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition, MIT Press 2010.
4. Mohammed J. Zaki and Wagner Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, First South Asia edition, 2016.
5. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016.
6. Neural Networks and Deep Learning, Aggarwal, Charu C., c Springer International Publishing AG, part of Springer Nature 2018.

SEMESTER-II SYLLABUS

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242TCS100	ADVANCED DATA STRUCTURES AND ALGORITHMS	DISCIPLINE CORE 2	3	0	0	3

Preamble: The course introduces advanced data structures and algorithms in different domains. The goal of this course is to provide a solid background in the design and analysis of the major classes of algorithms. The course helps the learners to develop their own versions for a given computational task and to compare and contrast their performance.

Course Outcomes: After the completion of the course the student will be able to:*

CO 1	Analyse the relevance of amortized analysis and applications. (Cognitive Level: Apply)
CO 2	Illustrate string matching algorithms. (Cognitive Level: Apply)
CO 3	Illustrate advanced data structures like Binomial heap, Fibonacci heap, Disjoint set and string matching algorithms. (Cognitive Level: Apply)
CO 4	Illustrate network flow algorithms and applications. (Cognitive Level: Apply)
CO 5	Make use of probabilistic algorithms and approximation algorithms in computing. (Cognitive Level: Apply)
CO 6	Design, develop and implement software using advanced data structures and algorithms. (Cognitive Level: Create)

* The COs shown are only indicative. For each course, there can be 4 to 6 COs.

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓		✓	
CO 2	✓		✓	✓		✓	
CO 3	✓		✓	✓		✓	
CO 4	✓		✓	✓		✓	
CO 5	✓		✓	✓		✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	80%
Analyse	20%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks Course
based task/Seminar/Quiz : 10 marks
Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted. Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain how the accounting method of amortized analysis can be applied to stack operations.

2. Suppose we perform a sequence of n operations on a data structure in which the i^{th} operation costs i if i is an exact power of 2, and 1 otherwise. Use aggregate analysis to determine the amortized cost per operation.

3. What is the total cost of executing n of the stack operations PUSH, POP, and MULTIPOP, assuming that the stack begins with s_0 objects and finishes with s_n objects? Use potential method.

Course Outcome 2 (CO2)

1. Use an aggregate analysis to show that the running time of KMP-MATCHER is θ(n).

2. Working modulo q = 11, how many spurious hits does the Rabin-Karp matcher encounter in the text T = 3141592653589793 when looking for the pattern P = 26?

3. Compute the prefix function π for the pattern ababbabbabbababbabb

Course Outcome 3(CO3):

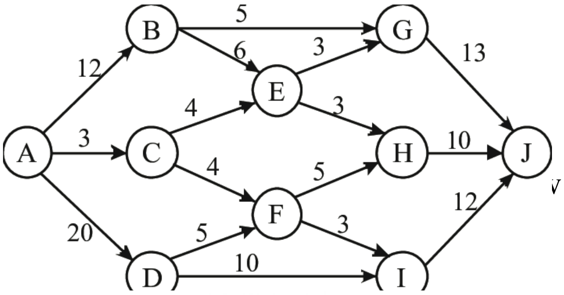
1. Analyse the time complexity of decrease-key operation of Fibonacci heap.

2. Illustrate extract-min operation of Binomial heap.

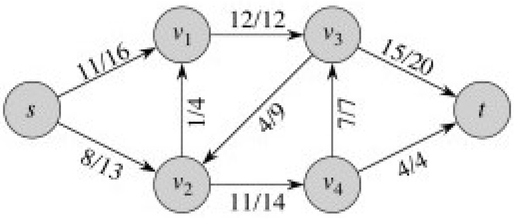
3. Explain the heuristics used in disjoint set data structure to improve the running time.

Course Outcome 4 (CO4):

1. Show the execution of the Edmonds-Karp algorithm on the given flow network (source: A and sink: J).



v across the cut ({s, v_2, v_4}, {v_1, v_3, t}). What is the



3. State and prove max flow min cut theorem.

Course Outcome 5 (CO5):

1. Illustrate Miller-Rabin primality testing method.

2. Explain probabilistic selection algorithm.

Course Outcome 6 (CO6):

1. Explain the approximation algorithm for subset sum problem.

2. Consider each of the following words as a set of letters: {*arid, dash, drain, heard, lost, nose, shun, slate, snare, thread*}. Show which set cover GREEDY-SET-COVER produces when we break ties in favor of the word that appears first in the dictionary.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242TCS100

Course Name: ADVANCED DATA STRUCTURES AND ALGORITHMS

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. Explain accounting method of amortized analysis with a suitable example.

2. Explain the algorithm for uniting two binomial heaps and analyse the running time.

3. Maximum matching in a bipartite graph G corresponds to a maximum flow in its corresponding flow network G'. Comment on this statement. Explain how maximum flow problem can be used to solve maximum bipartite matching problem.

4. Explain thbe probabilistic algorithm for verifying matrix multiplication problem.

5. Explain the approximation algorithm for traveling salesperson problem.

(5x5=25)

Part B

(Answer any five questions. Each question carries 7 marks)

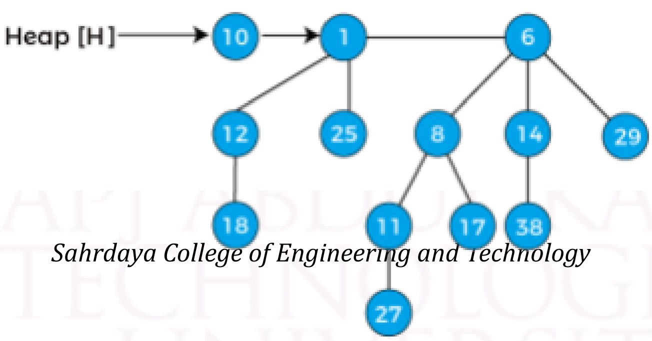
6. Describe Knuth-Morris-Pratt algorithm and illustrate using given text T = AABAACAADAABAABA and pattern P = AABA. **(7)**

7. (a) Using potential method, compute the amortized cost of incrementing a binary counter. **(3)**

(b) Suppose we perform a sequence of n operations on a data structure in which the i^{th} operation costs i if i is an exact power of 2, and 1 otherwise. Use accounting method of amortized analysis to determine the amortized cost per operation. **(4)**

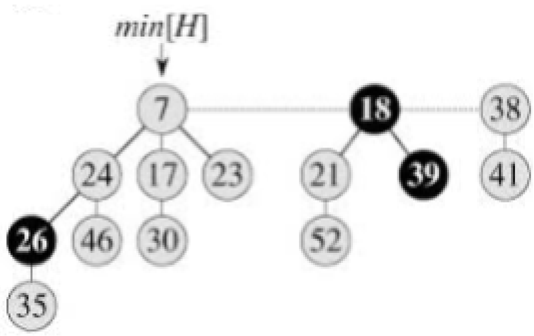
8. (a) Explain how disjoint set data structure is used to find connected components on an undirected graph. **(3)**

(b) Show the binomial heap that results when a node with key 11 is deleted from the binomial heap shown in figure. **(4)**

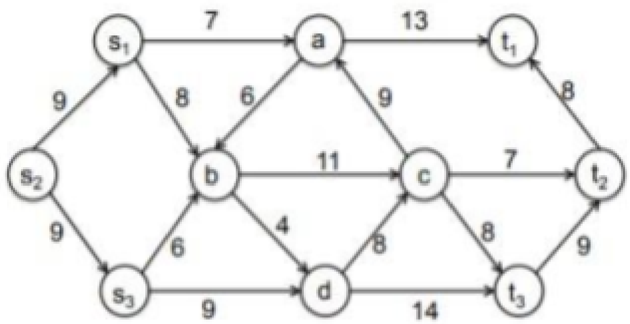


9. (a) Explain the structure of Fibonacci heap. **(2)**

(b) Apply extract minimum operation on the Fibonacci heap shown in figure and show the result. **(5)**



10 Describe Ford-Fulkerson algorithm and apply on the following network. Also obtain minimum cut across the network. **(7)**



11 (a) Apply Miller-Rabin algorithm to test whether the number 341 is prime or not. **(4)**

(b) Explain probabilistic quick sort algorithm. **(3)**

12	(a)	Describe polynomial-time approximation scheme and fully polynomial-time approximation scheme.	**(3)**
	(b)	Give an example of a graph for which APPROX-VERTEX-COVER always yields a suboptimal solution.	**(4)**

Syllabus

Module – 1 (Amortized analysis and String matching)

Overview of asymptotic notations and complexity analysis, Amortized analysis – aggregate analysis, accounting method, potential method.

String matching – introduction, Rabin-Karp algorithm, Knuth-Morris-Pratt algorithm.

Module – 2 (Advanced data structures)

Overview of binary heap operations, Binomial tree and heap, Binomial heap operations, Fibonacci heap structure, Fibonacci heap operations, Disjoint set – overview, linked list representation, disjoint set forests.

Module – 3 (Network flow)

Network flow properties, examples, residual network, augmenting path, cut of network, maxflow-mincut theorem, Ford-Fulkerson algorithm, Edmonds-Karp algorithm, maximum bipartite matching.

Module – 4 (Probabilistic algorithms)

Introduction, types of probabilistic algorithms, Numerical algorithms – Numerical integration, Probabilistic counting, Monte-Carlo algorithms – Verifying matrix multiplication.

Number theory fundamentals – modular arithmetic, modular exponentiation, Euler's Theorem and Fermat's Theorem, Primality testing – Miller-Rabin test.

Las Vegas algorithms – Probabilistic selection and quick sort.

Module – 5 (Approximation algorithms)

Introduction, Vertex-cover problem, Traveling-salesman problem, Set-covering problem, Subset-sum problem.

Course Plan

No	Topic	No. of Lectures (37)
1	**Module – 1 (Amortized analysis and String matching)**	
1.1	Overview of asymptotic notations and complexity analysis	1
1.2	Amortized analysis – aggregate analysis	1
1.3	accounting method	1
1.4	potential method	1
1.5	String matching – introduction	1
1.6	Rabin-Karp algorithm	1
1.7	Knuth-Morris-Pratt algorithm (1)	1
1.8	Knuth-Morris-Pratt algorithm (2)	1
2	**Module – 2 (Advanced data structures)**	
2.1	Overview of binary heap operations	1
2.2	Binomial tree and heap	1
2.3	Binomial heap operations (1)	1
2.4	Binomial heap operations (2)	1
2.5	Fibonacci heap structure	1
2.6	Fibonacci heap operations (1)	1
2.7	Fibonacci heap operations (2)	1
2.8	Disjoint set – overview, linked list representation	1
2.9	disjoint set forests	1
3	**Module – 3 (Network flow)**	
3.1	Network flow properties, examples	1
3.2	residual network, augmenting path, cut of network	1
3.3	maxflow-mincut theorem	1
3.4	Ford-Fulkerson algorithm	1
3.5	Edmonds-Karp algorithm	1
3.6	Maximum bipartite matching	1
4	**Module – 4 (Probabilistic algorithms)**	
4.1	Introduction, types of probabilistic algorithms	1
4.2	Numerical algorithms – Numerical integration, Probabilistic counting	1
4.3	Monte-Carlo algorithms – Verifying matrix multiplication	1
4.4	Number theory fundamentals – modular arithmetic, modular exponentiation	1
4.5	Euler's Theorem and Fermat's Theorem	1
4.6	Primality testing – Miller-Rabin test (1)	1

4.7	Primality testing – Miller-Rabin test (2)	1
4.8	Las Vegas algorithms – Probabilistic selection and quick sort	1
5	**Module – 5 (Approximation algorithms)**	
5.1	Introduction	1
5.2	Vertex-cover problem	1
5.3	Traveling-salesman problem	1
5.4	Set-covering problem	1
5.5	Subset-sum problem (1)	1
5.6	Subset-sum problem (2)	1

Reference Books

1. T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, "Introduction to Algorithms", MIT Press, 3rd edition, 2009.

2. Gilles Brassard and Paul Bratley, "Fundamentals of algorithms", Prentice-hall of India Private Limited, 2001.

3. Rajeev Motwani, Prabhakar Raghavan, "Randomized Algorithms", Cambridge University Press, 2000.

4. Dexter C. Kozen, "The Design and Analysis of Algorithms", Springer.

5. Jon Kleinberg and Eva Tardos, "Algorithm Design", Pearson Education, 2006.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242TCS001	ADVANCED OPERATING SYSTEMS	PROGRAM CORE 3	3	0	0	3

Preamble: Study of this course enables the learners to understand the configuration and functions of OS Kernel and have an overview on concepts implemented in modern operating systems. The course focus on providing information on the design and implementation of the Linux kernel modules. This course will help the learners to suggest solutions/ modify the existing architectural features.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Illustrate the concepts of process management and process scheduling mechanisms employed in the Linux operating system. **(Cognitive Knowledge Level: Apply)**
CO 2	Describe the set of interfaces by which the process running in user space can interact with the system and how the Kernel manages the various interrupts. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply various synchronization methods to write race free code. **(Cognitive Knowledge Level: Apply)**
CO 4	Demonstrate how the kernel handles memory and implementation of the file system. **(Cognitive Knowledge Level: Apply)**
CO 5	Analyze how kernel manages block devices and their requests and identify the issues to be considered in writing portable codes **(Cognitive Knowledge Level: Analyze)**
CO 6	Design and implement different kernel modules. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks Course
based task/Seminar/Quiz : 10 marks
Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted. Test
paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. How does Linux process scheduler ensure a fair share of its computational power among the executing process.
2. By default the child process starts execution in Linux. Discuss the reason behind this.
3. How is TASK_INTERRUPTIBLE state different from TASK_UNINTERRUPTIBLE state ? Suppose a process is waiting for a specific event to happen. Which state should it be assigned?

Course Outcome 2 (CO2):

1. Illustrate the role of system calls in Linux.
2. Can we use work queues in an interrupt context? Justify your answer.

3. In interrupt processing, if the deferred work needs to run in interrupt context and should guarantee that no two of same type can run concurrently, which bottom half can be preferred. Write notes on its implementation and usage.

Course Outcome 3(CO3):

1. What trouble arises if a kernel code which has acquired a spin lock is interrupted by an interrupt handler trying to acquire the same spin lock? Is it possible to avoid such a situation ? How ?
2. If your code needs to sleep, which is often the case when synchronizing with user-space: Explain which synchronization method is most suitable for the above scenario?
3. List out the various aspects of scalability and contention.

Course Outcome 4 (CO4):

1. Differentiate vmalloc() and kmalloc()? Which of these must be used for allocating memory for hardware devices? Why?
2. Explain the design of the slab layer. Also explain the advantage of slab layer allocation of kernel objects.
3. Illustrate the importance of zones in Linux memory management.

Course Outcome 5 (CO5):

1. Before the 2.6 kernel, buffer head was used for block I/O whereas in the later kernels, it has been replaced by biostructure as the basic container for block I/O. Why?
2. Discuss the significance of an I/O scheduler. What drawback of the deadline I/O scheduler is overcome by the anticipatory I/O scheduler?
3. Describe the measures to avoid data alignment issues.

Course Outcome 6 (CO6):

1. Design and implement a process scheduler.

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES : 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242TCS001

Course Name: ADVANCED OPERATING SYSTEMS

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Discuss about Linux implementation of threads.	**(5)**
2.	Describe how the work queue mechanism differs from other bottom half mechanisms.	**(5)**
3.	Explain pre-emption disabling and list its advantages.	**(5)**
4.	Specify the context in which kfree() is needed.	**(5)**
5.	Elaborate the working of the deadline I/O scheduler and describe Linus Elevator.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.	(a)	How does Linux process scheduler ensure a fair share of its computational power among the executing process.	**(4)**
	(b)	Consider a scenario with two runnable tasks: a text editor and a video encoder. Analyse the scheduling policy in action.	**(3)**
7.	(a)	Discuss how system calls are implemented and parameters are verified.	**(3)**
	(b)	Describe the methods used in enabling and disabling interrupts? Justify the significance of disabling an interrupt line.	**(4)**
8.		Explain the trouble that arises if a kernel code which has acquired a spin lock is interrupted by an interrupt handler trying to acquire the same spin lock? Is it possible to avoid such a situation ? How ?	**(7)**
9.	(a)	Illustrate the importance of zones in Linux memory management.	**(3)**

	(b)	Explain the design of slab layer and the advantage of slab layer allocation of kernel objects.	**(3)**
10	(a)	Explain data alignment. Describe the measures to avoid data alignment issues.	**(4)**
	(b)	List the difference between big Endian and little Endian byte ordering	**(3)**
11	(a)	How does the TASK_INTERRUPTIBLE state differ from TASK_UNINTERRUPTIBLE state ? If a process is waiting for a specific event to happen, which state should be assigned?	**(4)**
	(b)	Explain the situation where individual members of a task list are deallocated?	**(3)**
12	(a)	Explain the significance of the bottom half mechanism. In interrupt processing, if the deferred work needs to run in interrupt context and should guarantee that no two of same type run concurrently, which bottom half is preferred. Justify.	**(7)**

Syllabus

Module 1: Process Management and Scheduling

Introduction to the Linux kernel, Process Management – Process, Process descriptor and the task structure, Process creation, The Linux implementation of threads, Process termination. Process Scheduling – Multitasking, Linux's process scheduler, Policy, Linux scheduling algorithm, Preemption and context switching, Real-time scheduling policies.

Module 2: System calls and Interrupts

System Calls - Communicating with the Kernel, Syscalls, System call handler, System call implementation, System call context . Interrupts and Interrupt Handlers – Interrupts, Interrupt handlers, Top halves versus bottom halves, Registering an interrupt handler, Writing an interrupt handler, Interrupt context, Interrupt control. Bottom Halves – Task queues, Softirqs, Tasklets, Work queues.

Module 3: Kernel Synchronization

Kernel Synchronization – Critical regions and race conditions, Locking, Deadlocks, Contention and scalability. Kernel Synchronization Methods – Atomic operations, Spin locks, Semaphores, Mutexes, Completion variables, BKL: The Big Kernel Lock, Sequential locks, Preemption disabling.

Module 4: Memory Management and Virtual File System

Memory Management – Pages, Zones, kmalloc(), kfree(), vmalloc(), Slab layer – design, Per- CPU allocations. The Virtual File system – VFS objects, data structures, relationship and functionalities.

Module 5: Block I/O Layer and Portability

The Block I/O Layer – Buffers and buffer heads, Request queues, I/O schedulers – Types, Scheduler selection. Portability – Word size and data types, Data alignment, Byte order, Time, Processor ordering

Course Plan

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Process Management and Scheduling**	**8**
1.1	Introduction to the Linux kernel , Process	1
1.2	Process descriptor and the task structure	1
1.3	Process creation, The Linux implementation of threads	1
1.4	Process termination, Multitasking	1
4.5	Linux's process scheduler, Policy	1
1.6	Linux scheduling algorithm	1
1.7	Preemption and context switching	1
1.8	Realtsime scheduling policies	1
2	**Module 2: System Calls and Interrupts**	9
2.1	Communicating with the Kernel, Syscalls	1
2.2	System Call Handler, System Call Implementation	1
2.3	System call context	1
2.4	Interrupts, Interrupt handlers, Top halves versus bottom halves	1
2.5	Registering an Interrupt Handler	1
2.6	Writing an Interrupt Handler	1
2.7	Interrupt Context, Interrupt Control	1
2.8	Bottom Halves – Task Queues, Softirqs	1
2.9	Tasklets, Work Queues	1
3	**Module 3: Kernel Synchronization**	9
3.1	Critical regions and race conditions	1
3.2	Locking, Deadlocks	1
3.3	Contention and scalability	1
3.4	Kernel Synchronization Methods – Atomic operations	1
3.5	Spin locks	1
3.6	Semaphores	1
3.7	Mutexes, Completion variables	1

3.8	BKL: The Big Kernel Lock	1
3.9	Sequential locks, Preemption disabling	1
4	**Module 4: Memory Management and Virtual File System**	7
4.1	Pages, Zones	1
4.2	kmalloc()	1
4.3	kfree(), vmalloc()	1
4.4	Slab Layer - Design	1
4.5	Per-CPU Allocations	1
4.6	The Virtual File system – VFS objects	1
4.7	Data structures, relationship and functionalities	1
5	**Module 5 : The Block I/O Layer**	7
5.1	Buffers and buffer heads	1
5.2	Request queues	1
5.3	I/O Schedulers	1
5.4	Types, Scheduler Selection	1
5.5	Portability – Word size and data types	1
5.6	Data Alignment, Byte Order	1
5.7	Time, Processor ordering	1

Reference

1. Robert Love, “Linux Kernel Development”, 3/e, Addison-Wesley, 2010.
2. Daniel Bovet, Marco Cesati, “Understanding the Linux Kernel”, 3/e, OReilly Media Inc., 2005.
3. Linux Kernel Architecture – Wolfgang Mauerer.
4. Reilly Christian Benvenuti, “Understanding Linux Network Internals”, 1/e, OReilly Media Inc.,2005.
5. Jonathan Corbet, Alessandro Rubini, Greg Kroah-Hartman, “Linux Device Drivers”, 3/e, OReilly Media Inc., 2005
6. Operating Systems Concepts, 9th Edition- Silberschatz, Galvin, Gagne

COURSE CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242PCS100	MINI PROJECT	PROJECT	0	0	4	2

Mini project can help to strengthen the understanding of student's fundamentals through application of theoretical concepts and to boost their skills and widen the horizon of their thinking. The ultimate aim of an engineering student is to resolve a problem by applying theoretical knowledge. Doing more projects increases problem solving skills.
The introduction of mini projects ensures preparedness of students to undertake dissertation. Students should identify a topic of interest in consultation with PG Programme Coordinator that should lead to their dissertation/research project. Demonstrate the novelty of the project through the results and outputs.The progress of the mini project is evaluated based on three reviews, two interim reviews and a final review. A report is required at the end of the semester.
Evaluation Committee - Programme Coordinator, One Senior Professor and Guide.

Sl. No	Type of evaluations	Mark	Evaluation criteria
1	Interim evaluation 1	20	
2	Interim evaluation 2	20	
3	Final evaluation by a Committee	35	Will be evaluating the level of completion and demonstration of functionality/ specifications, clarity of presentation, oral examination, work knowledge and involvement
4	Report	15	the committee will be evaluating for the technical content, adequacy of references, templates followed and permitted plagiarism level(not more than 25%)
5	Supervisor/Guide	10	
	Total Marks	100	

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242LCS100	COMPUTING LAB 2	LABORATORY 2	0	0	2	1

Preamble: This course provides an in-hand experience related to various database management systems. Also equips them to design and implement a database application built over the concepts. This course helps the learners to develop applications that manage data efficiently with the help of suitable data models and techniques.

Course Outcomes: After the completion of the course the student will be able to

CO#	Course Outcomes
CO1	Able to perform the distributed processing of large data sets across clusters of computers using simple programming models with the help of Hadoop. **(Cognitive Knowledge Level: Apply)**
CO2	Perform data summarization and ad hoc querying using Hive. **(Cognitive Knowledge Level: Create)**
CO3	Operates on document databases and techniques using DynamoDB. **(Cognitive Knowledge Level: Apply)**
CO4	Capable of implementing XML and XML queries for data management. **(Cognitive Knowledge Level: Create)**
CO5	Apply emerging technologies in column store along with Cassandra. **(Cognitive Knowledge Level: Apply)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓	✓	✓	✓	✓	✓	✓
CO2	✓	✓	✓	✓	✓		
CO3	✓	✓	✓	✓	✓		
CO4	✓	✓	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓		

Continuous Internal Evaluation Pattern:

The laboratory courses will be having only Continuous Internal Evaluation and carries 100 marks. Final assessment shall be done by two examiners; one examiner will be a senior faculty from the same department.

Continuous Evaluation: 60 marks
Final internal assessment: 40 marks

Lab Report:

All the students attending the Lab should have a Fair Report. The report should contain details of experiments such as Objective, Algorithm/Design, Description, Implementation, Analysis, Results, and Outcome. The report should contain a print out of the respective code with inputs addressing all the aspects of the algorithm described and corresponding outputs. All the experiments noted in the fair report should be verified by the faculty regularly. The fair report, properly certified by the faculty, should be produced during the time of the final assessment.

Syllabus

Basic concepts of Big Data, Configuration of Hadoop, Parallel Database, Distributed Database, Semi-structured Data and XML Databases: XML Data Model – XML Schema- DTD-XSD – XPath and XQuery, JDOQL (Java Data Object-based Query Language), No SQL Databases: Key value stores – DynamoDB, Column Based- HBase, Cassandra, Web Page ranking algorithm.

Practice Questions

1. Study and Configure Hadoop for Big Data.
2. Study of NoSQL Databases such as Hive/HBase/Cassandra/DynamoDB.
3. Design Data Model using NoSQL Databases such as Hive/HBase/Cassandra/DynamoDB.
4. Implement any one Partitioning technique in Parallel Databases.
5. Implement Two Phase commit protocol in Distributed Databases.
6. Design Persistent Objects using JDO and implement min 10 queries on objects using JDOQL in Object DB NOSQL DATABASE.
7. Create XML, XML schemas, DTD for any database application and implement min 10 queries using XQuery FLOWR expression and XPath.
8. Design database schemas and implement min 10 queries using Hive/ HBase/ Cassandra column-based databases.
9. Design database schemas and implement min 10 queries using Dynamo DB key-value based databases.
10. Implement Web Page ranking algorithm.
11. Create a database infrastructure like GitHub backend DB for a project team to collaborate for coding, code review, code commenting and approval, code rejection for changes workflow.

Reference Books:

1. Tom White, Hadoop: The Definitive Guide, O'Reilly Media 4th Edition, April 2015.
2. Joe Fawcett, Danny Ayers, Liam R. E. Quin, Beginning XML, 5/e, John Wiley & Sons, 2012.
3. Jeff Carpenter, Eben Hewitt, “Cassandra: The Definitive Guide”, 3/e, O’Reilly, 2020.
4. Tanmay Deshpande, “DynamoDB Cookbook”, Packt Publishing, September 2015.

SEMESTER II
PROGRAM ELECTIVE III

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS100	BIG DATA ANALYTICS	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: This course helps the learner to understand the basic concepts of data analytics. This course covers mathematics for data analytics, Big data and its applications, techniques for managing big data and data analysis & visualization using R programming tool. It enables the learners to perform data analysis on a real-world scenario using appropriate tools.

Prerequisite: Nil

Course Outcomes: After the completion of the course the student will be able to

CO 1	Illustrate the concept for data analytics using basic mathematics **(Cognitive Knowledge Level: Apply)**
CO 2	Illustrate the concepts of data analytics **(Cognitive Knowledge Level: Apply)**
CO 3	Apply various Big Data Analytics Techniques using R **(Cognitive Knowledge Level: Apply)**
CO 4	Access and Process Data on Distributed File System and to Manage Job Execution in Hadoop Environment **(Cognitive Knowledge Level: Apply)**
CO 5	Analyze the Big Data using Advanced analytical methods such as text analysis **(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, Implement and Present innovative ideas on distributed algorithms and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Data Science has a big role to play than BI in big data analytics. Give your view.
2. The number of members of a millionaires' club were as follows:

Year	2011	2012	2013	2014	2015	2016

Members	23	24	27	25	30	28

i. What is the average growth rate of the membership?

ii. Based on the results of (i), how many members would one expect in 2018?

3. A On-line store has integrated Big Data across multiple IT systems to gather customer transaction and interactions data in order to predict customer recommendations better. Provide your view on how BI and Data science handles the recommendations.

Course Outcome 2 (CO2):

1. Assume you have more than two populations to consider? Which hypothesis testing method would you adopt. State the initial and alternate hypothesis for the same, How to accept/reject these hypotheses?

2. Differentiate between type-1 error and type-2 error. Do you think one is always more serious than the other. Justify.

3. How various t-tests can be done in R? Give the R functions for the same and their usage.

Course Outcome 3(CO3):

1. With a neat sketch illustrate the architecture diagram of Hadoop File System.

2. Illustrate the working of a Map Reduce program with example.

3. Explain the Avro and File-Based Data structures.

Course Outcome 4 (CO4):

1. List and explain various Hadoop ecosystem components.

2. Explain the data analytic architecture with a diagram.

3. Give a code in Hive which illustrate the functions to read and write data in to a database.

Course Outcome 5 (CO5):

1. Given a movie review dataset. Propose a technique by which sentiments can be analysed from rated movie reviews.

2. Which numerical statistic is intended to reflect how important a 'word' is to a document in a collection or corpus? Justify your answer and sort out the same using the following example.

   Given a document containing 100 words wherein the word ‘cat’ appears 3 times. Assume there are ‘10 million’ documents and the word ‘cat’ appears in ‘one thousand’ of these.

Course Outcome 6 (CO6):

1. Assume that a non-stationary time series with seasonal fluctuations of last 100 months is given. Applying various methods illustrate how prediction can be done from this series for next 12 months.

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS100

Course Name: BIG DATA ANALYTICS

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Data Science has a big role to play than BI in big data analytics. Give your view	**(5)**
2.	Assume you have more than two populations to consider? Which hypothesis testing method would you adopt. State the initial and alternate hypothesis for the same, How to accept/reject these hypotheses?	**(5)**
3.	With a neat sketch illustrate the architecture diagram of Hadoop File System	**(5)**
4.	List and explain various Hadoop ecosystem components.	**(5)**
5.	Given a movie review dataset. Propose a technique by which sentiments can be analysed from rated movie reviews.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) The number of members of a millionaires' club were as follows: (4)

Year	2011	2012	2013	2014	2015	2016
Member s	23	24	27	25	30	28

iii. What is the average growth rate of the membership?

iv. Based on the results of (i), how many members would one expect in 2018?

(b) A On-line store has integrated Big Data across multiple IT systems to gather customer transaction and interactions data in order to predict customer recommendations better. Provide your view on how BI and Data science handles the recommendations. (3)

7. Differentiate between type-1 error and type-2 error. Do you think one is always more serious than the other. Justify. (7)

8. How various t-tests can be done in R? Give the R functions for the same and their usage.(7)

9. Illustrate the working of a Map Reduce program with example **(7)**

10 Explain the Avro and File-Based Data structures **(7)**

11 Explain the data analytic architecture with a diagram **(7)**

12 Give a code in Hive which illustrate the functions to read and write data in to a database

(7)

Syllabus

Module 1: Introduction to Data Analytics

Mathematics for Data Analytics - Descriptive statistics - Measures of central tendency and dispersion, Association of two variables, Probability calculus - probability distributions, Inductive statistics - Point estimation, Interval estimation, Hypothesis Testing - Basic definitions, t-test. Introduction To Big Data: Big data characteristics, Features of Big Data, Evolution of Big data, Analyst Perspective on Data Repositories , State of the Practice in Analytics, BI Versus Data Science.

Module 2: Introduction to R

Review of basic data analytic methods using R : Introduction to R, R graphical user interface- data import and export-attribute and data type. Exploratory data analysis-Visualization, Dirty data, single and multiple variables, data exploration vs presentation. Statistical methods for evaluation-Hypothesis testing, difference of means, Wilcoxon rank sum test, type I and II errors, power and sample size,

Module 3: Hadoop & HDFS

History of Hadoop, Apache Hadoop, Analysing Data with Unix tools, Analysing Data with Hadoop, Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Infosphere BigInsights and Big Sheets.

The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures

Module 4: Map reduce & Hadoop Ecosystem

Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features.

Pig : Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. Hive : Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL.

Module 5: Advanced Analytical Theory, Methods and Text Analysis

Advanced analytical theory and methods: Time Series Analysis- Overview of Time Series Analysis, Box-Jenkins Methodology ARIMA Model, Autocorrelation Function (ACF), Autoregressive Models, Moving Average Models ARMA and ARIMA Models Building and Evaluating an ARIMA Model, Reasons to Choose and Cautions.

Course Plan

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Introduction to Data Analytics**	**8**
1.1	Mathematics for Data Analytics - Descriptive statistics	1
1.2	Measures of central tendency and dispersion, Association of two variables	1
1.3	Probability calculus - probability distributions, Inductive statistics	1
1.4	Point estimation, Interval estimation, Hypothesis Testing - Basic definitions, t-test	1
1.5	Introduction To Big Data: Big data characteristics	1
1.6	Features of Big Data,	1
1.7	Evolution of Big data, Analyst Perspective on Data Repositories	1
1.8	State of the Practice in Analytics BI Versus Data Science	1

2	**Module 2: Introduction to R**	8
2.1	Review of basic data analytic methods using R: Introduction to R	1
2.2	R graphical user interface-data import and export-attribute and data type	1
2.3	Exploratory data analysis-Visualization	1
2.4	Dirty data, single and multiple variables, data exploration vs presentation.	1
2.5	Statistical methods for evaluation	1
2.6	Hypothesis testing, Difference of means.	1
2.7	Wilcoxon rank sum test, type I and II errors	1
2.8	Power and sample size.	1
3	**Module 3: Hadoop & HDFS**	**8**
3.1	History of Hadoop, Apache Hadoop	1
3.2	Analysing Data with Unix tools, Analysing Data with Hadoop	1
3.3	Hadoop Streaming	1
3.4	Hadoop Echo System, IBM Big Data Strategy	1
3.5	Introduction to Infosphere Big Insights and Big Sheets.	1
3.6	The Design of HDFS, HDFS Concepts, Command Line Interface	1
3.7	Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives	1
3.8	Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures	1
4	**Module 4: Map reduce & Hadoop Ecosystem**	**8**
4.1	Anatomy of a Map Reduce Job Run	1
4.2	Failures, Job Scheduling.	1
4.3	Shuffle and Sort, Task Execution	1
4.4	Map Reduce Types and Formats, Map Reduce Features.	1
4.5	Pig: Introduction to PIG, Execution Modes of Pig	1
4.6	Comparison of Pig with Databases, Grunt, Pig Latin	1
4.7	User Defined Functions, Data Processing operators.	1
4.8	Hive: Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL.	1
5	**Module 5: Advanced Analytical Theory, Methods and Text Analysis**	**8**
5.1	Advanced analytical theory and methods: Time Series Analysis	1
5.2	Overview of Time Series Analysis.	1
5.2	Box-Jenkins Methodology ARIMA Model	1
5.3	Autocorrelation Function (ACF)	1
5.4	Autoregressive Models	1
5.5	Moving Average Models	1
5.6	ARMA and ARIMA Models Building	1

5.7	Evaluating an ARIMA Model	1
5.8	Reasons to Choose and Cautions	1

References

1. Christian Heumann and Michael Schomaker, “Introduction to Statistics and DataAnalysis”, Springer, 2016
2. David Dietrich, Barry Heller, Biebie Yang, Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, EMC Education Services, John Wiley &Sons, Inc
3. Jaiwei Han, MichelineKamber, “Data Mining Concepts and Techniques", Elsevier, 2006.
4. Tom White “ Hadoop: The Definitive Guide” Third Edit on, O’reily Media, 2012.
5. Seema Acharya, Subhasini Chellappan, "Big Data Analytics" Wiley 2015

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS001	WIRELESS SENSOR NETWORKS	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: This course gives an introduction to wireless sensor networks, which has helped to develop a vast array of applications in diverse areas like agriculture, military, health and civil infrastructures. The course deals with hardware aspects like sensors with their built in transceivers and associated electronics as well as the softwares needed to run them. The MAC layer with its myriad protocols, as well as routing layer protocols, are discussed in the course. Localization of sensor nodes, deployment and coverage, as well as security of wireless sensor networks also form part of the course. This course enables the learners to design and develop wireless sensor protocols and applications.

Course Outcomes: After the completion of the course the student will be able to

CO 1	List the applications, hardware and software components of wireless sensor networks, and the challenges faced in design of sensor networks. (**Cognitive knowledge: Analyze**)
CO 2	Apply design principles and formulate necessary service interfaces while designing a wireless sensor network (**Cognitive knowledge: Apply**)
CO 3	Design MAC protocols for wireless sensor networks taking into account the specific requirements of the network. (**Cognitive knowledge: Create**)
CO 4	Explain localization techniques, coverage problem and security issues in wireless sensor networks (**Cognitive knowledge: Evaluate**)
CO 5	Design energy efficient routing protocols for wireless sensor networks (**Cognitive knowledge**: **Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

the-art tools to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✔		✔	✔		✔	
CO 2	✔		✔	✔	✔	✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Test paper (1 number) : 10 marks
iii. Course based task/Seminar /Data collection and interpretation : 15 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Enumerate the challenges faced while designing a wireless sensor network.
2. List the important hardware elements of a wireless sensor and explain the function of each.
3. What are the challenges faced in sensor network programming?

Course Outcome 2 (CO2):

1. Examine the design principles that should be followed while designing a wireless sensor network.
2. Evaluate the performance of IEEE 802.11 DCF using Markov chain.
3. Illustrate the functionalities that a service interface should provide when interfacing an application to a protocol stack.

1. Analyze any three techniques which MAC protocols use to reduce energy consumption.
2. Differentiate between contention free and schedule based MAC protocols.
3. Design a schedule based MAC protocol.

Course Outcome 4 (CO4):

1. Propose a method to find the best coverage path between two nodes in a wireless sensor network.
2. List the types of security attacks that can occur in WSNs.

Course Outcome 5 (CO5):

1. What is data-centric routing? Why is data-centric routing more feasible compared to routing based on addresses?
2. Under what circumstances GPSR enters the perimeter routing mode ?
3. Design a routing protocol for a wireless sensor network that supports an application with continuous data transfer.

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES : 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 24ECS001

Course Name: Wireless Sensor Networks

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Describe important applications of wireless sensor networks.	**(5)**
2.	Differentiate between single hop and multihop wireless sensor networks.	**(5)**
3.	How does T-MAC take care of the shortcomings of S-MAC? Comment on T- MAC's ability to adapt to traffic density.	**(5)**
4.	Evaluate security challenges in wireless sensor networks?	**(5)**
5	Explain any three ways in which gossiping overcomes the challenges of flooding.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

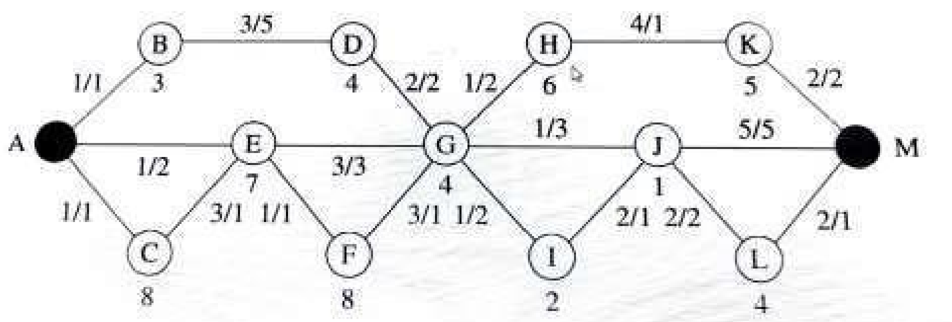
6.	(a)	Explain how to calculate the amount of energy saved if a sensor node is allowed to sleep.	**(5)**
	(b)	A cross layered approach is more suitable for wireless sensor networks than a layered approach. Explain	**(2)**
7.		Demonstrate the different types of in-network processing.	**(7)**
8.	(a)	Show how the design of the MAC protocol affects the energy efficiency of a sensor node.	**(3)**
	(b)	What are the advantages and disadvantages of TRAMA protocol? What is the	

purpose of NP component? **(4)**

9. (a) Why is overhearing a problem in a wireless sensor network.How does PAMAS solve this problem ? C(S3 1)

(b) Define "early sleeping problem" and discuss the solution provided by T-MAC. **(4)**

10. **(7)**



For each link in the above figure, the ratio gives latency/energy cost for transmitting a single packet, while the number under each node gives the node's remaining energy capacity.

Node G's routing table is given as

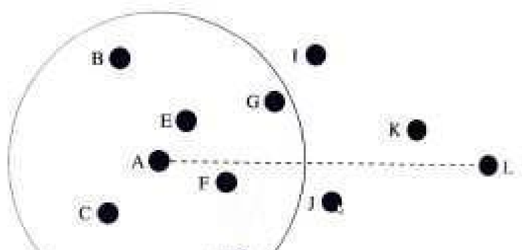
Event	Distance	Direction
E1	3	F
E2	4	I

G sends queries toward events E1, E2 using rumor routing.Show how it is made possible.

11. (a) Enumerate the issues and challenges while designing a routing protocol for wireless sensor network(3)

(b)

In the above figure, if node A wants to forward a packet to node L using greedy



forwarding, which neighbor should it choose? The communication range of node A is given by the circle.(4)

12. (a) Localization in wireless sensor networks. (4)

(b) Discuss a method to find the maximum breach path in a wireless sensor network (3)

Syllabus

Module 1: Introduction to Wireless Sensor Networks

Motivation, Challenges and Constraints, Applications.

Node Architecture – Hardware elements, Sensors and Actuators, Power supply, Energy Consumption of sensor nodes.

Challenges in sensor network programming, Operating systems and execution environments- embedded OS, issues, Programming models.

Module 2: Sensor Network Architecture

Sensor network scenarios, Optimization goals and figures of merit

Design principles for WSNs, Service interfaces of WSNs. Gateway concepts, WSN- Internet Communication.

Discrete time Markov Chain, Performance analysis of IEEE 802.11 DCF using Markov Chain

Module 3: MAC Protocols for WSNs

Characteristics and design goals, low duty cycle protocols and wakeup concepts. Contention based protocols: PAMAS, STEM, T-MAC,

Schedule based protocols: SMACS, TRAMA, Hybrid MAC protocols: Z-MAC

Case Studies: S-MAC, 802.15.4, 802.15.6

Module 4: Routing Protocols for WSNs

Introduction, Routing challenges and design issues, Routing Strategies

Routing Techniques: Flooding and Gossiping, SPIN, LEACH, PEGASIS, Directed Diffusion, Rumour Routing, Geographic Routing- Forwarding Strategies.

Module 5: Localization, Coverage and Security in WSNs

Localization: approaches-proximity-trilateration and triangulation- scene analysis

Coverage and deployment: sensing models, coverage measures, uniform random deployments, coverage determination

Security: Security challenges in WSNs, Security attacks in WSNs

No	Topic	No. of Lectures (40)
1	**Module 1: Introduction to Wireless Sensor Networks**	**9**
1.1	Motivation, Challenges and Constraints	1
1.2	Applications	1
1.3	Node Architecture	1
1.4	Hardware elements	1

1.5	Sensors and Actuators, Power supply	1
1.6	Energy Consumption of sensor nodes	1
1.7	Challenges in sensor network programming	1
1.8	Operating systems and execution environments- embedded OS, issues	1
1.9	Programming models	1
2	**Module 2: Sensor Network Architecture**	**7**
2.1	Sensor network scenarios	1
2.2	Optimization goals and figures of merit	1
2.3	Design principles for WSNs	1
2.4	Service interfaces of WSNs	1
2.5	Gateway concepts, WSN- Internet Communication	1
2.6	Discrete time Markov Chain	1
2.7	Performance analysis of IEEE 802.11 DCF using Markov Chain	1
3	**Module 3: MAC Protocols for WSNs**	**8**
3.1	Characteristics and design goals	1
3.2	Low duty cycle protocols and wakeup concepts	1
3.3	Contention based protocols: PAMAS	1
3.4	STEM, T-MAC	1
3.5	Schedule based protocols: SMACS TRAMA	1
3.6	Hybrid MAC protocols: Z-MAC	1
3.7	Case Studies: S-MAC	1
3.8	802.15.4, 802.15.6	1
4	**Module 4: Routing Protocols for WSNs**	**8**
4.1	Introduction, Routing challenges and design issues	1
4.2	Routing Strategies	1
4.3	Routing Techniques: Flooding and Gossiping, SPIN	1
4.4	LEACH, PEGASIS	1
4.5	Directed Diffusion	1
4.6	Rumour Routing	1
4.7	Geographic Routing	1
4.8	Forwarding Strategies	1
5	**Module 5: Localization, Coverage and Security in WSNs**	**8**
5.1	Localization: approaches-proximity-trilateration	1
5.2	Triangulation, scene analysis	1
5.3	Coverage and deployment: sensing models	1
5.4	Coverage measures	1

5.5	Uniform random deployments	1
5.6	Coverage determination	1
5.7	Security: Security challenges in wireless sensor networks	1
5.8	Security attacks in wireless sensor networks	1

References

1. W. Dargie and C. Poellabauer, Fundamentals of Wireless Sensor Networks, Theory and Practice,Wiley, 2010
2. Holger Karl and Andreas Willig, Protocols and Architectures for Wireless Sensor Networks,Wiley, 2005
3. K. Sohraby, D. Minoli and T. Znati, Wireless Sensor Networks, Technology, Protocols, and Applications,Wiley-Interscience, 2007
4. C. Siva Ram Murthy and B. S Manoj, Adhoc Wireless Networks Architectures and Protocols,Prentice Hall, 2004
5. Feng Zhao and Leonidas Guibas, Wireless Sensor Networks An Information Processing Approach, Morgan Kaufman, 2005

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS002	DEEP LEARNING	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: Study of this course provides the learners an overview of the concepts and algorithms involved in deep learning. The course covers the basic concepts in deep learning, optimization techniques, regularization techniques, convolutional neural networks, recurrent neural networks, graphical models, deep generative models. This course helps the students to implement deep learning algorithms to solve real-world problems.

Prerequisite : Nil

Course Outcomes: After the completion of the course the student will be able to

CO 1	Use the standard regularization and optimization techniques for the effective training of deep neural networks. **(Cognitive Knowledge Level: Apply)**
CO 2	Build convolutional Neural Network (CNN) models for different use cases. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply the concepts of Recurrent Neural Network (RNN), Long Short Term Memory(LSTM), Gated Recurrent Unit (GRU) for solving problems. **(Cognitive Knowledge Level:Apply)**
CO 4	Construct Bayesian networks, Markov networks and apply computational techniques to draw inferences. **(Cognitive Knowledge Level: Apply)**
CO 5	Illustrate the concepts of auto encoder, sampling algorithms, deep generative models and transfer learning. **(Cognitive Knowledge Level: Apply)**
CO 6	Design, develop, implement and present innovative ideas on deep learning concepts and techniques to solve real-world problems. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards.

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

:COMPUTER SCIENCE AND ENGINEERING-CS1

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Write an algorithm for backpropagation which uses stochastic gradient descent method. Comment on the effect of adding momentum to the network.
2. You are doing full batch gradient descent using the entire training set (not stochastic gradient descent). Is it necessary to shuffle the training data? Explain your answer.
3. You would like to train a dog/cat image classifier using mini-batch gradient descent. You have already split your dataset into train, validation and test sets. The classes are balanced. You realize that within the training set, the images are ordered in such a way that all the dog images come first and all the cat images come after. A friend tells you: ”you absolutely need to shuffle your training set before the training procedure.” Is your

friend right? Explain.

4. Explain how L1 regularization method leads to weight sparsity.
5. Data augmentation is often to increase the amount of data you have, Should you apply data augmentation to the test set? Explain why.

Course Outcome 2 (CO2):

1. You are given a dataset of 10 x 10 grayscale images. Your goal is to build a 5-class classifier. Explain which one of the following two options you would choose and why ?
a) the input is flattened into a 100-dimensional vector, followed by a fully-connected layer with 5 neurons, b) the input is directly given to a convolutional layer with five 10 x 10 filters.

2. Weight sharing allows CNNs to deal with image data without using too many parameters. Does weight sharing increase the bias or the variance of a model? Explain

3. A convolutional neural network has 4 consecutive layers as follows:
3 x 3 conv (stride 2) - 2 x 2 Pool - 3 x 3 conv (stride 2) - 2 x 2 Pool. Determine how large is the set of image pixels which activate a neuron in the 4th non- image layer of this network?

4. Consider the convolutional neural network defined by the layers in the left column below. Determine the shape of the output volume and the number of parameters at each layer. You can write the activation shapes in the format (H, W, C), where H, W, C are the height, width and channel dimensions, respectively. Unless specified, assume padding 1, stride 1 where appropriate.

Notation:

- CONVx-N denotes a convolutional layer with N filters with height and width equal to x.
- POOL-n denotes a n×n max-pooling layer with stride of n and 0 padding.
- FLATTEN flattens its inputs.
- FC-N denotes a fully-connected layer with N neurons.

Layer	Activation Volume Dimensions	Number of parameters
Input	$32 \times 32 \times 3$	0
CONV3-8		
Leaky ReLU		
POOL-2		
BATCHNORM		
CONV3-16		

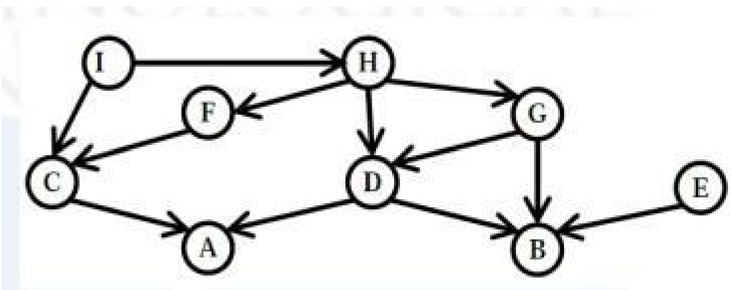
Leaky ReLU		
POOL-2		
FLATTEN		
FC-10		

Course Outcome 3(CO3):

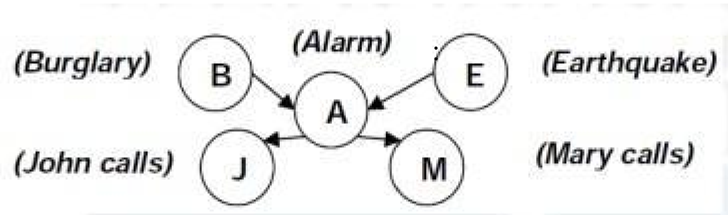
1. Illustrate the workings of the RNN with an example of a single sequence defined on a vocabulary of four words.
2. List the differences between LSTM and GRU
3. Explain your understanding of unfolding a recursive or recurrent computation into a computational graph.
4. The vanishing gradient problem is more pronounced in RNN than in traditional neural networks. Give reason. Discuss a solution for the problem.
5. Show the steps involved in an LSTM to predict stock prices. Give one advantage of using an RNN rather than a convolutional network.

Course Outcome 4 (CO4):

1. Construct the Bayesian Network that corresponds to this conditional probability: P(A | B,C,E) P(B | D,E) P(C | F,H) P(D | G) P(E| G,H) P(F | H) P(G) P(H)
2. Write down the factored conditional probability expression that corresponds to the graphical Bayesian Network shown below.



3. Shown below is the Bayesian network corresponding to the Burglar Alarm problem, P(J | A) P(M | A) P(A | B, E) P(B) P(E). The probability tables show the probability that variable is True, e.g., P(M) means P(M = t). Calculate P(J = t ∧ M = f ∧ A = f ∧ B = f ∧ E = t).



P(E)
.002

A	P(M)
t	.70
f	.01

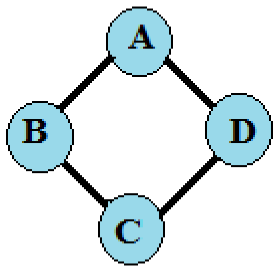
B	E	P(A)
t	t	.95
t	f	.94
f	t	.29
f	f	.001

P(B)
.001

A	P(J)
t	.90
f	.05

4. Consider the simple Markov network given below. Let A,B,C and D be binary random variables representing four people's beliefs as to whether the earth is round (1 for believes, 0 for does not believe). Determine the probability of only person A and D believes that the earth is round.

A	B	ϕ_{AB}
0	0	50
0	1	5
1	0	5
1	1	50



A	D	ϕ_{AD}
0	0	5
0	1	50
1	0	50
1	1	5

B	C	ϕ_{BC}
0	0	1
0	1	5
1	0	45
1	1	50

C	D	ϕ_{CD}
0	0	1
0	1	15
1	0	40
1	1	50

Course Outcome 5 (CO5):

1. Define effective sample size (ESS). Large ESS necessary but not sufficient for good MCMC mixing. Justify
2. Is an autoencoder for supervised learning or for unsupervised learning? Explain briefly.
3. Explain how does the variational auto-encoder(VAE) architecture allow it to generate new data points, compared to auto-encoder, which cannot generate new data points?
4. Generative Adversarial Networks(GANs) include a generator and a discriminator. Sketch a basic GAN using those elements, a source of real images, and a source of randomness.

5. Write down the formula for the energy function (E) of a Restricted Boltzmann Machine (RBM).
6. List the difference between Boltzmann Machine and Deep Belief Network.
7. One of your friends has trained a cat vs. non-cat classifier. It performs very well and you want to use transfer learning to build your own model. Explain what additional hyperparameters (due to the transfer learning) you will need to tune.

Course Outcome 6 (CO6):

1. Implement image classification using CNN
2. Implement ECG classification using LSTM

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES: 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS002

Course Name: Deep Learning

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. A 2 × 2 image is represented by the following pixel value matrix. (5)

$$\begin{bmatrix} 5 & 4 \\ 2 & 7 \end{bmatrix}$$

This image is given to a 3-layer neural network, that is, two hidden layers and one output layer. Draw schematic diagram of the network.

Assuming all inter-connection weights having values 1, bias having value 0, the hidden layers having 3 neurons each, and a simple activation function of the form **(5)**

2. In Convolutional Neural Networks, there is no need to perform feature extraction. Justify with an example.
3. Explain your understanding of unfolding a recursive or recurrent computation into a computational graph. List three differences between LSTM and GRU.

4. Sketch the core idea of the Monte Carlo method. What is a sample? What is a direct sampling method? Why can't it be used directly to do any inference? What is rejection sampling? What is its major disadvantage?

5. How does the variational auto-encoder(VAE) architecture allow it to generate new data points, compared to auto-encoder which cannot generate new data points? **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. Differentiate gradient descent with and without momentum. Give equations for weight updation in GD with and without momentum. Illustrate plateaus, saddle points and slowly varying gradients. **(7)**

7. (a) Consider an activation volume of size 13×13×64 and a filter of size 3×3×64. Discuss whether it is possible to perform convolutions with strides 2, 3 and 5. What happens if the stride of the convolutional layer increases? What can be the maximum stride? Justify your answer. **(4)**

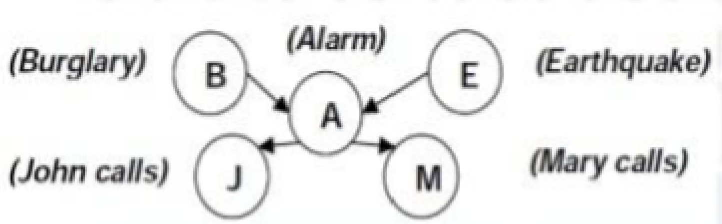
(b) How backpropagation is used to learn higher-order features in a convolutional Network? **(3)**

8. The vanishing gradient problem is more pronounced in RNN than in traditional neural networks. Give reason. Discuss a solution for the problem. **(7)**

9. Shown below is the Bayesian network corresponding to the Burglar Alarm problem, P(J | A) P(M | A) P(A | B, E) P(B) P(E). The probability tables show the probability that variable is True, e.g., P(M) means P(M = t). Calculate **(7)**

i) P(J ∧ M ∧ A ∧ ¬B ∧ ¬E)

ii) P(J)



P(E)
.002

A	P(M)
t	.70
f	.01

B	E	P(A)
t	t	.95
t	f	.94
f	t	.29
f	f	.001

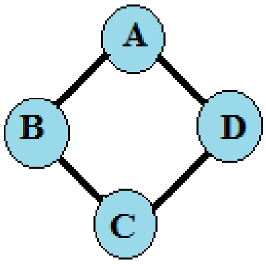
A	P(J)
t	.90
f	.05

P(B)
.001

10 Compare Boltzmann Machine with Deep Belief Network. **(7)**

11	(a)	Training error of the deep learning model trained for the classification problem was found to be very low but generalization error was high. Identify the problem and suggest techniques to reduce this generalization error.	**(4)**
	(b)	Initializing the weights of a neural network with very small or large random numbers is not advisable. Justify.	**(3)**
12		Consider the simple Markov network given below. Let A,B,C and D be binary random variables representing four people's beliefs as to whether the earth is round (1 for believes, 0 for does not believe). Determine the probability of only person D believes that the earth is round.	**(7)**

A	B	ϕ_{AB}
0	0	50
0	1	5
1	0	5
1	1	50



A	D	ϕ_{AD}
0	0	5
0	1	50
1	0	50
1	1	5

B	C	ϕ_{BC}
0	0	1
0	1	5
1	0	45
1	1	50

C	D	ϕ_{CD}
0	0	1
0	1	15
1	0	40
1	1	50

Syllabus

Module 1: Introduction to Deep learning

Introduction to deep learning, Deep feed forward network, Training deep models - introduction, setup and initialization issues, Vanishing and exploding gradient problems, Optimization techniques - Gradient Descent (GD), Stochastic GD, GD with momentum, GD with Nesterov momentum, AdaGrad, RMSProp, Adam. Regularization Techniques - L1 and L2 regularization, Early stopping, Dataset augmentation, Parameter tying and sharing, Ensemble methods, Dropout.

Module 2: Convolutional Neural Networks

Convolutional Neural Networks –Architecture, Convolution and Pooling operation, Motivation, Variants of convolution functions, Structured outputs, Data types, Efficient convolution algorithms, Training a Convolutional Network, Applications of Convolutional Networks, Case study of Convolutional Architectures – AlexNet

Module 3: Recurrent neural networks

Recurrent neural networks – Computational graphs, RNN design, encoder – decoder sequence to sequence architectures, deep recurrent networks, recursive neural networks, challenges of training Recurrent Networks, gated RNNs LSTM and GRU, Applications of RNNs.

Module 4: Graphical Models and Sampling

Graphical models - Bayesian network, Markov networks, Inference on chains and factor graphs. Monte Carlo Methods – Basics of Monte Carlo Sampling, Importance sampling, Markov chain Monte Carlo methods(MCMC), Gibbs sampling.

Module 5: Advanced Deep learning Topics

Autoencoders, Variational AutoEncoder , Deep generative models - Boltzmann machines, Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines,

,Generative Adversarial Networks, Auto-Regressive Networks. Transfer Learning and Domain Adaptation.

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Introduction to deep learning**	**9**
1.1	Introduction to deep learning, Deep feed forward network	1
1.2	Training deep models - Introduction, setup and initialization issues	1
1.3	Vanishing and exploding gradient problems	1
1.4	Concepts of optimization, Gradient Descent (GD)	1
1.5	Stochastic GD, GD with momentum, GD with Nesterov momentum	1
1.6	AdaGrad, RMSProp, Adam	1
1.7	Concepts of Regularization, L1 and L2 regularization	1
1.8	Early stopping, Dataset augmentation	1
1.9	Parameter tying and sharing, Ensemble methods, Dropout	1
2	**Module 2 : Convolutional Neural Network**	**9**
2.1	Convolutional Neural Networks, Architecture	1
2.2	Convolution and Pooling operation with example	1

2.3	Motivation	1
2.4	Variants of convolution functions	1
2.5	Structured outputs, Data types	1
2.6	Efficient convolution algorithms	1
2.7	Training a Convolutional Network	1
2.8	Applications of Convolutional Networks	1
2.9	Case study of Convolutional Architectures – AlexNet	1
3	**Module 3 : Recurrent Neural Network**	**7**
3.1	Recurrent neural networks – Computational graphs	1
3.2	RNN design, Encoder – decoder sequence to sequence architectures	1
3.3	Deep recurrent networks, Recursive neural networks	1
3.4	Challenges of training Recurrent Networks	1
3.5	LSTM	1
3.6	GRU	1
3.7	Applications of RNN	1
4	**Module 4 : Graphical Models and Sampling**	**6**
4.1	Graphical models - Bayesian network	1
4.2	Markov network	1
4.3	Inference on chains and factor graphs	1
4.4	Monte Carlo Methods – Basics of Monte Carlo Sampling	1
4.5	Importance sampling	1
4.6	Markov chain Monte Carlo methods(MCMC), Gibbs sampling	1
5	**Module 5 : Advanced Deep learning Topics**	**9**
5.1	Autoencoders	1
5.2	Variational Autoencoder	1
5.3	Deep generative models - Boltzmann machines	1

5.4	Restricted Boltzmann Machines	1
5.5	Deep Belief Networks	1
5.6	Deep Boltzmann Machines	1
5.7	Generative Adversarial Networks	1
5.8	Auto-Regressive Networks	1
5.9	Transfer Learning and Domain Adaptation.	1

References

1. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016.
2. Neural Networks and Deep Learning, Aggarwal, Charu C., c Springer International Publishing AG, part of Springer Nature 2018.
3. Christopher M. Bishop. Pattern recognition and machine learning. Springer 2006.
4. David Foster. Generative Deep Learning - Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Inc., June 2019.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS003	COMPUTER VISION	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: Study of this course provides the leaners to explain the fundamental concepts of Computer Vision and Image Processing, and major approaches that address them. This course provides concepts in computer vision including image acquisition and image formation models. This course helps the learners to model fuzzy image processing, probabilistic model of image formation and different image compression techniques.

Prerequisite : Nil

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply domain knowledge in vision system, radiometry, shadows, shading, views and Stereopsis for image formation. **(Cognitive Knowledge Level: Apply)**
CO 2	Apply affine structures and Geometry for reconstructing images. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply Bayesian decision theory for pattern classification. **(Cognitive Knowledge Level: Apply)**
CO 4	Analyze the contents of the image using image processing, probabilistic modelling and fuzzy image processing. **(Cognitive Knowledge Level: Analyze)**
CO 5	Analyze various techniques used for image restoration, compression and segmentation. **(Cognitive Knowledge Level: Analyze)**
CO 6	Design, develop, implement and present innovative ideas on computer vision concepts and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyse	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. What is BRDF? How are area sources different from line sources?
2. What shapes can the shadow of a sphere take, if it is cast on a plane, and the source is a point source?
3. What are the different criteria for evaluating the performance of the computer vision algorithms?

Course Outcome 2 (CO2) :

1. How can a 3D structure captured using a sequence of video frames be recovered?
2. Discuss the methods that are used to segment the data points into independently-moving objects

Course Outcome 3(CO3):

1. Explain Minimum Squared Error Method (MSE) for Classification.

2. Explain how Bayesian Decision Theory can be used to determine whether a given email is "spam" or "non-spam".

Course Outcome 4 (CO4):

1. Explain about the basic relationships and distance measures between pixels in a digital image?
2. Explain Object recognition as probabilistic modelling.
3. How to define an image as a Fuzzy Set?

Course Outcome 5 (CO5):

1. Explain how an image is restored using an inverse filter. What are its drawbacks?
2. Given an image with the intensity distribution as

Intensity Value	Percentage(%)
0	20
32	10
128	25
144	5
152	5
160	10
168	5
249	5
250	10
251	5

Use the Huffman code to compress the image. Draw the Huffman tree.

Course Outcome 6 (CO6):

1. Implement any one applications- Object Recognition with Intelligent Cameras/ Fast 3-D Full Body Scanning for Humans and Other Objects/ Motion Tracking

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS003

Course Name: COMPUTER VISION

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. State different limitations of a pinhole camera and how to overcome these limitations. Write a short note on thin lenses. **(5)**

2. How is conversion from affine to euclidean images performed? **(5)**

3. Differentiate "Bayesian decision theory discrete feature" and "Bayesian decision theory continuous feature" with illustration. **(5)**

4. Illustrate the relevance of smoothing filters in image processing. **(5)**

5. Explain the method of global thresholding for image segmentation. **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) Describe the role of Stereopsis in computer vision? What are the two processes involved in it? **(7)**

(b) Explain why it is difficult to use shadow boundaries to infer shape, particularly if the shadow is cast onto a curved surface.

7. A video sequence of a 3D scene is provided. Suggest an algorithm for deriving affine shape from motion. **(7)**

8. What is meant by a pose? How can you hypothesize a correspondence between a collection of image features and a collection of object features, using pose consistency? **(7)**

Discuss about weak perspective projection and orthographic projection in all.

10. Consider an 8 level 64x64 image with normalized gray levels in the range [0, 1].The normalized histogram is given below. Perform histogram equalization and plot the equalized histogram. (7)

r_k	n_k	$P(r_k)$
$r_0=0$	790	0.19
$r_1=1$	1023	0.25
$r_2=2$	850	0.21
$r_3=3$	656	0.16
$r_4=4$	329	0.08
$r_5=5$	245	0.06
$r_6=6$	122	0.03
$r_7=7$	81	0.02

11. Given the messages x_1, x_2, x_3, x_4, x_5, and $x_6$with respective probabilities 0.30, 0.25, 0.20, 0.12, 0.08 and 0.05. Determine the binary code by applying Huffman encoding procedure and hence find efficiency? (7)

12. The arithmetic decoding process is the reverse of the encoding procedure. Decode the message 0.23355 given the coding model. (

Symbol	Probability
a	0.2
e	0.3
i	0.1
o	0.2
u	0.1
!	0.1

Syllabus

Module 1: Image formation and modelling

Components of a vision system, Imaging systems, Signal processing for computer vision, Pattern recognition for computer vision, Performance evaluation of algorithms.

Image formation and Image model- Camera model and camera calibration- Radiometry- Light in space- Light in surface - Sources, shadows and shading, Multiple images-The Geometry of multiple views- Stereopsis.

Module 2: Affine structures

Affine structure from motion- Elements of Affine Geometry, Affine structure and motion from two images- Affine structure and motion from multiple images- From Affine to Euclidean images.

High level vision- Geometric methods- Model based vision- Obtaining hypothesis by pose consistency, pose clustering and using Invariants, Verification.

Module 3: Bayesian Decision Theory

Bayesian Decision Theory- Minimum error rate classification Classifiers, discriminant functions, decision surfaces- The normal density and discriminant-functions for the Normal density.

Module 4: Introduction to Digital Image Processing

Introduction to Digital Image Processing- fundamental steps in Digital Image Processing, relationship between pixels, intensity transformations and spatial filtering: basic intensity transformation functions, histogram processing, spatial filtering, smoothing and sharpening filters

Probabilistic Modelling and Fuzzy Image Processing: Introduction of Probabilistic Modelling in Computer Vision, why probabilistic models, Object recognition as probabilistic modelling, Introduction, Fuzzy image understanding, Fuzzy image processing systems, Theoretical components of fuzzy image processing.

Module 5: Processing on Images

Image restoration: noise models, restoration in the presence of noise only, periodic noise reduction.

Image compression: fundamentals, compression models and standards, basic compression methods: Huffman coding, arithmetic coding, LZW coding, run-length coding.

Image segmentation: point, line and edge detection, thresholding region based segmentation

Case Study: Any two applications-Object Recognition with Intelligent Cameras/ Fast 3-D Full Body Scanning for Humans and Other Objects/ Motion Tracking/ Multicolour Classification of Astronomical Objects.

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Image formation and modelling**	**7**
1.1	Components of a vision system	1
1.2	Application for computer vision and performance evaluation.	1
1.3	Cameras- camera model and camera calibration	1
1.4	Radiometry- Light in space- Light in surface	1
1.5	Sources, shadows and shading	1
1.6	Multiple images-The Geometry of multiple views	1
1.7	Stereopsis	1
2	**Module 2: Affine structures**	**8**
2.1	Elements of Affine Geometry	1
2.2	Affine structure	1
2.3	Motion from two images	1
2.4	Affine structure and motion from multiple images	1
2.5	From Affine to Euclidean images.	1
2.6	High level vision- Geometric methods-	1
2.7	Model based vision- Obtaining hypothesis by pose consistency,	1
2.8	Pose clustering and using Invariants, Verification.	1
3	**Module 3: Bayesian Decision Theory**	**6**
3.1	Bayesian Decision Theory	1
3.2	Minimum error rate classification Classifiers	1
3.3	Discriminant functions	1
3.4	Decision surfaces.	1
3.5	The normal density and discriminant	1
3.6	Functions for the Normal density.	1
4	**Module 4: Introduction to Digital Image Processing**	**10**
4.1	Fundamental steps in Digital Image Processing	1
4.2	Relationship between pixels	1
4.3	Intensity transformations	1
4.4	Spatial filtering	1
4.5	basic intensity transformation functions	1
4.6	histogram processing	1
4.7	Spatial filtering	1
4.8	Smoothing and sharpening filters.	1
4.9	Probabilistic Modelling in Computer Vision	1

4.10	Fuzzy Image processing	1
5	**Module 5: Processing on Images**	**9**

5.1	Image restoration: noise models	1
5.2	Restoration in the presence of noise only.	1
5.3	Periodic noise reduction.	1
5.4	Image compression: fundamentals	1
5.5	Compression models and standards	1
5.6	Basic compression methods: Huffman coding, arithmetic coding,	1
5.7	LZW coding, run-length coding,	1
5.8	Image segmentation: point, line and edge detection,	1
5.9	Thresholding , region based segmentation	1

References

1. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

2. R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, John Wiley, 2001.

3. Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, Second Edition, Cambridge University Press, 2004.

4. S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th Ed., Academic Press, 2009.

5. Gonzalez R. C. & Woods R. E., Digital Image Processing, 3rd ed, PHI Learning, 2008

6. Jain A K, Fundamentals of Digital Image Processing, Prentice-Hall India, 2007.

7. Bernd Jahne, Horst Haubecker, "Computer Vision and Applications", Academic Press.

8. David A. Forsyth, Jean Ponce, "Computer Vision: A Modern Approach", 2nd Ed., 2011.

9. Richard Szeliski, "Computer Vision: Algorithms and Applications", Springer, 1st Ed., 2010.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS004	SEMANTIC WEB ARCHITECTURE	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: This course helps the learners to understand the advanced concepts in emerging trends in Web architecture. The learner will be able to grasp the basics of metadata, OWL, RDF and ontologies. This course enables the learners to build an ontology for a given domain.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Distinguish between Traditional web Environment and Semantic web Environment and in the behavior of a search engine. **(Cognitive Knowledge Level: Understand)**
CO 2	Recognize and Apply the concept of RDF and RDFS and the benefits of Ontology and Taxonomy **(Cognitive Knowledge Level: Apply)**
CO 3	Identify how to use OWL to rewrite the ontology and new features in OWL. **(Cognitive Knowledge Level: Analyze)**
CO 4	Analyze the Web services like UDDI and Real World Examples like Swoogle and FOAF. **(Cognitive Knowledge Level: Analyze)**
CO 5	Apply concepts OWL-S for Web service annotation and mapping OWL-S to UDDI **(Cognitive Knowledge Level: Apply)**
CO 6	Design, develop, implement or present innovative ideas on Semantic Web Architecture and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓		✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓		✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Give the specific instances to illustrate the issue in the markup of the webpage?
2. “The Semantic Web is an extension of the current Web”. Justify the statement.
3. Determine how metadata can be embedded in to existing web pages

Course Outcome 2 (CO2):

1. Differentiate RDF and RDFS

2. A Resource in the world of RDF schema has the same semantics as "class". Explain the statement with suitable example.
3. Illustrate the main differences in the concept of Taxonomy and Ontology.

Course Outcome 3(CO3):

1. Compare and contrast OWL and RDF?
2. OWL has created a new class called "owl:Class" to define classes in OWL documents; it is a subclass of "rdfs:Class". The relationship between all these top classes is summarized in Figure. Define all the classes in our camera ontology

Course Outcome 4 (CO4):

1. Determine in what manner ranking of documents using metadata done in Swoogle.
2. Illustrate why the Person class is the core of the FOAF vocabulary. Explain with an Example
3. Swoogle is expected to be used more often by the researchers and developers in the Semantic Web community. Why?

Course Outcome 5 (CO5):

1. Determine in general what is necessary for the concept of degree of matching? Explain the four degree of match between two concepts.
2. "OWL-S is not about learning a new language; instead, it is about understanding and using three new ontologies." Justify the statement.
3. Discuss the issues of mapping OWL-S profile information into UDDI registry

Course Outcome 6 (CO6):

1. Make a Project on Semantic Web that creates a platform to store data on the Web, build vocabularies, and write rules for handling data. Linked data are empowered by technologies such as RDF & OWL.
2. Take a Seminar on the Topic " Latest technologies in the Semantic Web" (for example Web 3.0).

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES : 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS004

Course Name: SEMANTIC WEB ARCHITECTURE

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. "The Semantic Web is an extension of the current Web". Explain the statement. **(5)**

2. Differentiate RDF and RDFS. **(5)**

3. Compare and contrast OWL and RDF? Illustrate the need for OWL in Semantic Web? **(5)**

4. Determine in what manner ranking of documents using metadata done in Swoogle. **(5)**

5. Discuss the issues of mapping OWL-S profile information into UDDI registry **(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. Compare search engines work in traditional and Semantic web? Explain in detail. **(7)**

7. **A Set of RDF Statements** **(7)**

Subject	**Predicate**	**Object**
mySLR:Nikon-D70	mySLR:weight	1.4 lb
mySLR:Nikon-D70 M	mySLR:pixel	6.1
mySLR:Nikon-D50	mySLR:weight	1.3 lb

What properties did we define in order to describe Nikon D70? Using the table Write a pseudocode to help the computer to answer the question?

8. OWL has created a new class called "owl:Class" to define classes in OWL documents; it is a subclass of "rdfs:Class". The relationship between all these top classes is summarized in Figure. Define all the classes in our camera ontology **(7)**

9. Illustrate why the "Person class" is the core of the FOAF vocabulary. Explain with an Example **(7)**

10 Swoogle is expected to be used more often by the researchers and developers in the Semantic Web community. Why? **(7)**

11 Determine the necessity about the concept of degree of matching? Explain the four degrees of match between two concepts. **(7)**

12 "OWL-S is not about learning a new language; instead, it is about understanding and using three new ontologies." Justify the statement. **(7)**

Syllabus

Module 1 : Introduction to semantic web technology:
Traditional web to semantic web-WWW- First Look at the Semantic Web – meta data- Search Engine for the Traditional Web- Search Engine for the Semantic Web

Module 2: Resource Description Framework
Elements - Resource – Property- Statement -rules of RDF – tools- RDFS core elements-- Syntax and Examples - More about Properties - XML Schema and RDF Schema –Taxonomy and ontology concepts.

Module 3: Web ontology language: OWL:
Define classes- set operators –enumerations- defining properties- Symmetric Properties- Transitive Properties- Functional Properties- Inverse Property- Inverse Functional Property- Validating OWL ontology- Related Development Tools- Validate OWL Ontology by Using Web Utilities- Using Programming APIs to Understand OWL Ontology.

Module 4: Web services and Real world examples: Web services – web services standards – web services to semantic web services- UDDI. Swoogle- architecture and usage of meta data; FOAF – vocabulary – creating documents – overview of semantic markup – semantic web search engines- Implementation Details.

Module 5: Concept of OWL-S – building blocks of OWL-S- OWL-S Profile Ontology- OWL-S Process Ontology- OWL-S Grounding Ontology - mapping OWL-S to UDDI - WSDL

Course Plan

No	Topic	No. Of Lectures (40 Hours)
1	**Module 1: Introduction to semantic web technology**	**6**
1.1	Traditional web to semantic web	1
1.2	World Wide Web	1
1.3	First Look at the Semantic Web	1
1.4	Meta data	1
1.5	Search Engine for the Traditional Web	1
1.6	Search Engine for the Semantic Web	1
2	**Module 2: Resource Description Framework**	**10**
2.1	Elements- Resource	1
2.2	Property- Statement	1
2.3	Rules of RDF	1
2.4	Tools	1
2.5	RDFS core elements	1
2.6	Syntax and Examples	1
2.7	Syntax and Examples	1
2.8	More about Properties	1
2.9	XML Schema and RDF Schema	1
2.10	Taxonomy and ontology concepts .	1
3	**Module 3: Web ontology language: OWL**	**9**
3.1	OWL: define classes	1
3.2	Set operators, enumerations	1
3.3	Defining properties- Symmetric Properties	1
3.4	Transitive Properties- Functional Properties	1
3.5	Inverse Property- Inverse Functional Property	1
3.6	Validating OWL ontology.	
3.7	Related Development Tools	1
3.8	Validate OWL Ontology by Using Web Utilities	1
3.9	Using Programming APIs to Understand OWL Ontology	1
4	**Module 4: Web services and Real world examples**	**10**
4.1	Web services	1
4.2	Web services standards	1
4.3	Web services to semantic web services	1
4.4	UDDI	1
4.5	Swoogle	1
4.6	Architecture and usage of meta data	1
4.7	FOAF – vocabulary – creating documents	1

4.8	Overview of semantic markup	1
4.9	Semantic web search engines.	1
4.10	Implementation Details	1
5	**Module 5: Concept of OWL-S**	**5**
5.1	Concept of OWL-S	1
5.2	Building blocks of OWL-S- OWL-S Profile Ontology	1
5.3	OWL-S Process Ontology- OWL-S Grounding Ontology	1
5.4	Mapping OWL-S to UDDI	1
5.5	WSDL	1

References

1. Liyang Yu, Introduction to the Semantic Web and Semantic web services. Chapman & Hall/CRC, Taylor & Francis group, 2007.

2. Johan Hjelm. Creating the Semantic Web with RDF. Wiley, 2001

3. Grigoris Antoniou and Frank van Harmelen. A Semantic Web Primer. MIT Press

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS005	PROGRAM ANALYSIS	PROGRAM ELECTIVE 3	3	0	0	3

Preamble: The course enables the learners to perform static analysis of programs. It explores methods such as abstract interpretation, intraproceduttral and pointer analysis, program slicing and type checking. The learner will be able to do program analysis in several applications such as compilers, tools that help programmers understand and modify programs, and tools that help programmers verify that programs satisfy certain properties of interest.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Examine the various mathematical concepts needed for program analysis, including lattice, chains and Galois connections. (**Cognitive Knowledge Level: Analyze**)
CO 2	Investigate the role of abstract interpretation in static analysis of programs. (**Cognitive Knowledge Level: Analyze**)
CO 3	Use various methods such as interprocedural analysis, pointer analysis and program slicing for program analysis. (**Cognitive Knowledge Level: Analyze**)
CO 4	Use simply typed lambda calculus for developing type checking systems. (**Cognitive Knowledge Level: Analyze**)
CO 5	Design, develop and implement solutions based on the concepts of program analysis. (**Cognitive Knowledge Level: Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tools to model, analyse and solve practical engineering problems.

stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Show that the inverse of a partial order is a partial order.

Course Outcome 2 (CO2):

1. For a statement n: x := y+5, considering the Constant Propagation (CP)transfer function f_n, show an element d_1 belonging to the CP abstract lattice such that $\gamma(f_n(d_1)) = nstate'(\gamma(d_1))$.

Course Outcome 3(CO3):

1. Using program slicing, show that the following program segments are equivalent.

c = initialState(c)
 	
i = 0;
 	
j = 0;
 	
while (i < 100) {
 	
  i = i+2;
 	
  j = j-2;
 	
  c = c+i+j;
 	
}
 	
finalUse(c);
	c= initialState(c)
	
	i = 0;
	
	j = 0;
	
	while (i < 100)
	
	  { j = j-2;
	
	  i = i+2;
	
	  c = c+i+j;
	
	}
	
	finalUse(c);

Course Outcome 4 (CO4):

1. Prove that the type of a variable is preserved under substitution.

Course Outcome 5 (CO5):

1. Given a Java program, implement a tool that performs interval analysis.

Model Question Paper

Second Semester M. Tech. Degree Examination, Month, Year

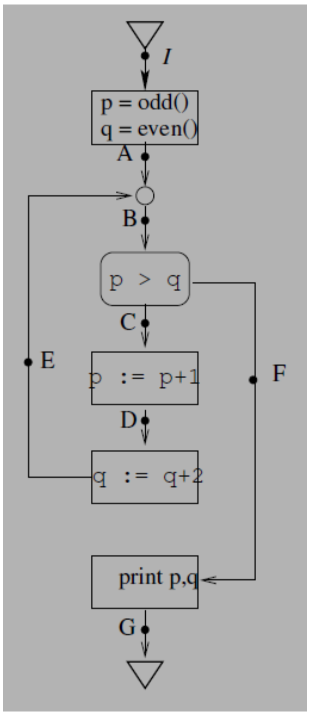
242ECS005 – Program Analysis

Time: 2.5 Hours **Max. Marks: 60**

Part A

(Answer all questions. Each question carries 5 marks) 5x5 = 25 Marks

1. For a partially ordered set$(L, \leq)$, prove that the following claims are equivalent:
    a. L is a complete lattice.
    b. Every subset of L has a least upper bound
    c. Every subset of L has a greatest lower bound.
2. Give the collecting semantics of the following program.



3. What is the problem associated with the Join Over All Paths (JOP) approach in analysing programs with procedure calls? Explain with an example.
4. Illustrate strong and weak updates in points-to analysis.
5. Suppose *t* is a closed, well-typed term. Then show that either *t* is a value or there is some *t'* with $t \rightarrow t'$.

Part B

(Answer any 5 questions. Each question carries 7 marks) 5x7 = 35 Marks

6. What is a chain? Give an example. Prove that a partially ordered set $(L, \leq)$ has finite height if and only if it satisfies both the Ascending Chain and Descending Chain Conditions.
7. Show that if (α, β) is a Galois connection, then α uniquely determines β and, similarly, β uniquely determines α. (By the statement α uniquely determines β, we mean that if (α, β) and (α, β') are Galois connections, then $\beta = \beta'$.)
8. Define consistent abstraction? Give a suitable example.
9. Give Kildall's algorithm to compute over-approximation of Join Over All Paths (JOP). Apply it on the Constant Propagation Lattice.
10. Explain the functional approach to static analysis of programs with procedure calls.
11. With a suitable example, explain the procedure to compute a program slice.
12. Give the algorithm for type checking. Comment on the soundness of the algorithm.

Module 1 - Mathematical Foundations for Program Analysis

–Revision of Partially Ordered Sets, Lattice, Chains, Fixed Points, Knaster-Tarski Fixed Point Theorem, Galois Connections and their properties, Introduction to Program Analysis.

Module 2- Data Flow Analysis / Abstract Interpretation

Collecting Semantics, Abstract Interpretation, Join Over all Paths, Abstract Interpretation for Constant propagation, Correctness of Abstract Interpretation, Kildall's algorithm.

Module 3- Interprocedural Analysis

Call Strings approach, Join OverInterprocedurally Valid Paths, Sharir and Pneuli's approaches to Interprocedural Analysis, Functional Approach.

Module 4 - Pointer Analysis, PDGs and Slicing

Pointer Analysis, Program Dependence Graph, Computing a Program Slice.

Module 5 - Simply Typed Lambda Calculus

Introduction to Lambda Calculus, Type Systems, Algorithms for Type Checking.

Course Plan

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Mathematical Foundations for Program Analysis**	**7**
1.1	Revision of Partially Ordered Sets, Complete Lattice	1
1.2	Chains, Fixed Points	1
1.3	Knaster-Tarski Fixed Point Theorem	1
1.4	Galois Connections	1
1.5	Properties of Galois Connections	1
1.6	Introduction to Program Analysis (Lecture 1)	1
1.7	Introduction to Program Analysis (Lecture 2)	1
2	**Module 2: Data Flow Analysis / Abstract Interpretation**	**8**
2.1	Collecting Semantics	1
2.2	Abstract Interpretation, Collecting abstract values	1
2.3	Comparison of abstract Join Over all Paths (JOP) states and collecting states	1
2.4	Abstract Interpretation for Constant Propagation	1
2.5	Correctness of Abstract Interpretation using Galois Connections	1
2.6	Kildall's algorithm for over-approximate JOP (Lecture 1)	1
2.7	Kildall's algorithm for over-approximate JOP (Lecture 2)	1

2.8	Kildall's algorithm for over-approximate JOP (Lecture 2)	1
3	**Module 3: Interprocedural Analysis**	**8**
3.1	Sharir-Pnueli's Call-Strings approach handling programs with procedure calls, Problem with JOP	1
3.2	Interprocedurally valid paths and their call-strings, Join Over Interprocedurally-valid Paths (JVP)	1
3.3	Sharir and Pnueli's approaches to Interprocedural Analysis (Lecture 1)	1
3.4	Sharir and Pnueli's approaches to Interprocedural Analysis (Lecture 2)	1
3.5	Sharir and Pnueli's approaches to interprocedural Analysis (Lecture 3)	1
3.6	Sharir-Pneuli's Functional Approach – Equation Solving Approach (Lecture 1)	1
3.7	Equation Solving Approach (Lecture 2)	1
3.8	Iterative / Tabular Approach	1
4	**Module 4: Pointer Analysis, PDGs and Slicing**	**9**
4.1	Points-To Analysis	1
4.2	May-Point-To Analysis	1
4.3	Andersen's Analysis	1
4.4	Program Dependence Graph (PDG)	1
4.5	Def-order Dependencies	1
4.6	Sequence of values at a node, Adequacy of PDGs, PDG Isomorphism	1
4.7	Computing a Program Slice (Lecture 1)	1
4.8	Computing a Program Slice (Lecture 2)	1
4.9	Computing a Program Slice (Lecture 3)	1
5	**Module 5: Simply Typed Lambda Calculus**	**7**
5.1	Introduction, Syntax, Semantics	1
5.2	Examples	1
5.3	Type Systems	1
5.4	Typing Rules, Using a derivation tree to prove that a term is well-typed	1
5.5	Properties – Progress	1
5.6	Properties - Preservation	1
5.7	Algorithm for type checking	1
5.8	Soundness of the algorithm	1

References

1. Nielson, Nielson, and Hankin, Principles of Program Analysis, Springer-Verlag.
2. Alfred Tarski, A lattice-theoretic fixpoint theorem and its applications, Pacific J. Mathematics, 5, pages 285--309, 1955.
3. Gary A. Kildall, A unified approach to global program optimization. In POPL '73: Proceedings of the 1st annual ACM SIGACT-SIGPLAN symposium on Principles of programming languages, pages 194--206, New York, NY, USA, 1973. ACM Press.
4. Anders Moller and Michael I. Schwartzbach, Static Program Analysis.
5. MichaSharir and Amir Pnueli: Two approaches to interprocedural data-flow analysis, in Program Flow Analysis, Ed. Muchnik and Jones (1981).
6. Tom Reps, Susan Horwitz and MoolySagiv, Precise interprocedural dataflow analysis via graph reachability, In POPL '95, pages 49-61.
7. Manuvir Das, Sorin Lerner and Mark Seigle, ESP: path-sensitive program verification in polynomial time, ACM SIGPLAN Notices, Volume 37, Issue 5, May 2002, pages 57–68
8. Thomas Reps and Wuu Yang, The Sematics of Program Slicing, University of Wisconsin – Madison.
9. S. Horwitz, T. Reps, and D. Binkley. Interprocedural slicing using dependence graphs, Wisconsin – Madison.
10. Jeanne Ferrante, Karl J Ottenstein, and Joe D Warren, The program dependence graph and its use in optimization, ACM Transactions on Programming Languages and Systems, Vol. 9, No. 3, July 1987, Pages 319 – 349.
11. Karl J Ottenstein and Linda M Ottenstein, The program dependence graph in a software development environment.
12. Benjamin C. Pierce, "Types and Programming Languages". Relevant chapters: Chapters 1-4 (preliminaries), 5 and 6 (basics of lambda calculus), and 8-9 (simple type systems, including simply typed lambda calculus).
13. L. Damas and R. Milner, Principal type-schemes for functional programs, in POPL '82, pages 207--212.

SEMESTER II

PROGRAM ELECTIVE IV

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS006	BLOCKCHAIN TECHNOLOGY AND IOT	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: The objective of this course is to enable to the student to build an IoT application. It also provides conceptual understanding of how blockchain technology can be used to innovate and improve business processes. This course covers the interconnection and integration of the physical world and the cyber space. This course will provide the students with the advanced competitive skills required to contribute to the development of the IoT.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Analyze various protocols for IoT (**Cognitive Knowledge Level: Analyze**)
CO 2	Design IoT based applications using Arduino or Raspberry PI boards. (**Cognitive Knowledge Level: Apply**)
CO 3	Identify the need of blockchains to find the solution to the real-world problems. (**Cognitive Knowledge Level: Analyze**)
CO 4	Recognize the underlying technology of transactions, blocks, proof-of-work, and consensus building. (**Cognitive Knowledge Level: Analyze**)
CO 5	Design and implement new ways of using blockchain for real time applications. (**Cognitive Knowledge Level: Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓			
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓		✓	
CO 4	✓		✓	✓		✓	
CO 5	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

comprehension, application, analysis, synthesis, evaluation and understanding of the students.
End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Describe how the environment can be more protected with the help of IoT technology in the following categories: (i) Air pollution monitoring (ii) Noise pollution monitoring (iii) Forest fire detection (iv) River flood detection.

Course Outcome 2 (CO2):

1. Design an automatic refrigerator light system with LED, switch & raspberry pi and write a python program to support the working of that design.

Course Outcome 3(CO3):

1. Identify the steps that are involved in the Blockchain project implementation.

Course Outcome 4 (CO4):

1. Write a crowd-sale smart contract code in Solidity programming language.

Course Outcome 5 (CO5):

1. Design and implement private blockchain for any given enterprise use-cases.

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS006

Course Name: BLOCKCHAIN TECHNOLOGY AND IOT

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Explain how Information is shared between the devices through IoT.	**(5)**
2.	Illustrate the role of actuators in IoT?	**(5)**
3.	Define interoperability in the context of blockchain implementation.	**(5)**
4.	How does Bitcoin use Blockchain?	**(5)**
5.	Describe the main Use Cases of Solidity?	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.		Formulate the significant use of Raspberry Pi in Smart cities and Industrial appliances.	**(7)**
7.		Design the protocol layer of IoT and explain various protocols used in each layer.	**(7)**
8.		Construct the Design of Smart home with Raspberry Pi and other hardware devices with a neat sketch.	**(7)**
9.		Formulate the significant use of Raspberry Pi in Smart cities and Industrial appliances.	**(7)**
10	(a)	Describe how the PoW consensus works.	**(4)**
	(b)	An organization wants to develop smart contracts, based on blockchain technology. The organization does not wish to burden employees with maintaining the security of the blockchain. What blockchain technology fits the organization best?	**(3)**
11		Define a Merkle hash tree, describe its role in blockchain, and explain the meaning of a Merkle root in the block header.	**(7)**

12 How Blockchain is Revolutionizing the Traditional Business Network? Explain with examples. **(7)**

Syllabus

Module 1: Introduction to IoT

Sensor basics, sensing and actuation, basics of networking - wired, wireless, MANET, PAN, wireless and wired protocols, Communication protocols - IEEE standards, 5G era, sensor communications, connectivity challenges, fading and attenuations.

Module 2: IoT Architecture and Programming

basic architectures, Data processing mechanisms, scalability issues and visualization issues, analytic basics, utility of cloud computing, fog computing, edge computing, Raspberry Pi and Arduino programming, Applications- IoT for industrial automation (Industry 4.0).

Module 3: Introduction to Blockchain

Blockchain concepts, evolution, structure and characteristics, benefits and challenges, Blockchain as public ledgers - Transactions, Elements of Cryptography - Cryptographic Hash functions, Merkle Tree.

Module 4: Blockchain architecture and Use Cases

Design methodology for Blockchain applications, Blockchain application templates, Blockchain application development, Ethereum, Solidity,Bitcoin, Sample use cases from Industries,Business problems.

Module 5: Smart Contracts and Decentralized Applications (DApps)

Smart contract smart contract examples, structure of a contract, smart contract examples, smart contract patterns, implementing Dapps, Ethereum Dapps, case studies related to Dapps.

Course Plan

No	Topic	No. of Lecture (40 hrs)
1	**Module 1: Introduction to IoT**	**10**
1.1	Sensor basics	1
1.2	sensing and actuation	1
1.3	basics of networking - wired, wireless	1
1.4	MANET	1
1.5	PAN	1

1.6	Wireless and wired protocols	1
1.7	Communication protocols - IEEE standards	1
1.8	5G era, sensor communications	1
1.9	Connectivity challenges	1
1.10	Fading and attenuations	1
2	**Module 2: IoT architectures and programming**	**10**
2.1	Basic architectures	1
2.2	Data processing mechanisms	1
2.3	Scalability issues and visualization issues	1
2.4	Analytic basics	1
2.5	Utility of cloud computing	1
2.6	Fog computing	1
2.7	Edge computing	1
2.8	Raspberry Pi and Arduino programming	1
2.9	Applications	1
2.10	IoT for industrial automation (Industry 4.0)	1
3	**Module 3: Introduction to Blockchain**	**6**
3.1	Blockchain concepts	1
3.2	Evolution, structure and characteristics	1
3.3	Benefits and challenges	1
3.4	Blockchain as public ledgers - Transactions	1
3.5	Elements of Cryptography - Cryptographic Hash functions	1
3.6	Merkle Tree	1
4	**Module 4: Blockchain Architecture and Use cases**	**7**
4.1	Design methodology for Blockchain applications	1
4.2	Blockchain application templates	1
4.3	Blockchain application development	1
4.4	Ethereum, Solidity	1
4.5	Bitcoin	1
4.6	Sample use cases from Industries	1

4.7	Business problems	1

5	**Module 5: Smart contracts and Decentralized applications (Dapps)**	**7**
5.1	Smart contract smart contract examples	1
5.2	Structure of a contract	1
5.3	Smart contract examples	1
5.4	Smart contract patterns	1
5.5	Implementing Dapps	1
5.6	Ethereum Dapps	1
5.7	Case studies related to Dapps	1

References

1. The Internet of Things: Enabling Technologies, Platforms, and Use Cases, Pethuru Raj and Anupama C. Raman, CRC Press, First edition, 2017.

2. The Internet of Things in the Cloud: A Middleware Perspective, Honbu Zhou, CRC press, First edition, 2012.

3. Internet of Things: A Hands-on Approach, Arshdeep Bahga and Vijay Madisetti, Universities Press, First edition, 2014.

4. Fog for 5G and IoT (Information and Communication Technology Series, Mung Chiang, Bharath Balasubramanian, Flavio Bonomi, Wiley series, First edition, 2017.

5. Blockchain applications: a hands-on approach, Bahga A., Madisetti V., VPT, 2017.

6. Beginning Blockchain, A Beginner's Guide to Building Blockchain Solutions, Bikramaditya Singhal, Gautam Dhameja, Priyansu Sekhar Panda, Apress, 2018.

7. Blockchain A Practical Guide to Developing Business, Law, and Technology Solutions, Joseph J. Bambara and Paul R. Allen, McGraw Hill, 2018.

8. Blockchain enabled Applications Vikram Dhillon, David Metcalf and Max Hooper, Apress, 2017.

9. The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology, William Mougayar, Wiley, 2016.

10. Blockchain Science: Distributed Ledger Technology, Roger Wattenhofer, Inverted Forest Publishing; 3rd edition, 2019.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS007	SOCIAL NETWORK ANALYSIS	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: This course provides an exposure to the concepts and techniques in Social Network Analysis. This course covers various types of modelling, visualization and mining techniques used in social networks. This course helps the learners to analyse social media data using appropriate data/web mining techniques.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Understand the concepts and properties of social networks. **(Cognitive Knowledge Level: Understand)**
CO 2	Analyze the concepts of evolution and privacy in social networks. **(Cognitive Knowledge Level: Analyze)**
CO 3	Model and visualize social networks. **(Cognitive Knowledge Level: Apply)**
CO 4	Mine the behaviour of users in the Social Networks. **(Cognitive Knowledge Level: Analyze)**
CO 5	Use Multimedia Information Networks in Social Media. **(Cognitive Knowledge Level: Analyze)**
CO 6	Design, Develop, Implement and Present innovative ideas on Social network analysis concepts and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tools to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects.

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓					✓	
CO 2	✓		✓		✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	✓
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task/ Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain the properties of dynamic weighted and unweighted graphs.

Course Outcome 2 (CO2):

1. Describe privacy breaches and privacy preserving mechanisms in social networks.

Course Outcome 3(CO3):

1. Classify the visualization of social networks.

Course Outcome 4 (CO4):

1. Given a social network with labels on some nodes, how to provide a high quality labelling for every node?
2. What are the types of changes in the evolution of the web community?

Course Outcome 5 (CO5):

1. Describe the different kinds of tags.
2. Explain the applications of tags.

Model Question Paper

QP CODE:

Reg No: _______________

Name: **PAGES: 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS007

Course Name: Social Networks Analysis

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1. What are the static properties of unweighted graphs? **(5)**

2. Briefly describe various Node-Edge Diagrams **(5)**

3. Explain the techniques for keyword search over graph data **(5)**

4. What are the applications of community mining algorithms? **(5)**

5. Differentiate between Categorizers and Describers. **(5)**

PART B

(Answer any five questions. Each question carries 7 marks)

6.		Explain the properties of dynamic weighted and unweighted graphs.	**(7)**
7.		Describe privacy breaches and privacy preserving mechanisms in social networks.	**(7)**
8.		Classify the visualization of social networks.	**(7)**
9.	(a)	Given a social network with labels on some nodes, how to provide a high quality labelling for every node?	**(3)**
	(b)	What are the types of changes in the evolution of the web community?	**(4)**
10	(a)	Describe the different kinds of tags.	**(4)**
	(b)	Explain the applications of tags.	**(3)**
11		Describe Ethnography, Netnography and event maps	**(7)**
12	(a)	Multimedia information networks can be viewed as a marriage of multimedia content and social networks. Justify.	**(3)**
	(b)	Explain ontology based learning.	**(4)**

Syllabus

Module 1: Introduction

Introduction to Social Network Data Analytics: Introduction, Online Social Networks: Research Issues, Research Topics in Social Networks. Statistical Properties of Social Networks: Preliminaries, Static Properties, Dynamic Properties. Random Walks in Social Networks and their Applications: Random Walks on Graphs: Background, Application in Computer Vision, Text Analysis, Collaborative Filtering.

Module 2: Evolution in Social Networks

Evolution in Social Networks: Framework, Challenges of Social Network Streams, Incremental Mining for Community Tracing, Tracing Smoothly Evolving Communities. Models and Algorithms for Social Influence Analysis: Influence Related Statistics, Social Similarity and Influence. Privacy in Social Networks: Privacy breaches in social networks, Privacy-preserving mechanisms.

Module 3: Visualizing Social Networks

Visualizing Social Networks: A Taxonomy of Visualizing Social Networks: A Taxonomy of Visualizations. Data Mining in Social Media: Methods for Social Media, Ethnography and

Netnography, Event Maps. Text Mining in Social Networks: Keyword Search, Classification and Clustering Algorithms, Transfer Learning in Heterogeneous Networks.

Module 4: Mining Communities

Aggregating and reasoning with social network data, Advanced Representations - Extracting evolution of Web Community from a Series of Web Archive -Detecting Communities in Social Networks - Evaluating Communities – Core Methods for Community Detection & Mining - Applications of Community Mining Algorithms - Node Classification in Social Networks

Module 5: Multimedia Information Networks in Social Media

Multimedia Information Networks in Social Media: Links from Semantics, Links from Community Media. Network of Personal Photo Albums, Geographical Information, Inference Methods. Social Tagging and Applications: Tags: Why What, Tagging System Design, Tag analysis, Visualization of Tags, Applications of Tags.

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Introduction**	**8**
1.1	Introduction to Social Network Data Analytics	1
1.2	Online Social Networks: Research Issues	1
1.3	Research Topics in Social Networks	1
1.4	Statistical Properties of Social Networks: Preliminaries, Static Properties, Dynamic Properties	1
1.5	Random Walks in Social Networks and their Applications	1
1.6	Random Walks on Graphs: Background	1
1.7	Application in Computer Vision	1
1.8	Text Analysis, Collaborative Filtering	1
2	**Module 2: Evolution in Social Networks**	**8**
2.1	Framework	1
2.2	Challenges of Social Network Streams	1
2.3	Incremental Mining for Community Tracing	1
2.4	Tracing Smoothly Evolving Communities	1
2.5	Models and Algorithms for Social Influence Analysis: Influence Related Statistics	1
2.6	Social Similarity and Influence	1
2.7	Privacy in Social Networks: Privacy breaches in social networks	1

2.8	Privacy-preserving mechanisms	1
3	**Module 3: Visualizing Social Networks**	**8**
3.1	A Taxonomy of Visualizing Social Networks	1
3.2	A Taxonomy of Visualizations	1
3.3	Data Mining in Social Media: Methods for Social Media	1
3.4	Ethnography and Netnography, Event Maps	1
3.5	Text Mining in Social Networks: Keyword Search	1
3.6	Text Mining in Social Networks: Keyword Search	1
3.7	Classification and Clustering Algorithms	1
3.8	Transfer Learning in Heterogeneous Networks	1
4	**Module 4: Mining Communities**	**8**
4.1	Aggregating and reasoning with social network data	1
4.2	Advanced Representations - Extracting evolution of Web Community from a Series of Web Archive	1
4.3	Detecting Communities in Social Networks	1
4.4	Evaluating Communities	1
4.5	Core Methods for Community Detection & Mining	1
4.6	Core Methods for Community Detection & Mining	1
4.7	Applications of Community Mining Algorithms	1
4.8	Node Classification in Social Networks	1
5	**Module 5: Multimedia Information Networks in Social Media**	**8**
5.1	Links from Semantics, Links from Community Media	1
5.2	Network of Personal Photo Albums	1
5.3	Geographical Information	1
5.4	Inference Methods	1
5.5	Social Tagging and Applications: Tags: Why What	1
5.6	Tagging System Design	1
5.7	Tag analysis	1
5.8	Visualization of Tags, Applications of Tags	1

References

1. Charu C. Aggarwal, “Social Network Data Analytics”, Springer.
2. Peter Mika, “Social Networks and the Semantic Web”, Springer, 1St edition 2007.
3. BorkoFurht, “Handbook of Social Network Technologies and Applications”, Springer, 1st edition, 2010.
4. GuandongXu, Yanchun Zhang and Lin Li, “Web Mining and Social Networking Techniques and applications”, Springer, 1St edition, 2011.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS008	MODERN DATABASE MANAGEMENT	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: This course provides an exposure to the concepts and techniques in modern data management. Different types of NoSQL databases, their architecture and use cases are discussed in this course. A better understanding of data management is provided through MongoDB, Cassandra and Neo4J databases. This course helps the learners to develop applications that manage data efficiently with the help of suitable data models and techniques.

Course Outcomes: After the completion of the course the student will be able to:

CO 1	Analyze issues and solutions in modern databases. **(Cognitive Knowledge Level: Understand)**
CO 2	Employ and operate on document databases and techniques. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply spatial database concepts to efficiently organise and retrieve spatial data. **(Cognitive Knowledge Level: Apply)**
CO 4	Research, analyze and use emerging technologies in column store. **(Cognitive Knowledge Level: Analyze)**
CO 5	Practice techniques in Graph database. **(Cognitive Knowledge Level: Analyze)**
CO 6	Design, Develop, Implement and Present innovative ideas on modern database concepts and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**	**PO 7**
CO 1	✔		✔	✔		✔	
CO 2	✔		✔	✔	✔	✔	
CO 3	✔		✔	✔	✔	✔	
CO 4	✔		✔	✔	✔	✔	
CO 5	✔		✔	✔	✔	✔	
CO 6	✔	✔	✔	✔	✔	✔	✔

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Give specific instances to illustrate vertical and horizontal fragmentation in distributed databases.
2. The figure shows version v1 of a system. Values 'berlin' and '142' are assigned to city and number fields of the address object respectively.

i. Explain how the system can maintain its consistency using object versioning after the change.

Course Outcome 2 (CO2):

1. While creating Schema in MongoDB what are the points that need to be taken in consideration? Explain with an example.

Course Outcome 3(CO3):

1. Use R-Tree to store the following points in which only 3 points/node. (10,10), (15,10), (20,10), (30,20), (50,10), (50,80), (60,30), (60,60), (70,20), (75,10).

Course Outcome 4 (CO4):

1. How does Cassandra handle large volumes of data efficiently? Describe the use of Sorted String Tables and consistent hashing with virtual nodes in Cassandra.

Course Outcome 5 (CO5):

1. Write the Neo4j Cypher Query which returns all the players (nodes) that belong to the country India and have scored runs greater than 210. Assume that the below nodes are created using Neo4j.

CREATE (Dhawan: player {name: "Shikar Dhawan", YOB: 1985, runs:363, country: "India"})

CREATE (Jonathan: player {name: "Jonathan Trott", YOB:1981, runs:229, country: "South Africa"})

CREATE (Sangakkara: player {name: "KumarSangakkara", YOB:1977, runs:222, country: "Sri Lanka"})

CREATE (Virat: player {name:"ViratKohli", YOB: 1988, runs:176, country: "India"})

Course Outcome 6 (CO6):

1. Implement E-Commerce stream analysis using Cassandra.

Model Question Paper

QP CODE:

Reg No: _______________

Name: __________________ **PAGES: 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS008

Course Name: Modern Database Management

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

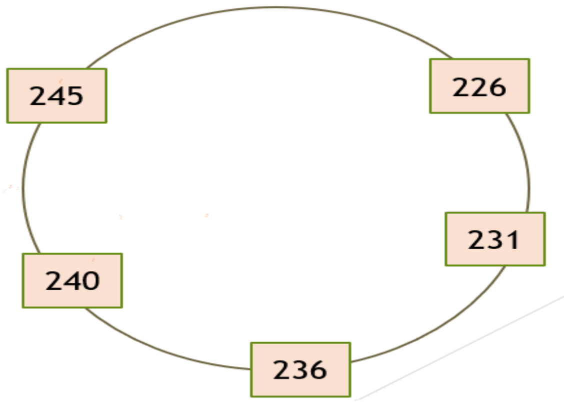
Answer All Questions. Each Question Carries 5 Marks

1.	Give specific instances to illustrate vertical and horizontal fragmentation in distributed databases.	**(5)**
2.	How does MongoDB provide consistency? Explain with the help of a scenario.	**(5)**
3.	What is the purpose of spatial indexing? How spatial indexing is useful in geospatial applications like Uber ride sharing apps.	**(5)**
4.	How Cassandra implements Hinted Handoff. Illustrate with an example.	**(5)**
5.	Suggest a suitable data model for storing the details of movies, actors with queries list all pairs of actors who have co-starred, movies of a particular director in a particular year, build a simple recommendation engine, starting from a known movie that you liked, and find similar movies that you might also enjoy. Give justification for your answer.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6. (a) In hotel taj each customer is served by a waiter according to consistent hashing . Each waiter keeps the information about their customers' favourite coffee with them. Each customer is mapped to their corresponding waiter using a hash function. The key values represent the waiter. 1



Hash function:

H(x)=(x-65) mod 26 x=key

H(x)=(n-65) mod 26 n= node

i. Map the following customers to their corresponding waiters. [F, G, P, S, T]

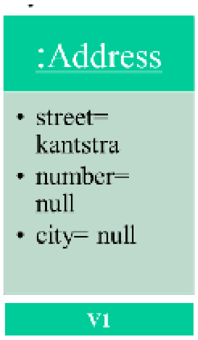
[Hint: ASCII- F:70, G:71, P:80, S:83, T:84]

ii. Employee 236 takes break and customer ' L' appears

iii. What is the arrangement when new employee 228 is added?

iv. If that waiter went for a smoke break while a customer arrives, how will the situation be handled?

(b) The figure shows version v1 of a system. Values ‘berlin’ and ‘142’ are assigned t city and number fields of the address object respectively. **(3)**



i. Explain how the system can maintain its consistency using object versioning after the change.

ii. Draw version 2 of the system.

7. While creating Schema in MongoDB what are the points need to be taken in consideration? Explain with an example. **(7)**

8. Use R-Tree to store the following points in which only 3 points/node. (10,10), (15,10), (20,10), (30,20), (50,10), (50,80), (60,30), (60,60), (70,20), (75,10). **(7)**

9. How does Cassandra handle large volumes of data efficiency . Describe the use of Sorted String Tables and consistent hashing with virtual nodes in Cassandra.

10 How is a bloom filter space efficient? Do blooms filters have false negatives? Consider a Bloom filter that consists of m= 13 memory bits and hash function f () defined as f(k)=(3*k) mod m where k is the given key. Assume that all m bits of the Bloom filter are initially set to 0. Show the Bloom filter bits following the insertion of the keys 7, 11 and 9. Show result after each insertion. **(7)**

11 Write the Neo4j Cypher Query which returns all the players (nodes) that belong to the country India and have scored runs greater than 210. Assume that the below nodes are created using Neo4j. **(7)**

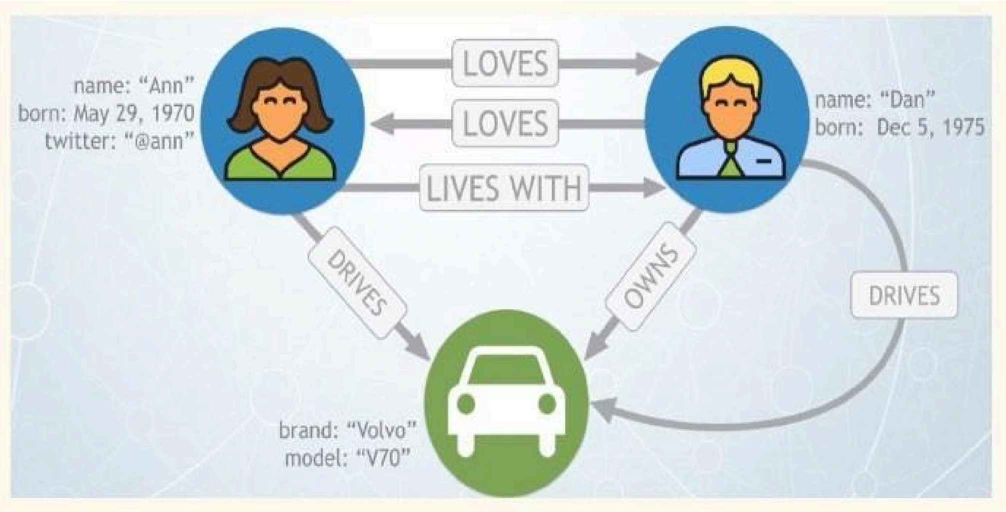
CREATE (Dhawan: player {name: "Shikar Dhawan", YOB: 1985, runs:363, country: "India"})

CREATE (Jonathan: player {name: "Jonathan Trott", YOB:1981, runs:229, country: "South Africa"})

CREATE (Sangakkara: player {name: "KumarSangakkara", YOB:1977, runs:222, country: "Sri Lanka"})

CREATE (Virat: player {name:"ViratKohli", YOB: 1988, runs:176, country: "India"})

12 Analyse the following instance of a graph database: **(7)**



Write Neo4j cypher query for the following cases:

i. Create a relation Love between Persons with names Ann and Dane
ii. Who drives a car owned by a lover
iii. Whom does Ann love?
iv. Find Ann’s car?
v. Describe the car using SET clause (Brand- Volvo, Model-V70)
vi. Ensure uniqueness for the person with name Ann
vii. Create a relationship between Ann and her pet dog with name Sam using MERGE clause

Syllabus

Module 1: NoSQL Databases

Review of Distributed Databases- Fragmentation, Replication, Transparencies in design, CAP theorem- BASE transactions and eventual consistency, consistent hashing, object versioning & vector clocks, Consensus algorithms, Logging & Snapshots, Properties of NoSQL databases, Types of NoSQL Databases.

Module 2: Document based databases

MongoDB- Documents- JSON & BSON format, representing relationships, CRUD operations, Indexing, Aggregation, Sharding architecture and Replication strategies, consistency and locking.

Module 3: Spatial Databases

Types of Spatial Data and Queries- Point and Region Data- Queries, Spatial Indexing: Space Filling Curves- Z ordering, Quad Trees, R-Trees, Geospatial queries & geospatial indexes in MongoDB.

Module 4: Column databases

Column family, Cassandra Architecture-Gossiping, Snitches, Rings and Tokens, Virtual Nodes, Replication Strategies, Consistency, Hinted handoff, Lightweight Transactions, Bloom filter, Compaction; Fault tolerance, Caching, SSTable & MemTable, Cassandra Query Language

Module 5: Graph Databases

Neo4j- Introduction, Example graphs, Data Modeling, Traversal, Indexing, Features, operations, Cypher Queries- Create, Match clause.

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: NoSQL Databases**	**8**
1.1	Review of Distributed Databases	1
1.2	CAP theorem- BASE transactions and eventual consistency	1
1.3	Consistent hashing	1
1.4	Object versioning	1
1.5	Vector clocks	1
1.6	Consensus algorithms	1

1.7	Logging & Snapshots :COMPUTER SCIENCE AND ENGIN	EERI NG-CS 1
1.8	Properties of NoSQL databases, Types of NoSQL Databases.	1
2	**Module 2: Document based databases**	**10**
2.1	Introduction to Document Database & MongoDB	1
2.2	MongoDB Documents- JSON & BSON format	1
2.3	Representing relationships	1
2.4	CRUD operations	1
2.5	Indexing	1
2.6	Aggregation	1
2.7	Sharding architecture	1
2.8	Replication strategies	1
2.9	Consistency	1
2.10	Locking	1
3	**Module 3: Spatial Database**	**8**
3.1	Types of Spatial Data	1
3.2	Types of Spatial Queries	1
3.3	Spatial Indexing	1
3.4	Space Filling Curves- Z ordering	1
3.5	Quad Trees	1
3.6	R-Trees	1
3.7	Geospatial queries in MongoDB	1
3.8	Geospatial indexes in MongoDB	1
4	**Module 4: Column databases**	**8**
4.1	Introduction to Column family & Cassandra, Cassandra Architecture-Gossiping	1
4.2	Snitches, Rings and Tokens	1
4.3	Virtual Nodes & Replication Strategies	1
4.4	Consistency	1
4.5	Hinted handoff & Lightweight Transactions	1
4.6	Bloom filter, Compaction	1
4.7	Fault tolerance, Caching, SSTable & MemTable	1
4.8	Cassandra Query Language	1
5	**Module 5: Graph Databases**	**6**
5.1	Introduction to graph databases	1
5.2	Neo4j, Example graphs	1
5.3	Data Modeling	1
5.4	Traversal, Indexing, Features and operations	1
5.5	Cypher Queries- CREATE	1
5.6	Cypher Queries- MATCH	1

1. Guy Harrison, "Next Generation Databases: NoSQL, NewSQL, and Big Data", Apress, 2016
2. Thomas Cannolly and Carolyn Begg, "Database Systems, A Practical Approach to Design, Implementation and Management", 6/e, Pearson Education, 2015.
3. Henry F Korth, Abraham Silberschatz, S. Sudharshan, "Database System Concepts", 7/e, Tata McGraw Hill, 2019.
4. Pramod Sadalage, Martin Fowler," NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley, 2012
5. Shannon Bradshaw, Eoin Brazil, Kristina Chodorow, "MongoDB: The Definitive Guide", 3/e, O'Reilly, 2019
6. Ian Robinson, Jim Webber, Emil Eifrem, "Graph databases", O'Reilly, 2013.
7. Jeff Carpenter, Eben Hewitt, "Cassandra: The Definitive Guide", 3/e, O'Reilly, 2020
8. Dan Sullivan, "NoSQL for Mere Mortals", Addison-Wesley, 2015
9. Alex Petrov, "Database Internals: A Deep Dive into How Distributed Data Systems Work", O'Reilly, 2019
10. Luc Perkins, Eric Redmond, Jim Wilson, "Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement", 2/e, O'Reilly, 2018
11. Jérôme Baton, Rik Van Bruggen, "Learning Neo4j 3.x", 2/e, Packt, 2017.
12. Aleksa Vukotic, Nicki Watt, "Neo4j in Action", Manning Publications, 2015.
13. Shashi Shekhar, Sanjay Chawla, Spatial databases A Tour, Pearson Education, Indian Edition, First Impression 2009.
14. Web Resource: https://neo4j.com/docs/operations-manual/current/

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS009	DISTRIBUTED ALGORITHMS	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: The purpose of this course is to introduce the distributed algorithms with an emphasis on principles and theory. This course provides the concepts and techniques of distributed algorithms in synchronous and asynchronous distributed computing systems. This course helps the leaners to gain a good idea regarding various system models and their capabilities, which will help to design new algorithms.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Understand the working of problem domain in synchronous distributed computing system. **(Cognitive Knowledge Level: Understand)**
CO 2	Develop algorithms for synchronous distributed computing system**(Cognitive Knowledge Level: Apply)**
CO 3	Examine the consensus problem in a distributed computing system**(Cognitive Knowledge Level: Apply)**
CO 4	Illustrate the use of different communication models in asynchronous distributed computing systems. **(Cognitive Knowledge Level: Apply)**
CO 5	Develop algorithms for various resource allocation problems in Asynchronous distributed computing systems. **(Cognitive Knowledge Level: Apply)**
CO 6	Design, Develop, Implement and Present innovative ideas on distributed algorithms and techniques. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓					✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓		✓	✓	
CO 4	✓		✓		✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

Course based task / Seminar/ Data collection and interpretation :15 marks

Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain the complexity measures of synchronous distributed computing system.
2. Compare invariant assertion and simulations.
3. Illustrate the execution of a system.

Course Outcome 2 (CO2):

1. For the LCR algorithm,
   a. Give a UID assignment for which Ω(n 2) messages are sent.

b. Give a UID assignment for which only O(n) messages are sent.

2. Design a unidirectional leader-election algorithm that works with unknown ring size. Your algorithm should manipulate the UIDs using comparisons only.

3. Suppose that LubyMIS is executed in a ring of size n. Estimate the probability that any particular edge is removed from the graph in one iteration of the algorithm.

Course Outcome 3(CO3):

1. Explain Consensus in a distributed computing system.

2. Compare Byzantine failure and stopping failure.

3. Describe the required conditions for the K agreement problem.

Course Outcome 4 (CO4):

1. Describe a specific I/O automaton having no input actions, whose output actions are {0, 1, 2,... }, and whose fair traces are exactly the sequences in set S, defined as follows. S consists of all the sequences of length 1 over the output set, that is, all the sequences consisting of exactly one nonnegative integer.

2. Write a pre condition effect code for Reliable reordering channel.

Course Outcome 5 (CO5):

1. Describe a fair execution of the DijkstraME algorithm in which a particular process is locked out.

2. Illustrate Right-Left Dining Philosophers Algorithm.

Course Outcome 6 (CO6):

1. Implement FloodMax algorithm in a synchronous distributed computing system.

2. Implement Peterson Leader-Election Algorithm in an asynchronous distributed computing system.

Model Question Paper

QP CODE:

Reg No: _______________

Name: __________________ **PAGES : 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS009

Course Name: DISTRIBUTED ALGORITHMS

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.	Let A be a system of n processes, n > 1, arranged in a bidirectional ring. If all the processes in A are identical, then A does not solve the leader-election problem. Justify	**(5)**
2.	Analyse the complexity of the bellman ford algorithm in a general synchronous network	**(5)**
3.	Determine the required conditions for the Commit problem	**(5)**
4.	Write an I/O automaton A representing a reliable message channel that accepts and delivers messages from the union of two alphabets, M1 and M2. The message channel is supposed to preserve the order of messages from the same alphabet. Also, if a message from alphabet M1 is sent prior to another message from alphabet M2, then the corresponding deliveries must occur in the same order. However, if a message from M1 is sent after a message from M2, then the deliveries are permitted to occur in the opposite order. Your automaton should actually exhibit all of the allowable external behaviours.	**(5)**
5.	Illustrate a fair execution of the DijkstraME algorithm in which a particular process is locked out.	**(5)**

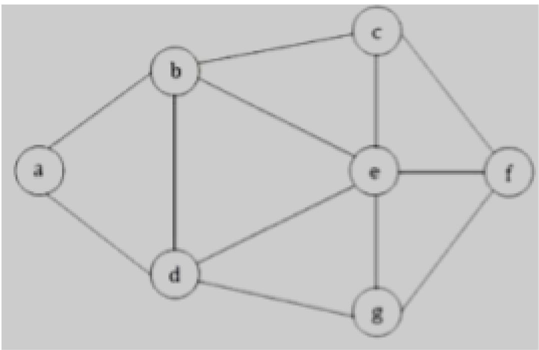
(5x5=25)

Part B

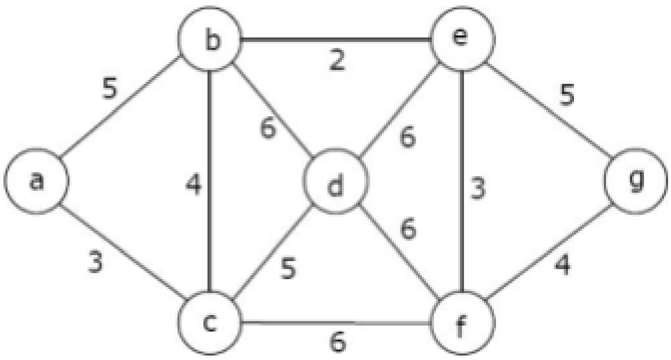
(Answer any five questions. Each question carries 7 marks)

6.	Demonstrate the leader election algorithm with an O(nlogn) communication complexity.	**(7)**

7. Compute the maximal independent of the given graph 1**(7)**



8. Construct the minimum spanning tree of the given graph **(7)**



9. Calculate the largest number of faulty processes that can be tolerated by Byzantine agreement algorithms that run in the following network graphs? **(7)**

i. A ring of size n.

ii. A three-dimensional cube, m nodes on a side, in which nodes are connected only to their neighbours in the three dimensions.

10 Consider the PetersonLeader algorithm in a ring with n = 16 nodes, in which the UIDs for processes P I , . . . , P16 are, respectively, 25, 3, 6, 15, 19, 8, 7, 14, 4, 22, 21, 18, 24, 1, 10, 23. Predict which process is elected as leader? **(7)**

11 Write an algorithm for broadcast and acknowledgment in asynchronous networks, in which the time complexity depends on the number of nodes. **(7)**

12 Demonstrate the precondition-effect code for an implementation of the Dijkstra's Mutual Exclusion Algorithm in the asynchronous network setting. **(7)**

Module 1: Synchronous distributed computing system

Synchronous distributed computing system, Leader Election in a Synchronous Ring, LCR algorithm, HS algorithm, Time Slice Algorithm, Variable Speeds Algorithm, Lower Bound for Comparison-Based Algorithms.

Module 2: Algorithms in General Synchronous Networks

Algorithms in General Synchronous Networks. Leader election in a General Network - Simple Flooding Algorithm, Basic Breadth- First Search Algorithm, Bellman-Ford algorithm, Minimum Spanning Tree, Maximal Independent Set, LubyMIS algorithm.

Module 3: Distributed Consensus

Distributed Consensus, Distributed Consensus with Link Failures – The coordinated Attack Problem-Deterministic version, Randomized version. Distributed Consensus with Process Failures and Process Failures. Algorithms for stopping Failures, Algorithms for Byzantine Failure. Byzantine agreement in general graphs. Weak Byzantine agreement. Consensus Problems-K Agreement, Approximate Agreement, Commit Problem.

Module 4: Asynchronous distributed computing system

Asynchronous distributed computing system, Asynchronous Network Model - Send/Receive systems, Broadcast systems, Multicast systems. Asynchronous Network algorithms- Peterson Leader-Election Algorithm, Local Synchronizer, Safe Synchronizer.

Module 5: Asynchronous Shared Memory Systems

Asynchronous Shared Memory Systems, Environment Model, Shared Variable Types. Mutual Exclusion - Asynchronous Shared Memory Model, Dijkstra's Mutual Exclusion Algorithm. Resource Allocation - Nonexistence of Symmetric Dining Philosophers Algorithms, Right-Left Dining Philosophers Algorithm, Mutual exclusion and Consensus, Relationship between shared memory and network models.

No	Topic	No. of Lectures (40 hrs)
1	**Module 1: Synchronous distributed computing system**	**8**
1.1	Synchronous distributed computing system	1
1.2	Synchronous network model	1
1.3	Leader Election in a Synchronous Ring	1
1.4	LCR algorithm	1
1.5	HS algorithm	1
1.6	Time Slice Algorithm	1
1.7	Variable Speeds Algorithm	1
1.8	Lower Bound for Comparison-Based Algorithms	1
2	**Module 2: Algorithms in General Synchronous Networks**	**8**
2.1	Algorithms in General Synchronous Networks	1
2.2	Leader election in a General Network	1
2.3	Simple Flooding Algorithm	1
2.4	Basic Breadth- First Search Algorithm	1
2.5	Bellman-Ford algorithm	1
2.6	Minimum Spanning Tree	1
2.7	Maximal Independent Set	1
2.8	LubyMIS algorithm.	1
3	**Module 3: Distributed Consensus**	**8**
3.1	Distributed Consensus	1
3.2	The coordinated Attack Problem- Deterministic version, Randomized version	1
3.3	Distributed Consensus with Process Failures and Process Failures	1
3.4	Algorithms for stopping Failures	1
3.5	Algorithms for Byzantine Failure	1
3.6	Byzantine agreement in general graphs	1
3.7	Weak Byzantine agreement	1
3.8	Consensus Problems-K Agreement, Approximate Agreement, Commit Problem	1
4	**Module 4: Asynchronous distributed computing system**	**7**
4.1	Asynchronous distributed computing system	1
4.2	Asynchronous Network Model - Send/Receive systems, Broadcast systems	1
4.3	Multicast systems	1
4.4	Asynchronous Network algorithms	1
4.5	Peterson Leader-Election Algorithm	1
4.6	Local Synchronizer	1
4.7	Safe Synchronizer	1

5	**Module 5: Asynchronous Shared Memory Systems**	**9**
5.1	Asynchronous Shared Memory Systems.	1
5.2	Environment Model	1
5.3	Shared Variable Types	1
5.4	Mutual Exclusion - Asynchronous Shared Memory Model	1
5.5	Dijkstra's Mutual Exclusion Algorithm	1
5.6	Resource Allocation - Nonexistence of Symmetric Dining Philosophers Algorithms	1
5.7	Right-Left Dining Philosophers Algorithm	1
5.8	Mutual exclusion and Consensus	1
5.9	Relationship between shared memory and network models	1

References

1. Nancy A. Lynch, Distributed Algorithms, Morgan Kaufmann Publishers, Inc,1996.
2. Wolfgang Reisig, W. Reisig, Elements Of Distributed Algorithms: Modeling And Analysis With Petri Nets, Springer-verlag, First Edition, 1998.
3. Tel Gerard , Introduction To Distributed Algorithms, Second Edition, Cambridge University Press,2000
4. Valmir C. Barbosa, An Introduction To Distributed Algorithms, Mit Press,1996
5. Randy Chow, Theodore Johnson, Distributed Operating Systems and Algorithm Analysis, Pearson Education,2009.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS010	CYBER FORENSICS AND INFORMATION SECURITY	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: This course provides an exposure to the concepts and techniques in Computer Forensics Technologies. Different types of Computer Forensics systems, evidence of data gathered, network attacks, evidence data analysis and forensics tools are discussed in this course. This course helps the learners to design, and develop innovative ideas on different digital forensic investigation models

Course Outcomes: After the completion of the course the student will be able to

CO 1	Apply forensics technologies, data recovery , evidence collection and handling different forensics issues. **(Cognitive Knowledge Level: Apply)**
CO 2	Solve the issues of Data Acquisition and Data Recovery. **(Cognitive Knowledge Level: Apply)**
CO 3	Investigate network intrusions and attacks. **(Cognitive Knowledge Level: Analyze)**
CO 4	Validating Forensics data and process crime, incident scene. **(Cognitive Knowledge Level: Analyze)**
CO 5	Exploring file structures and perform forensics investigation. **(Cognitive Knowledge Level: Analyze)**
CO6	Design, Develop, Implement and Present innovative ideas on different digital forensic investigation model. **(Cognitive Knowledge Level: Create)**

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the- art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration consistency ,social, economical,ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓			✓	✓	✓	
CO 2	✓			✓	✓		
CO 3	✓			✓	✓	✓	
CO 4	✓		✓	✓	✓		
CO 5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. An email threat has been sent to the President at 4 am. The police succeeded in tracking the person down and they found that the IP address belonged to Bob. They reached his home around 8 am and seized the computer (computer was ON) and arrested Bob. Suppose you are the Cyber Forensic expert who reached the venue at 9.am and started the investigation.

   a. Identify the steps need to be taken and justify your answer by giving the reason for each step.
   b. List the methods used by the police to track Bob down?

2. How is data seizure relevant in computer forensics and evidence collection?

Course Outcome 2 (CO2):

1. After the Covid-19 virus lockdown, the American based laboratory found that someone had tampered with their central server and they think that one of the staff members did this using the company system. The Manager sends the computer to a CFS team for a forensic examination. How did the CFST go about conducting their examination?

Course Outcome 3(CO3):

1. A bank suspected an employee of downloading sensitive files from the bank's computer network using his bank laptop computer from home while on leave of absence. The bank sent the computer for a computer forensic examination. How can this issue be addressed by the professionals?
2. A cyber-attack happened in a Govt. Official server and the network has been breached. In this case how will the network forensics and investigation of the logs happen?

Course Outcome 4 (CO4):

1. Analyse the computer image verification and authentication.

2. When a network breach happens, what all are the obstacles in collecting digital evidence from the network?

Course Outcome 5 (CO5):

1. A new trainee joined a cyber-security firm. His trainer gave a talk about how to poison DNS. So what may be the techniques he has explained to him?

2. Explain the most commonly used cyber forensic tools for forensic investigation. Also write on the surveillance tools used for information warfare of the future.

Course Outcome 6 (CO6):

1. Develop a digital forensic investigation model with the idea of Venter that digital forensics investigation can be conducted by even non-technical persons

Model Question Paper

QP CODE:

Reg No: _______________

Name: _________________ **PAGES : 2**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SECOND SEMESTER MTECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: 242ECS010

Course Name: CYBER FORENSICS AND INFORMATION SECURITY

Max. Marks : 60 **Duration: 2.5 Hours**

PART A

Answer All Questions. Each Question Carries 5 Marks

1.		Analyse the procedure for corporate high-tech investigations.	**(5)**
2.		Give specific instances to illustrate data seizure.	**(5)**
3.		Explain the steps involved in investigating router attacks.	**(5)**
4.		Illustrate the validation of forensics data.	**(5)**
5.		Explain the different hardware tools for computer forensics.	**(5)**

Part B

(Answer any five questions. Each question carries 7 marks)

6.	**a**	Analyse the procedure for corporate high-tech investigations.	**(3)**
	b	How is data seizure relevant in computer forensics and evidence collection.	**(4)**
7		Describe the different cyber forensic malicious software.	**(7)**
8	**a**	Discuss Duplication and Preservation of Digital Evidence.	**(3)**
	b	A bank suspected an employee of downloading sensitive files from the bank's computer network using his bank laptop computer from home while on leave of absence. The bank sent the computer for a computer forensic examination. How can this issue be addressed by the professionals? .	**(4)**
9		A new trainee joined a cyber-security firm. His trainer gave a talk about how to poison DNS. So what may be the techniques he has explained to him?	**(7)**
10		A cyber-attack happened in a Govt. Official server and the network has been breached. In this case how will the network forensics and investigation of the logs happen?	**(7)**
11		Explain cyber surveillance and criminal tracking.	**(7)**

12	**a**	Discuss about any two software tools used for cyber forensics.	
	b	Develop a digital forensic investigation model with the idea of Venter that digital forensics investigation can be conducted by even non-technical persons.	**(4)**

Syllabus

Module 1: Introduction to Computer Forensics

Computer Forensics Fundamentals, Types of Computer Forensics Technology – Types of Computer Forensics Systems-Data Recovery and Evidence Collection– Forensic duplication and preservation of DE, Understanding Computer Investigation.

Module 2: Evidence Data Gathering

Data Acquisition. - Data Recovery- Evidence Collection and Data Seizure - Duplication and Preservation of Digital Evidence.

Module 3: Investigations

Network Traffic , Web Attacks, Router Forensics, DoS Attacks and Internet Crime.

Module 4: Evidence Data Analysis

Discovery of Electronic Evidence - Identification of Data - Determining and Validating Forensics Data – Data Hiding Techniques – Performing Remote Acquisition– Cell Phone and Mobile Devices Forensics-Processing Crime and Incident Scenes.

Module 5: Forensics Tools and Case Studies

Working with Windows and DOS Systems. - Understanding File systems, Exploring Microsoft file structures, Examining NTFS disks, Understanding whole disk encryption, windows registry, Microsoft start-up tasks.

Current Computer Forensics Tools: Software/ Hardware Tools. Computer forensics investigation – A case study

Course Plan

No	Topic	No. of Lectures (40 Hours)
1	**Module 1: Introduction to Computer Forensics**	**8**
1.1	Computer Forensics Fundamentals	1
1.2	Types of Computer Forensics Technology	1
1.3	Types of Military Computer Forensic Technology	1
1.4	Types of Computer Forensics Systems	1
1.5	Intrusion Detection Systems	1

1.6	Data Recovery and Evidence Collection	1
1.7	Forensic duplication and preservation of digital evidence	1
1.8	Understanding Computer Investigation	1
2	**Module 2: Evidence Data Gathering**	**11**
2.1	Data Acquisition	1
2.2	Data Recovery	1
2.3	Evidence Collection	1
2.4	Why Collect Evidence, Collection Options Obstacles	1
2.5	Types of Evidence, The Rules of Evidence	1
2.6	Volatile Evidence, General Procedure	1
2.7	Collection and Archiving	1
2.8	Methods of Collection, Artifacts	1
2.9	Collection Steps	1
2.10	Data Seizure	1
2.11	Duplication of Digital Evidence	1
3	**Module 3: Investigations**	**6**
3.1	Network Traffic :Investigating Network Intrusions	1
3.2	Network Forensics and Investigating logs	1
3.3	Web Attacks	1
3.4	Router Forensics	1
3.5	DoS Attacks	1
3.6	Internet Crime	1
4	**Module 4: Evidence Data Analysis**	**7**
4.1	Discovery of Electronic Evidence	1
4.2	Identification of Data	1
4.3	Determining and Validating Forensics Data	1
4.4	Data Hiding Techniques	1
4.5	Performing Remote Acquisition	1
4.6	Cell Phone and Mobile Devices Forensics	1
4.7	Processing Crime and Incident Scenes	1
5	**Module 5: Forensics Tools and Case Studies**	**8**
5.1	Working with Windows and DOS Systems	1
5.2	Understanding File systems	1
5.3	Exploring Microsoft file structures	1
5.4	Examining NTFS disks	1
5.5	Understanding whole disk encryption, windows registry	1
5.6	Microsoft start-up tasks.	1
5.7	Current Computer Forensics Software/Hardware Tools	1
5.8	Case Study	1

References

1. Man Young Rhee, “Internet Security: Cryptographic Principles”, “Algorithms and protocols”, Wiley Publications, 2003.
2. Nelson, Phillips, Enfinger, Steuart, “Computer Forensics and Investigations”, Cengage Learning, India Edition, 2008.
3. John R.Vacca, “Computer Forensics”, Cengage Learning, 2005.
4. Richard E.Smith, “Internet Cryptography”, 3rd Edition Pearson Education, 2008.
5. Marjie T.Britz, “Computer Forensics and Cyber Crime”: An Introduction”, 3rd Edition, Prentice Hall, 2013.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS011	SOFTWARE TESTING	PROGRAM ELECTIVE 4	3	0	0	3

Preamble: The course aims at introducing various concepts of software testing, including white box, black box, static and dynamic testing. Testing is an integral part of software development and the software has to be tested in all aspects to ensure trouble free execution in all environments. The learner will be able to use the various testing tools such as JUnit, Selenium, AFL and CBMC.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Explain various concepts in software testing including black box testing and white box testing. (**Cognitive Knowledge Level: Understand**)
CO 2	Develop test cases in JUnit to verify the behaviour of independent components of a program.(**Cognitive Knowledge Level: Evaluate**)
CO 3	Use Selenium to automate tests across web browsers to verify that the application behaves as expected.(**Cognitive Knowledge Level: Evaluate**)
CO 4	Use American Fuzzy Lop (AFL) to efficiently increase the coverage of test cases and to automatically discover clean, interesting test cases that trigger new internal states in the targeted binary. (**Cognitive Knowledge Level: Analyze**)
CO5	Perform bounded model checking for C programs using the tool CBMC. (**Cognitive Knowledge Level: Evaluate**)
CO 6	Design, develop and implement solutions based on the concepts of software testing. (**Cognitive Knowledge Level: Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓		
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

i. Course based task / Seminar/ Data collection and interpretation : 15 marks

ii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment

Questions Course Outcome 1

(CO1):

1. Differentiate between black box testing and white box testing, static testing and dynamic testing.

Course Outcome 2(CO2):

1. Write a Data Driven Test Class for the following class in Java and explain?

```
publicclassCalculate {

    publicintsum(intvar1, intvar2) { System.out.println("Adding
```

values: "+ var1 + " + "+ var2); **return**var1 + var2;}

Course Outcome 3 (CO3):

1. Demonstrate the use of Selenium WebDriver for the following use case. Launch Chrome browser.
Open KTU Login Page Click on the Login Button Close the web browser.

Course Outcome 4 (CO4):

1. This problem is a pen-and-paper problem, and is based on the AFL fuzzerpseudo-code. For simplicity assume that the program being tested has only one branch pair. Hence, the arrays **G** and **Shm** have only one entry (consisting of a single byte). Say at some point of the run of afl-fuzz the value in **G** is **0xEB**.
   a. Say a run of the program on a test input t1 occurs now, and the exact visit count of the branch-pair in this run (i.e., before rounding) is 16. Would this test input t1 be added to the queue Q or not? Justify your answer by showing the necessary calculations. If it would be added to Q, what would be the updated entry in G? (You can ignore the array R for this problem.)
   b. Consider the next test input t2. What is the smallest (exact) visitcount of the branch pair by t2 that would cause t2 to be not added to Q, and what is the smallest (exact) visit count of the branch pair by t2 that would cause t2 to be added to Q? Justify your answer.

Course Outcome 5 (CO5):

1. A buffer is a contiguously allocated chunk of memory, represented by an array or a pointer in C. Programs written in C do not provide automatic bounds checking on the buffer, which means a program can – accidentally or deliberately – write beyond a buffer. The following example is a perfectly valid C program (in the sense that a compiler compiles it without any errors):

```
intmain()
  {
        int buffer[10]; buffer[20]
        = 10;
  }
```

However, the write access to an address outside the allocated memory region can lead to unexpected behavior. In particular, such bugs can be exploited to overwrite the return address of a function, thus enabling the execution of arbitrary user-induced code.

Detect the above problem using CBMC.

Course Outcome 6 (CO6):

1. Use Selenium to test the web portal of KTU and report errors and bugs.

Second Semester M. Tech. Degree Examination, Month, Year

242ECS011 – Software Testing

Time: 2.5 Hours Max. Marks:

60 Part A

(Answer all questions. Each question carries 5 marks) 5x5 = 25 Marks

1. Give the generic checklist for code review.
2. Explain the usage of @DisplayName in Junit5 with an example.
3. Explain shared UI maps in Selenium.
4. Explain the concept of coverage guided fuzzing.
5. Explain bounded model checking.

Part B

(Answer any 5 questions. Each question carries 7 marks) 5x7 = 35 Marks

6. Explain the various tests conducted in dynamic black box testing.
7. Illustrate dynamic white box testing.
8. Explain the Junit Life Cycle API.
9. Explain briefly about Junit test class and write a Junit test class for the following class in Java.

```
publicclassCalculate {

    publicintsum(intvar1, intvar2) { System.out.println("Adding
      values: "+ var1 + " + "+ var2); returnvar1 + var2;
    }
}
```

10. What is Jenkins? Explain how continuous integration is achieved using Jenkins and Selenium.
11. Give the AFL fuzzing algorithm.
12. Explain how bounded model checking can be used for program debugging and repair.

Module 1: Testing Fundamentals

Introduction, black box testing, white box testing, static black box testing, dynamic black box testing, static white box testing, dynamic white box testing.

Module 2: Introduction to JUnit

Introduction to Unit Testing in Java, Test Class, Test Method, Assertions, JUnit Life Cycle API, Test Execution, JUnit Test Framework.

Module 3: Introduction to Selenium

Introduction, Selenium methods, Verification Point in Selenium, Shared UI Maps, Using functions, Using a configuration file, Data Driven Testing, UI Objects, Debugging, Exception Handling, Reporting, Batch Execution, Continuous Integration with Jenkins.

Module 4: Coverage Guided Fuzzing Using American Fuzzy Lop (AFL)

Basic test input generation, AFL as a grey box fuzzer, Characteristics of retained test cases, vulnerability detection, Measuring the code coverage, AFL algorithm.

Module 5: Bounded Model Checking with CBMC

Principles of BMC, The CBMC tool, Applications of BMC, CBMC Hands on.

Course Plan

No	Topic	No. of Lectures (37 hrs)
1	**Module 1: Testing Fundamentals**	**7**
1.1	Getting started, Black box and white box testing, static and dynamic testing	1
1.2	Static black box testing	1
1.3	Dynamic black box testing (Lecture 1)	1
1.4	Dynamic black box testing (Lecture 2)	1
1.5	Static white box testing	1
1.6	Dynamic white box testing (Lecture 1)	1
1.7	Dynamic white box testing (Lecture 1)	1
2	**Module 2: Introduction to JUnit**	**8**
2.1	Test Driven Development, Benefits of Unit Testing	1
2.2	JUnit Introduction, Java 8 Primer (Lambdas, Stream API, Optional <T>)	1
2.3	Project Setup, Writing the first test.	1
2.4	Test Class, Test Method	1
2.5	Assertions	1
2.6	JUnit Life Cycle API, Test Execution	1
2.7	Developing an application with JUnit 5 (lecture 1)	1
2.8	Developing an application with Junit 5	1
3	**Module 3: Introduction to Selenium**	**8**
3.1	Introduction to Selenium, Installation, Using Selenium IDE, Managing User Interface Controls	1
3.2	Creating First Selenium Web Driver Script, Selenium Methods	1
3.3	Verification Point in Selenium, Shared UI Map	1
3.4	Using Functions, Using a Configuration File	1

3.5	Data Driven Testing – Parameterization, Synchronizing WebDriver Scripts	1
3.6	Handling Pop-up Dialogs and Multiple Windows, Working with Dynamic UI Objects	1
3.7	Debugging scripts, Exception Handling in WebDriver, Reporting in Selenium	1
3.8	Batch Execution, Continuous Integration with Jenkins	1
4	**Module 4: Coverage Guided Fuzzing Using American Fuzzy Lop (AFL)**	**7**
4.1	Basics of test input generation using fuzzing	1
4.2	AFL as grey box fuzzer - Detailed Demo of AFL tool	1
4.3	Characteristics of Retained test cases	1
4.4	Vulnerability detection using AFL	1
4.5	Measuring the code coverage	1
4.6	AFL Algorithm - Selection of a test input for fuzzing, Scoring the test input	1
4.7	AFL Algorithm - Fuzzing the test input, Retention/discarding the generated test input	1
5	**Module 5: Bounded Model Checking with CBMC**	**7**
5.1	Principles of BMC (Lecture 1)	1
5.2	Principles of BMC (Lecture 2)	1
5.3	The CBMC Tool	1
5.4	Applications of BMC – Program Debugging and Repair, Concurrency	1
5.5	Applications of BMC – Testing, Security	1
5.6	CBMC Hands on (Lecture 1)	1
5.7	CBMC Hands on (Lecture 2)	1

References

1. Ron Patton, Software Testing, Second Edition, Sams Publishing, Pearson Education, 2007. (For module 1)
2. Shekhar Gulati and Rahul Sharma, Java unit Testing with JUnit 5, Apress, https://doi.org/10.1007/978-1-4842-3015-2 (For module 2)
3. Navneesh Garg, Test Automation Using Selenium WebDriver with java, AdactIn Group Pty Ltd, 2014. (For module 3)
1. The AFL home page (for module 4), https://lcamtuf.coredump.cx/afl/
2. The CPROVER Manual (for module 5), http://www.cprover.org/cprover-manual/

3. Biere, Armin; Cimatti, Alessandro; Clarke, Edmund M; Strichman, Ofer; Zhu, Yunshan (2018): Bounded Model Checking. Carnegie Mellon University. Journal contribution. https://doi.org/10.1184/R1/6603944.v1 (for module 5)
4. Edmund Clarke, Daniel Kroening & Flavio Lerda. A Tool for Checking ANSI-C Programs. TACAS 2004for module 5)
5. Rohan Bavishi, Awanish Pandey, Subhajit Roy. To be precise: regression aware debugging. OOPSLA 2016 for module 5)
6. Bernd Fischer, Omar Inverso, Gennaro Parlato. CSeq: A Concurrency Pre-processor for Sequential C Verification Tools. ASE 2013 for module 5)
7. Andreas Holzer, Christian Schallhart, Michael Tautschnig & Helmut Veith. Query-Driven Program Testing. VMCAI 2009 for module 5)
8. Vladimir Klebanov, Norbert Manthey & Christian Muise. SAT-Based Analysis and Quantification of Information Flow in Programs. QEST 2013 for module 5)

SEMESTER II

INTERDISCIPLINARY ELECTIVE

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS056	INTRODUCTION TO MACHINE LEARNING	INTERDISCIPLINARY ELECTIVE	3	0	0	3

Preamble: This course helps the learners to understand the concepts in Machine Learning. Students will be able to understand the basics of regression, classification and clustering. After completing this course students will be able to develop machine learning based solution for real world problems in multidisciplinary environments.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Illustrate the concept, purpose, scope, steps, and applications of ML techniques. **(Knowledge level : Apply)**
CO 2	Understand the concepts of supervised, unsupervised and reinforcement learning to apply in real world problems. **(Knowledge level : Apply)**
CO 3	Illustrate the working of classifiers and clustering techniques for typical machine learning applications. **(Knowledge level : Apply)**
CO 4	Acquire skills to improve the performance of Machine Learning models using ensemble techniques. **(Knowledge level : Apply)**
CO5	Design and Implement solution for a real world problem using Machine Learning algorithms (Cognitive Knowledge Level: **Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓		✓	✓	✓		
CO 2	✓		✓	✓	✓		
CO 3	✓		✓	✓	✓		
CO 4	✓		✓	✓	✓		
CO5	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	80%
Analyse	20%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Continuous Internal Evaluation : 40 marks

Micro project/Course based project : 20 marks

Course based task/Seminar/Quiz : 10 marks

Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects not permitted. Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

Total : 60 marks

The end semester examination will be conducted by the respective College. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For the healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the precision and recall for the data.

2. Distinguish between supervised learning and Reinforcement learning. Illustrate with an example.

3. Discuss any four examples of machine learning applications.

Course Outcome 2 (CO2)

1. State the mathematical formulation of the SVM problem. Give an outline of the method for solving the problem.

2. Show the final result of hierarchical clustering with complete link by drawing a dendrogram.

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	0.84	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	0.93	0.20	0.67	0

Course Outcome 3(CO3):

1. Identify the first splitting attribute for the decision tree by using the ID3 algorithm with the following dataset.

Major	Experience	Tie	Hired?
CS	programming	pretty	NO
CS	programming	pretty	NO
CS	management	pretty	YES
CS	management	ugly	YES
business	programming	pretty	YES
business	programming	ugly	YES
business	management	pretty	NO
business	management	pretty	NO

2. Consider the training data in the following table where Play is a class attribute. In the table, the Humidity attribute has values "L" (for low) or "H" (for high), Sunny has values "Y" (for yes) or "N" (for no), Wind has values "S" (for strong) or "W" (for weak), and Play has values "Yes" or "No".

Humidity	Sunny	Wind	Play
L	N	S	No
H	N	W	Yes
H	Y	S	Yes
H	N	W	Yes
L	Y	S	No

What is the class label for the following day (Humidity=L, Sunny=N, Wind=W), according to naïve Bayesian classification?

3. Explain DBSCAN algorithm for density based clustering. List out its advantages compared to K-means.

4. Explain how Support Vector Machine can be used for classification of linearly separable data.

5. Define Hidden Markov Model. What is meant by the evaluation problem and how is this solved?

6. Use K Means clustering to cluster the following data into two groups. Assume cluster centroid are m1=2 and m2=4. The distance function used is Euclidean distance. { 2, 4, 10, 12, 3, 20, 30, 11, 25 }

Course Outcome 4 (CO4):

1. Explain how the *Random Forests* give output for *Classification*, and *Regression* problems?

2. Is *Random Forest* an *Ensemble Algorithm*

3. Why is the training efficiency of Random Forest better than Bagging?

Model Question Paper

Reg No: ______________			
Name: _________________			**PAGES : 4**
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY			
SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR			
Course Code: 242ECS056			
Course Name: INTRODUCTION TO MACHINE LEARNING			
Max. Marks: 60			**Duration: 2.5 Hours**
PART A			
Answer All Questions. Each Question Carries 5 Marks			
1.	Bias-Variance trade-off is a design consideration while training the machine learning model. Justify.		
2.	Derive the expression for sigmoid function associated with Logistic Regression.		
3.	How Optimal Marginal Hyperplane contributes to the accuracy of predictions using SVM. Justify how Kernel functions are used in Linear Inseparable problems		
4.	Discuss how good DBSCAN is in clustering data points available in dense Euclidean space.		
5.	Suggest an ensemble method that generates one classifier per round		**(5x5=25)**
Part B			
(Answer any five questions. Each question carries 7 marks)			
6.		Explain various Cost Functions associated with Regression & Classification.	**(7)**
7.		Weight updation contributes to the performance of a Neural Network model. Justify the statement using the Back propagation algorithm.	**(7)**
8.		Compute the Principal Components for the 2D data: X=(x1,x2)={(1,2),(3,3),(3,5),(5,4),(5,6),(6,5),(8,7),(9,8)}	**(7)**

9. Using Naïve Bayes algorithm, predict whether a Red color car which is imported as a Sports category will be stolen or not. **(7)**

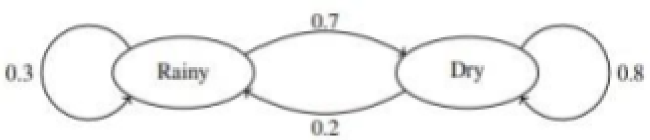
Color	Type	Origin	Stolen?
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

10 Construct Dendrograms based on Complete Linkage and Average Linkage.(7) **(7)**

ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0

11 Perform k-means algorithm on the data given in qn.8. (Given no. of clusters =2, iterations=2). **(7)**

12 (a) If P(Rain) = 0.4 and P(Dry) = 0.6 compute the probability for the sequence "Rain, Rain, Dry, Dry". **(3)**



(b) Elucidate the three basic problems of HMM **(4)**

	Syllabus	
Module	**Contents**	**Hours**
I	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning. Types of ML problems: Classification, Clustering and Regression, Cost functions: Definition and Types, Data PreProcessing, Bias-Variance trade off, Cross validation techniques, Classifier performance measures, ROC	6

	Curves	
II	Introduction to neural network : Linear Regression, Least square Gradients, Logistic Regression, Sigmoid function & differentiation, Logistic Regression – Regularization, Neural Networks – Concept of perceptron and Artificial neuron, Weight initialization techniques, Feed Forward Neural Network, Back Propagation algorithm	8
III	Classification Methods : Support Vector Machine, Optimal Separating hyper plane, Kernel trick, Kernel functions, Gaussian class conditional distribution, Bayes Rule, Naïve Bayes Model, Decision Tree – ID3, Maximum Likelihood estimation techniques	8
IV	Clustering Methods: K-means clustering , Hierarchical clustering techniques, Density Based clustering, Feature Selection techniques: Entropy, Correlation Coefficient, Chi-square Test, Forward & Backward Selection, Dimensionality Reduction: PCA, LDA, t-SNE	7
V	Basics of graphical models - Bayesian networks, Hidden Markov model, Ensemble methods – Boosting, Bagging, Random forest, XGBoost (Case study)	6

Lesson Plan

1	**Introduction to machine learning (Hours: 6)**	
1.1	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning	1
1.2	Types of ML problems: Classification, Clustering and Regression	1
1.3	Cost functions: Definition and Types	1
1.4	Data PreProcessing, Bias-Variance trade off	1
1.5	Cross validation techniques	1
1.6	Classifier performance measures, ROC Curves	1
2	**Introduction to neural network (Hours: 8)**	
2.1	Linear Regression, Least square Gradients	1
2.2	Logistic Regression	1
2.3	Sigmoid function & differentiation	1
2.4	Logistic Regression - Regularization	
2.5	Neural Networks – Concept of perceptron and Artificial neuron	1
2.6	Weight initialization techniques	1
2.7	Feed Forward Neural Network	1
2.8	Back Propagation algorithm	1
3	**Classification Methods (Hours: 8)**	
3.1	Support Vector Machine	1
3.2	Optimal Separating hyper plane	1
3.3	Kernel trick, Kernel functions	1
3.4	Gaussian class conditional distribution	1
3.5	Bayes Rule	1
3.6	Naïve Bayes Model	1

3.7	Decision Tree – ID3,	1
3.8	Maximum Likelihood estimation techniques	1
4	**Clustering Methods (Hours: 7)**	
4.1	K-means clustering	1
4.2	Hierarchical clustering techniques	1
4.3	BIRCH	1
4.4	Density Based clustering	1
4.5	Feature Selection techniques: Entropy, Correlation Coefficient, Chi-square Test	1
4.6	Forward & Backward Selection	1
4.5	Dimensionality Reduction: PCA	1
4.6	LDA	1
4.7	t-SNE	1
5	**Basics of graphical models (Hours: 6)**	
5.1	Basics of graphical models - Bayesian networks	1
5.2	Hidden Markov model	1
5.3	Ensemble methods - Boosting	1
5.4	Bagging	1
5.5	Random forest	1
5.6	XGBoost (Case study)	1

Reference Books

1. Ethem Alpaydın, "Introduction to Machine Learning (Adaptive Computation and Machine Learning)", MIT Press, 2004.

2. Kevin Murphy, Machine Learning: A Probabilistic Perspective (MLAPP), MIT Press, 2012

3. Han, Jiawei, and Micheline Kamber. Data Mining: Concepts and Techniques. San Francisco: Morgan Kaufmann Publishers

4. Christopher M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS057	DATA STRUCTURES	INTERDISCIPLINARY ELECTIVE	3	0	0	3

Preamble: The purpose of the syllabus is to create awareness about Data Structures and their applications. After the completion of the course, the learners should be able to either use existing data structures or design their own data structures to solve real world problems.

Course Outcomes: After the completion of the course the student will be able to

CO1	Design algorithms for a task and calculate the time complexity of that algorithm (**Cognitive Knowledge Level: Apply**)
CO2	Use arrays and linked lists for problem solving (Cognitive Knowledge Level: **Apply**)
CO3	Represent data using trees, graphs and manipulate them to solve computational problems. (**Cognitive Knowledge Level:Apply**)
CO4	Make use of appropriate sorting algorithms to order data based on the situation. (**Cognitive Knowledge Level: Apply**)
CO5	Design and Implement appropriate Data Structures for solving a real world problem (Cognitive Knowledge Level: **Create**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓		✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓	✓	✓	✓		✓	✓

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%
Analyse	30%
Evaluate	
Create	

Mark distribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Sample Course Level Assessment Questions

Course Outcome1(CO1):

1. Write an algorithm for matrix multiplication and calculate its time complexity.
2. Define Big-O notation. Derive the Big – O notation for $5n^3+2n^2+3^n$.
3. Check whether the following is true or not. 2n+1 is O(2n). Give reason.

Course Outcome 2(CO2):

1. How a linked list can be used to represent the polynomial
2. $5x^4y^6+24x^3y^4-17x^2y^3+15xy^2+45$.
   Write a procedure to add two Bivariate polynomials represented using linked lists.

3. Write an algorithm/pseudocode to convert a given infix expression to postfix expression. Trace the steps involved in converting the given infix expression

   ((A
   +B)^C)-((D*C)/F) to postfix expression.

4. Let L1 be a singly linked list in memory. Write an algorithm that
   i) Finds the number of non zero elements in L1
   ii) Adds a given value K to each element in L1

Course Outcome 3(CO3):

1. Create a Binary Tree with the following sequence 14, 15, 4, 18, 9, 16, 20, 17, 3, 7, 5, 2 and perform inorder, preorder and postorder traversals on the above tree and print the output of the traversals.
2. In a complete binary tree of depth d (complete including last level), give an expression to find the number of leaf nodes.

Course Outcome 4(CO4):

1. Write an algorithm/pseudocode to sort elements using Heap sort technique. Illustrate the working of Heap sort algorithm on the following input : 35,15,0,1,60

2. With the help of an algorithm/pseudocode and suitable example, explain how you would perform binary search on an array of n elements. Find the time complexity of binary search algorithm.

3. Suppose an array contains elements {10, 13, 21, 32, 35, 44, 55}. Give the steps to find an element "35" using i) linear search ii) binary search

Course Outcome 5(CO5):

Design a reservation system for railways that includes a waiting list. If the reservation is full, display "reservation full" and put them in the waiting list and give a waiting list number. If a passenger wishes to cancel his ticket, he may do it any time. Then the passenger at the front of the waiting list is allotted a berth automatically.

Model Question Paper

Reg No: _______________			
Name:			**PAGES : 4**
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY			
SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR			
Course Code: 242CS057			
Course Name: Data Structures			
Max. Marks: 60			**Duration: 2.5 Hours**
PART A			
Answer All Questions. Each Question Carries 5 Marks			
1.		Write an algorithm to add a new element in a particular position of an array.	
2.		Compare Circular Queue with a Normal Queue	
3.		How can a doubly linked list be used to find palindromes?	
4.		Write an iterative algorithm for in-order traversal of a Binary Tree	
5.		Trace the working of Quick sort on the following input 38,8,0,28,45,-12,89,66,42.	**(5x5=25)**
Part B			
(Answer any five questions. Each question carries 7 marks)			
6.	(a)	How is the performance of an algorithm evaluated?	**(4)**
	(b)	In the functions O(nlogn) and O(logn), which one is better in terms of computational complexity and why?	**(3)**
7.	(a)	Write algorithms to insert and delete elements from a double ended queue. Illustrate with examples	**(7)**
8.	(a)	Write an algorithm to multiply two polynomials represented using linked list	**(7)**
9.	(a)	List the properties of Binary Search Tree. Write an algorithm to search an element in a Binary Search Tree	**(7)**
10		Give algorithms for DFS and BFS of a graph and explain with examples.	**(7)**
11	(a)	Write and illustrate algorithms for Merge Sort and Quick Sort.	**(7)**
12	(a)	How is memory de-allocation done in memory management	**(7)**

	Syllabus	
Module	**Contents**	**Hours**
I	**Basic Concepts of Data Structures** System Life Cycle, Algorithms, Performance Analysis, Space complexity, Time Complexity, Asymptotic Notation, Complexity Calculation of Simple Algorithms	5
II	**Arrays and Searching** Polynomial representation using Arrays, Sparse matrix, Stacks, Queues, Circular Queues, Priority Queues, Double Ended Queues, Evaluation of Expressions, Linear Search and Binary Search	9
III	**Linked List and Memory Management** Self Referential Structures, Dynamic Memory Allocation, Singly Linked List-Operations on Linked List, Doubly Linked List, Circular Linked List, Stacks and Queues using Linked List, Polynomial representation using Linked List, Memory allocation and de-allocation,	8
IV	Trees, Binary Trees-Tree Operations, Binary Tree Representation, Tree Traversals, Graphs- Representation of Graphs, Depth First Search and Breadth First Search on Graphs, Applications of Graphs	7
V	Searching Techniques – Linear search , Binary search Sorting Techniques – Selection Sort, Insertion Sort, Quick Sort, Merge Sort and Heap Sort	5

Course Plan

Module 1 :Basic Concepts of Data Structures		(5 hours)
1.1	System Life Cycle,	1 hour
1.2	Algorithms	1 hour
1.3	Performance Analysis, Space Complexity, Time Complexity,	1 hours
1.4	Asymptotic Notation	1hour
1.5	Complexity Calculation of Simple Algorithms	1hour
Module 2 :Arrays and Searching		(9 hours)
2.1	Polynomial representation using Arrays	1 hour
2.2	Sparse matrix	1 hours

2.3	Stacks	1 hour
2.4	Queues, Circular Queues	1 hour
2.5	Priority Queues,	1 hour
2.6	Double Ended Queue	1 hours
2.7	Evaluation of Expressions	1 hour
2.8	Linear Search	1 hour
2.9	Binary Search	1 hour
Module 3 : Linked List and Memory Management		(8 hours)
3.1	Self Referential Structures	1 hour
3.2	Dynamic Memory Allocation	1 hour
3.3	Single Linked List-Operations on Linked List,	1 hour
3.4	Double Linked List	1 hour
3.5	Circular Linked List	1 hour
3.6	Stacks and Queues using Linked List	1 hour
3.7	Polynomial representation using Linked List	1 hour
3.8	Memory de-allocation	1 hour
Module 4 :Trees and Graphs		(7 hours)
4.1	Trees, Binary Trees	1hour
4.2	Tree Operations, Binary Tree Representation,	1hour
4.3	Tree Traversals	1hour
4.4	Graphs	1hour
4.5	Representation of Graphs	1hour
4.6	Depth First Search and Breadth First Search on Graphs	1hour
4.7	Applications of Graphs	1hour
Module 5 : Sorting and Hashing		(5 hours)
5.1	Sorting Techniques – Selection Sort	1hour
5.2	Insertion Sort	1hour
5.3	Quick Sort	1hour
5.4	Merge Sort	1hour
5.5	Heap Sort	1hour

Text Book

1. Ellis Horowitz, Sartaj Sahni and Susan Anderson-Freed, Universities Press, Fundamentals of Data Structures in C

Reference Books

1. Samanta D., Classic Data Structures, Prentice Hall India, 2/e, 2009.
2. Richard F. Gilberg, Behrouz A. Forouzan, Data Structures: A Pseudocode Approach with C, 2/e, Cengage Learning, 2005
3. Aho A. V., J. E. Hopcroft and J. D. Ullman, Data Structures and Algorithms, Pearson Publication,1983.
4. Tremblay J. P. and P. G. Sorenson, Introduction to Data Structures with Applications, Tata McGraw Hill, 1995.
5. Peter Brass, Advanced Data Structures, Cambridge University Press, 2008
6. Lipschuts S., Theory and Problems of Data Structures, Schaum's Series, 1986.
7. Wirth N., Algorithms + Data Structures = Programs, Prentice Hall, 2004.
8. Hugges J. K. and J. I. Michtm, A Structured Approach to Programming, PHI, 1987.
9. Martin Barrett, Clifford Wagner, And Unix: Tools For Software Design, John Wiley, 2008 reprint

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
242ECS058	**SOFTWARE PROJECT MANAGEMENT**	**INTERDISCIPLINARY ELECTIVE**	**3**	**0**	**0**	**3**

Preamble: This course provides fundamental knowledge in the Software Development Process. It covers Software Development, Quality Assurance, Project Management concepts and technology trends. This course enables the learners to apply state of the art industry practices in Software development.

Course Outcomes: After the completion of the course the student will be able to

CO1	Demonstrate Traditional and Agile Software Development approaches **(Cognitive Knowledge Level: Apply)**
CO2	Prepare Software Requirement Specification and Software Design for a given problem. **(Cognitive Knowledge Level**: **Apply**)
CO3	Justify the significance of design patterns and licensing terms in software development, prepare testing, maintenance and DevOps strategies for a project. **(Cognitive Knowledge Level: Apply**)
CO4	Make use of software project management concepts while planning, estimation, scheduling, tracking and change management of a project, with a traditional/agile framework**. (Cognitive Knowledge Level: Apply**)
CO5	Utilize SQA practices, Process Improvement techniques and Technology advancements in cloud based software models and containers & microservices. **(Cognitive Knowledge Level: Apply**)

Program Outcomes (PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the- art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO 1	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**	**PO 7**
CO 1	✓		✓	✓		✓	
CO 2	✓		✓	✓		✓	
CO 3	✓		✓	✓		✓	
CO 4	✓		✓	✓		✓	
CO 5	✓		✓	✓		✓	

Assessment Pattern

Bloom's Category	**End Semester Examination**
Apply	70%
Analyse	30%
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Course based task / Seminar/ Data collection and interpretation : 15 marks

iii. Test paper (1 number) : 10 marks

Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. What are the advantages of an incremental development model over a waterfall model?
2. Illustrate how the process differs in agile software development and traditional software development with a socially relevant case study. (Assignment question)

Course Outcome 2 (CO2):

1. How to prepare a software requirement specification?
2. Differentiate between Architectural design and Component level design.
3. How does agile approaches help software developers to capture and define the user requirements effectively?
4. What is the relevance of the SRS specification in software development?
5. Prepare a use case diagram for a library management system.

Course Outcome 3 (CO3):

1. Differentiate between the different types of software testing strategies.
2. Justify the need for DevOps practices?
3. How do design patterns help software architects communicate the design of a complex system effectively?
4. What are the proactive approaches one can take to optimise efforts in the testing phase?

Course Outcome 4 (CO4):

1. Illustrate the activities involved in software project management for a socially relevant problem?
2. How do SCRUM, Kanban and Lean methodologies help software project management?
3. Is rolling level planning in software project management beneficial? Justify your answer.
4. How would you assess the risks in your software development project? Explain how you can manage identified risks?

Course Outcome 5 (CO5):

1. Justify the importance of Software Process improvement?
2. Explain the benefits of cloud based software development, containers and microservices.
3. Give the role of retrospectives in improving the software development process.
4. Illustrate the use of project history data as a prediction tool to plan future socially relevant projects.

Model Question Paper

Reg No: _______________			
Name:			**PAGES : 4**
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY			
SECOND SEMESTER M.TECH DEGREE EXAMINATION, MONTH & YEAR			
Course Code: 242CS058			
Course Name: SOFTWARE PROJECT MANAGEMENT			
Max. Marks: 60			**Duration: 2.5 Hours**
PART A			
Answer All Questions. Each Question Carries 5 Marks			
1.	Explain Agile ceremonies and Agile manifesto.		
2.	Compare Software Architecture design and Component level design		
3.	Describe the formal and informal review techniques		
4.	Explain plan driven development and project scheduling.		
5.	Illustrate SPI process with an example.		**(5x5=25)**
Part B **(Answer any five questions. Each question carries 7 marks)**			
6.		Illustrate software process activities with an example.	**(7)**
7.		What are functional and nonfunctional requirements? Imagine that you are developing a library management software for your college, list eight functional requirements and four nonfunctional requirements.	**(7)**

8.		Explain Continuous Integration, Delivery, and Deployment CI/CD/CD)	**(7)**
9.		What is a critical path and demonstrate its significance in a project schedule with the help of a sample project schedule.	**(7)**
10.		What is algorithmic cost modeling? What problems does it suffer from when compared with other approaches to cost estimation?	**(7)**
11.		Explain elements of Software Quality Assurance and SQA Tasks.	**(7)**
12.		Compare CMMI and ISO 9001:2000.	**(7)**

Syllabus

Module 1 : Introduction to Software Engineering (7 hours)

Introduction to Software Engineering - Professional software development,

Software engineering

ethics. Software process models - The waterfall model, Incremental development. Process activities - Software specification, Software design and implementation, Software validation, Software evolution. Coping with change - Prototyping, Incremental delivery, Boehm's Spiral Model. Agile software development - Agile methods, agile manifesto - values and principles. Agile development techniques, Agile Project Management. Case studies : An insulin pump control system. Mentcare - a patient information system for mental health care.

Module 2 : Requirement Analysis and Design (8 hours)

Functional and non-functional requirements, Requirements engineering processes. Requirements elicitation, Requirements validation, Requirements change, Traceability Matrix. Developing use cases, Software Requirements Specification Template, Personas, Scenarios, User stories, Feature identification. Design concepts - Design within the context of software engineering, Design Process, Design concepts, Design Model. Architectural Design - Software Architecture, Architectural Styles, Architectural considerations, Architectural Design Component level design

- What is a component?, Designing Class-Based Components, Conducting Component level design, Component level design for web-apps. Template of a Design Document as per "IEEE Std 1016-2009 IEEE Standard for Information Technology Systems Design Software Design Descriptions". Case study: The Ariane 5 launcher failure.

Module 3 : Implementation and Testing (9 hours)

Object-oriented design using the UML, Design patterns, Implementation issues, Open-source development - Open-source licensing - GPL, LGPL, BSD. Review Techniques - Cost impact of Software Defects, Code review and statistical analysis. Informal Review, Formal Technical Reviews, Post-mortem evaluations. Software testing strategies - Unit Testing, Integration Testing, Validation testing, System testing, Debugging, White box testing, Path testing, Control Structure testing, Black box testing, Testing Documentation and Help facilities. Test automation, Test-driven development, Security testing. Overview of DevOps and Code Management - Code management, DevOps automation, Continuous Integration, Delivery, and Deployment (CI/CD/CD). Software Evolution - Evolution processes, Software maintenance.

Module 4 : Software Project Management (6 hours)

Software Project Management - Risk management, Managing people, Teamwork. Project Planning, Software pricing, Plan-driven development, Project scheduling, Agile planning. Estimation techniques, COCOMO cost modeling. Configuration management, Version management, System building, Change management, Release management, Agile software management - SCRUM framework. Kanban methodology and lean approaches.

Module 5 : Software Quality, Process Improvement and Technology trends (6 hours)

Software Quality, Software Quality Dilemma, Achieving Software Quality Elements of Software Quality Assurance, SQA Tasks , Software measurement and metrics. Software Process Improvement(SPI), SPI Process CMMI process improvement framework, ISO 9001:2000 for Software. Cloud-based Software - Virtualisation and containers, Everything as a service(IaaS, PaaS), Software as a service. Microservices Architecture - Microservices, Microservices architecture, Microservice deployment.

Text Books

1. Book 1 - Ian Sommerville, Software Engineering, Pearson Education, Tenth edition, 2015.
2. Book 2 - Roger S. Pressman, Software Engineering : A practitioner’s approach, McGraw Hill publication, Eighth edition, 2014
3. Book 3 - Ian Sommerville, Engineering Software Products: An Introduction to Modern Software Engineering, Pearson Education, First Edition, 2020.

References

2. IEEE Std 830-1998 - IEEE Recommended Practice for Software Requirements SpeciÞcations
3. IEEE Std 1016-2009 IEEE Standard for Information Technology—Systems Design— Software Design Descriptions
4. David J. Anderson, Kanban, Blue Hole Press 2010

5. David J. Anderson, Agile Management for Software Engineering, Pearson, 2003
6. Walker Royce, Software Project Management : A unified framework, Pearson Education, 1998
7. Steve. Denning, The age of agile, how smart companies are transforming the way work gets done. New York, Amacom, 2018.
8. Satya Nadella, Hit Refresh: The Quest to Rediscover Microsoft's Soul and Imagine a Better Future for Everyone, Harper Business, 2017
9. Henrico Dolfing, Project Failure Case Studies: Lessons learned from other people's mistakes, Kindle edition
10. Mary Poppendieck, Implementing Lean Software Development: From Concept to Cash, Addison-Wesley Signature Series, 2006
11. StarUML documentation - https://docs.staruml.io/
12. OpenProject documentation - https://docs.openproject.org/
13. BugZilla documentation - https://www.bugzilla.org/docs/
14. GitHub documentation - https://guides.github.com/
15. Jira documentation - https://www.atlassian.com/software/jira

Teaching Plan

No	Contents	No of Lecture Hrs
Module 1 : Introduction to Software Engineering (7 hours)		
1.1	Introduction to Software Engineering.[Book 1, Chapter 1]	1 hour
1.2	Software process models [Book 1 - Chapter 2]	1 hour
1.3	Process activities [Book 1 - Chapter 2]	1 hour
1.4	Coping with change [Book 1 - Chapter 2, Book 2 - Chapter 4]	1 hour
1.5	Case studies : An insulin pump control system. Mentcare - a patient information system for mental health care. [Book 1 - Chapter 1]	1 hour
1.6	Agile software development [Book 1 - Chapter 3]	1 hour
1.7	Agile development techniques, Agile Project Management.[Book 1 - Chapter 3]	1 hour
Module 2 : Requirement Analysis and Design (8 hours)		
2.1	Functional and non-functional requirements, Requirements engineering processes [Book 1 - Chapter 4]	1 hour

2.2	Requirements elicitation, Requirements validation, Requirements change, Traceability Matrix [Book 1 - Chapter 4]	1 hour
2.3	Developing use cases, Software Requirements Specification Template [Book 2 - Chapter 8]	1 hour
2.4	Personas, Scenarios, User stories, Feature identification [Book 3 - Chapter 3]	1 hour
2.5	Design concepts [Book 2 - Chapter 12]	1 hour
2.6	Architectural Design [Book 2 - Chapter 13]	1 hour
2.7	Component level design [Book 2 - Chapter 14]	1 hour
2.8	Design Document Template. Case study: The Ariane 5 launcher failure. [Ref - 2, Book 2 - Chapter 16]	1 hour
Module 3 : Implementation and Testing (9 hours)		
3.1	Object-oriented design using the UML, Design patterns [Book 1 - Chapter 7]	1 hour
3.2	Implementation issues, Open-source development - Open-source licensing - GPL, LGPL, BSD [Book 1 - Chapter 7]	1 hour
3.3	Review Techniques - Cost impact of Software Defects, Code review and statistical analysis. [Book 2 - Chapter 20]	1 hour
34	Informal Review, Formal Technical Reviews, Post-mortem evaluations. [Book 2 - Chapter 20]	1 hour
3.5	Software testing strategies - Unit Testing, Integration Testing, Validation testing, System testing and Debugging (basic concepts only). [Book 2 - Chapter 22]	1 hour
3.6	White box testing, Path testing, Control Structure testing, Black box testing. Test documentation [Book 2 - Chapter 23]	1 hour
3.7	Test automation, Test-driven development, Security testing. [Book 3 - Chapter 9]	1 hour
3.8	DevOps and Code Management - Code management, DevOps automation, CI/CD/CD. [Book 3 - Chapter 10]	1 hour
3.9	Software Evolution - Evolution processes, Software maintenance. [Book 1 - Chapter 9]	1 hour
Module 4 : Software Project Management (6 hours)		

4.1	Software Project Management - Risk management, Managing people, Teamwork [Book 1 - Chapter 22]	1 hour
4.2	Project Planning - Software pricing, Plan-driven development, Project scheduling, Agile planning [Book 1 - Chapter 23]	1 hour
4.3	Estimation techniques [Book 1 - Chapter 23]	1 hour
4.4	Configuration management [Book 1 - Chapter 25]	1 hour
4.5	Agile software management - SCRUM framework [Book 2 - Chapter 5]	1 hour
4.6	Kanban methodology and lean approaches.[Ref 9 - Chapter 2]	1 hour
Module 5 : Software Quality, Process Improvement and Technology trends (6 hours)		
5.1	Software Quality, Software Quality Dilemma, Achieving Software Quality. [Book 2 - Chapter 19]	1 hour
5.2	Elements of Software Quality Assurance, SQA Tasks , Software measurement and metrics. [Book 3 - Chapter 21]	1 hour
5.3	Software Process Improvement (SPI), SPI Process [Book 2 - Chapter 37]	1 hour
5.4	CMMI process improvement framework, ISO 9001:2000 for Software. [Book 2 - Chapter 37]	1 hour
5.5	Cloud-based Software - Virtualisation and containers, IaaS, PaaS, SaaS.[Book 3 - Chapter 5]	1 hour
5.6	Microservices Architecture - Microservices, Microservices architecture, Microservice deployment [Book 3 - Chapter 6]	1 hour

SEMESTER-III
SYLLABUS

CODE		CATEGORY	L	T	P	CREDIT
243AGE100	ACADEMIC WRITING	AUDIT COURSE	3	0	0	NIL

Preamble: Learning academic writing sharpens minds, teaches students how to communicate, and develops their thinking capacities and ability to understand others. Writing is thinking, and every student deserves to be a strong thinker. It can also make them think more carefully about what they write. Showing work to others can help to foster a better culture of learning and sharing among students. It also gives students a sense of how they are contributing to the body of work that makes up an academic subject.

Course Outcomes: The COs shown are only indicative. For each course, there can be 4 to 6 COs.

After the completion of the course the student will be able to

CO 1	Understand the principles of scientific/ academic writing
CO 2	Analyse the technique of scientific writing from the reader's perspective
CO 3	Apply the concepts of setting expectations and laying the progression tracks
CO 4	Evaluate the merits of a title, abstract , introduction, conclusion and structuring of a research paper
CO 5	Justify the need using a project proposal or a technical report
CO 6	Prepare a review paper, an extended abstract and a project proposal

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1		3	1				
CO 2		3	1				
CO 3		3	1			2	
CO 4		3	1				
CO 5		3	2	2		2	
CO 6	1	3	3	2		2	

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	40%
Analyse	30%
Evaluate	30%

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks

Course based task : 15 marks
Seminar/Quiz : 15 marks Test
paper, 1 no. : 10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern: 60 marks

The examination will be conducted by the respective College. The examination willbe for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any Vive. Each question can carry 12 marks.

Model Question paper

		SET1	**Total Pages:**
Reg No.:_______________		Name:___________________________	

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
THIRD SEMESTER M.TECH DEGREE EXAMINATION, MARCH 2024

Course Code: 243AGE100

Course Name: Academic Writing

Max. Marks: 60 | Duration: 2.5 Hours

Answer any five full questions, each carries 12 marks.

1 a)	Make clear-cut distinctions between 6 factors that take their toll on readers' memory.	6
1 b)	How can you sustain the attention of the reader to ensure continuous reading?	6
2 a)	What are the different methods by which you can create expectations in the reader?	6
2 b)	Give an account of the topic and non-topic based progression schemes.	6
3 a)	Bring out the differences between an abstract and the introduction of a research paper.	8
3 b)	How are the title of the research paper and its structure related?	4
4	What are 7 principles for including visuals in your research paper. What are the recommended constituents of a conclusion segment of a research paper?	12
5	Give a detailed description of the process and contents of a project proposal for funding.	12
6 a)	What are the contexts recommended for choosing between active and passive voices in technical writing?	8
6 b)	What are the different visual forms that are relevant in a research paper and how do you choose them?	4
7	Give the design of a research paper with the purposes each part serves.	12

Syllabus and Course Plan (For 3 credit courses, the content can be for 40 hrs and for 2 credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30 hours).

Syllabus:

CODE 243AG E100	**ACADEMIC WRITING**	**Audit**
Module No.	Topics in a module	Hours
1	Fundamentals of Academic writing from a reader's perspective: acronyms, synonyms, pronouns, disconnected phrases, background ghettos, abusive detailing, cryptic captions, long sentences : all that take their toll on readers' memory.	6
2	Fluid reading & reading energy consumption: setting expectations and laying Progression tracks; Reading energy consumption	6
3	How to write the Title, abstract, introduction ; Structure the writing with headings & subheadings	6
4	Visuals: Resources, Skills, and Methods; Conclusion; References; Bibliography; Grammar in technical writing	6
5	Techniques of writing: An extended abstract, a project proposal, a research paper, a technical report.	6

Course Plan:

No	Topic	No. of Lectures
1	Fundamentals of Academic writing from a reader's perspective: acronyms, synonyms, pronouns, disconnected phrases, background ghettos, abusive detailing, cryptic captions, long sentences all take their toll on readers' memory.	
1.1	The Reading tool-kit to reduce memory required; reduce reading time	1
1.2	Acronyms, Pronouns, Synonyms; Background, broken couple, words overflow	1
1.3	Sustain attention: Keep the story moving forward; Twists, shouts, Pause to clarify, recreate suspense	2
1.4	Keep the reader motivated: Fuel and meet Expectations; Bridge knowledge gap: ground level; Title words; Just In Time to local background	2
2	Fluid reading & reading energy consumption: setting expectations and laying Progression tracks; Reading energy consumption	
2.1	Setting expectations of the reader from Grammar, from theme	1
2.2	Progression tracks for fluid reading: Topic & stress; topic and non topic based progression tracks; pause in progression	2

2.3	Detection of sentence fluidity problems: No expectations/ Betrayed expectations	2
2.4	Controlling reading energy consumption: the energy bill; Energy fuelling stations: Pause	1
3	How to write the Title, abstract, introduction ; Structure the writing with headings & subheadings	
3.1	Title: Face of the paper: Techniques, Qualities & Purpose of title; Metrics	1
3.2	Abstract: Heart of the paper: 4 parts; coherence; tense of verbs, precision; purpose & qualities of the abstract; Metrics	2
3.3	Structure: Headings & sub-headings: Skeleton of the paper: principles for a good structure; Syntactic rules; Quality & Purpose of structures; Metrics	1
3.4	Introduction: Hands of the paper: Start, finish; scope, definitions; answers key reader questions; As a personal active story; Traps, qualities; Metrics	2
4	Visuals: Resources, Skills, and Methods; Conclusion; References; Bibliography; Grammar in technical writing	
4.1	Visuals as the voice of your paper: principles; purpose & qualities of visuals; metrics	2
4.2	Conclusion: contents; purpose, quality; metrics; Abstracts Vs. Conclusion; examples, counter-examples	1
4.3	References, Bibliography: Styles, punctuation marks, quotes, citations	1
4.4	Grammar in Technical writing: Articles, Syntax, Main and subordinate clauses; Active & passive voices; some commonly made mistakes in technical writing.	2
5	Techniques of writing: An extended abstract, a project proposal, a research paper, a technical report.	
5.1	Extended abstract: abstract and keywords, introduction and objective, method, findings and argument, conclusion and suggestions and references.	1
5.2	Project Proposal:Types, executive summary, background including status, objectives, solution, milestones, deliverables, timelines, resources, budgeting, conclusion	2
5.3	Research paper: writing an overview article: provide a comprehensive foundation on a topic; explain the current state of knowledge; identify gaps in existing studies for potential future research; highlight the main methodologies and research techniques	2
5.4	Writing Technical Reports: Title page; Summary; Table of contents; Introduction; Body; Figures, tables, equations and formulae; Conclusion; Recommendations.	1

Reference Books

1. SCIENTIFIC WRITING 2.0 A Reader and Writer's Guide: Jean-Luc Lebrun, World ScientiVic Publishing Co. Pte. Ltd., 2011

2. How to Write and Publish a ScientiVic Paper: Barbara Gastel and Robert A. Day, Greenwood publishers, 2016

3. Grammar, Punctuation, and Capitalisation; a handbook for technical writers and editors. www.sti.nasa.gov/publish/sp7084.pdf www.sti.nasa.gov/sp7084/contents.html

4. Everything You Wanted to Know About Making Tables and Figures. http://abacus.bates.edu/ %7Eganderso/biology/resources/writing/ HTWtableVigs.html

243AGE001	ADVANCED ENGINEERING MATERIALS	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: This course is designed in a way to provide a general view on typically used advanced classes of engineering materials including metals, polymers, ceramics, and composites.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Analyse the requirement and find appropriate solution for use of materials.
CO 2	Differentiate the properties of polymers, ceramics and composite materials.
CO 3	Recognize basic concepts and properties of functional materials.
CO 4	Comprehend smart and shape memory materials for various applications.
CO 5	Appraise materials used for high temperature, energy production and storage applications.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	□				□	□	
CO 2	□				□	□	
CO 3	□				□	□	
CO 4	□				□	□	
CO 5	□				□	□	

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	60%
Apply	20%
Analyse	20%

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks

Course based task : 15
marks Seminar/Quiz : 15
marks Test paper, 1 no. : 10

marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern: 60 marks

The examination will be conducted by the respective College. The examination willbe for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any Vive. Each question can carry 12 marks.

Model Question paper

AUDIT COURSE

223AGE001 - ADVANCED ENGINEERING MATERIALS

(Answer any five questions. Each question carries 12 Marks)

1.	a) State the relationship between material selection and processing.	**5**
	b) Write about the criteria for selection of materials with respect to the cost and service requirements for engineering applications.	**7**
2.	a) Differentiate thermosetting and thermoplastics with suitable examples.	**5**
	b) Briefly discuss about the properties and applications of polymer nano composite materials.	**7**
3.	a) Write about the potential application areas of functionally graded materials.	**5**
	b) With a neat sketch describe any one processing technique of functionally graded materials.	**7**
4.	a) “Smart materials are functional”? Justify the statement.	**5**
	b) Explain the terms electrostriction and magnetostriction with its application.	**7**

5.	a) What are the factors influencing functional life of components at elevated temperature?	**5**
	b) What are super alloys and what are their advantages?	**7**
6	a) What is a shape memory alloy? What metals exhibit shape memory characteristics?	**4**
	b) Explain about the detection capabilities and uses of pyroelectric sensors.	**8**
7	a) Differentiate between conventional batteries and fuel cells.	**4**
	b) Explain the construction and working of a Li-ion battery.	**8**

Syllabus

Module	**Content**	**Hours**	**Semester Exam Marks (%)**
I	Requirements / needs of advanced materials. Classification of materials, Importance of materials selection, Criteria for selection of materials; motivation for selection, cost basis and service requirements. Relationship between materials selection and processing.	**5**	**20**
II	Classification of non-metallic materials. Polymer, Ceramics: Properties, processing and applications. Nano Composites - Polymer nanocomposites (PNCs), Processing and characterisation techniques – properties and potential applications.	**7**	**20**
III	Functionally graded materials (FGMs), Potential Applications of FGMs, classification of FGMs, processing techniques. limitations of FGMs.	**6**	**20**
IV	Smart Materials: Introduction, smart material types - pyroelectric sensors, piezoelectric materials, electrostrictors and magnetostrictors, shape memory alloys – associated energy stimulus and response forms, applications.	**5**	**20**
V	High Temperature Materials: super alloys – main classes, high temperature properties of superalloys, applications.	**7**	**20**

	Energy Materials: materials for batteries.		

Course Plan

No	Topic	No. of Lectures
1	**Selection of materials for engineering applications**	
1.1	Benefits of advanced materials, classification of materials, importance of materials selection	2
1.2	Selection of materials for different properties, strength, toughness, fatigue and creep	1
1.3	Selection for surface durability, corrosion and wear resistance	1
1.4	Relationship between materials selection and processing	1
2	**Classification of non-metallic materials & nano composites**	
2.1	Rubber: properties, processing and applications.	1
2.2	Plastics: thermosetting and thermoplastics, applications and properties.	2
2.3	Ceramics: properties and applications.	1
2.4	Introduction to nano composites, classification	1
2.5	Processing and characterisation techniques applicable to polymer nanocomposites.	2
3	**Functionally graded materials**	
3.1	General concept, Potential Applications of FGMs	2
3.2	Classification of FGMs	1
3.3	FGMs processing techniques: powder metallurgy route, melt-processing route	2
3.4	Limitations of FGMs	1
4	**Smart materials**	
4.1	Introduction to smart materials, types	1
4.2	Pyroelectric sensors-material class, stimulus, detection capabilities and uses	1
4.3	Piezoelectric materials- material class, stimulus, sensing and actuating applications	1
4.4	Electrostrictors and magnetostrictors - material class, stimulus, micro positioning capabilities and applications	1
4.5	Shape memory alloys (SMAs) - material class, stimulus, temperature sensing and high strain responses, applications.	1
5	**High Temperature Materials and Energy Materials**	
5.1	Characteristics of high-temperature materials, superalloys as high-temperature materials	1
	superalloys - properties and applications	2

5.2	Introduction to lithium-ion battery (LIBs), operating mechanisms and applications	2
5.3	Introduction to Zn-based battery system, types and existing challenges	2

Reference Books

1. DeGarmo et al, "Materials and Processes in Manufacturing", 10th Edition, Wiley, 2008.
2. R.E. Smallman and A.H.W. Ngan, Physical Metallurgy and Advanced Materials, Seventh Edition, Butterworth-Heinemann, 2007
3. Vijayamohanan K. Pillai and Meera Parthasarathy, "Functional Materials: A chemist's perspective", Universities Press Hyderabad (2012).
4. M.V. Gandhi, B.S. Thompson: Smart Materials and Structures, Chapman & Hall, 1992.
5. G. W. Meetham and M. H. Van de Voorde, Materials for High Temperature Engineering Applications (Engineering Materials) Springer; 1 edition (May 19, 2000)
6. Inderjit Chopra, Jayant Sirohi, "Smart Structures Theory", Cambridge University Press, 2013

243AGE003	DATA SCIENCE FOR ENGINEERS	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	0

Preamble: This course covers essentials of statistics and Linear Algebra and how to prepare the data before processing in real time applications. The students will be able to handle missing data and detection of any outliers available in the dataset. This course explores data science, Python libraries and it also covers the introduction to machine learning for engineers.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Study Data Science Concepts and statistics
CO 2	Demonstrate Understanding of Mathematical Foundations needed for Data Science
CO 3	Understand Exploratory analysis and Data Visualization and Preprocessing on given dataset
CO 4	Implement Models such as Naive Bayes, K-Nearest Neighbors, Linear and Logistic Regression
CO 5	Build real time data science applications and test use cases

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO7
CO 1	2		2			2	
CO 2	2		2	1		2	
CO 3	2		2	2	2	2	
CO 4	2		2	2	3	2	
CO 5	2		2	3	3	3	2

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	50%
Apply	30%
Analyse	20%

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks

Course based task (Project/Assignments/Simulations/Case studies): 15 marks

Seminar/Quiz : 15 marks

Test paper, 1 no. : 10 marks

Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern:**60 marks**

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 mark.

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	**Statistics for Data science** Probability: Basic concepts of probability, conditional probability, total probability, independent events, Bayes' theorem, random variable, Population, Sample, Population Mean, Sample Mean, Population Distribution, Sample Distribution and sampling Distribution, Mean, Mode, Median, Range, Measure of Dispersion, Variance, Standard Deviation, Gaussian/Normal Distribution, covariance, correlation.	6	20
II	**Linear Algebra** Vectors and their properties, Sum and difference of Vectors, distance between Vectors, Matrices,Inverse of Matrix, Determinant of Matrix, Trace of a Matrix, Dot Product, Eigen Values, Eigen Vectors, Single Value Decomposition	6	20

III	**Hypothesis Testing** Understanding Hypothesis Testing, Null and Alternate Hypothesis, Non-directional Hypothesis, Directional Hypothesis Critical Value Method, P-Value Method, Types of Errors-Type1 Error,Type2 Error, Types of Hypothesis Test Z Test, Chi-Square	**6**	**20**
IV	**Exploratory Data Analysis** Data Collection –Public and Private Data, Data Cleaning-Fixing Rows and Columns, Missing Values, Standardizing values, invalid values, filtering data, Data-Integration,Data-Reduction,Data Transformation	**6**	**20**
V	**Machine Learning and Python for Data Science** Python Data structures-List, Tuple, Set, Dictionary, Pandas, Numpy, Scipy, Matplotlib, Machine Learning-Supervised Machine Learning, Unsupervised Machine Learning,Regression, Classification, Naïve-Bayes	**6**	**20**

Course Plan

No	Topic	No. of Lectures
1	**Statistics for Data science**	
1.1	Probability: Basic concepts of probability, conditional probability, total probability	1
1.2	independent events, Bayes' theorem, random variable, Population	1
1.3	Sample, Population Mean, Sample Mean, Population Distribution	1
1.4	Sample Distribution and sampling Distribution, Mean, Mode, Median, Range, Propositional logic and predicate logic	1
1.5	Measure of Dispersion, Variance, Standard Deviation	1
1.6	Gaussian/Normal Distribution, covariance, correlation.	1
2	**Linear Algebra**	
2.1	Vectors and their properties,	1
2.2	Sum and difference of Vectors, distance between Vectors	1
2.3	Matrices,Inverse of Matrix,	2

2.4	Determinant of Matrix, Trace of a Matrix, Dot Product, Eigen Values, Eigen Vectors, Single Value Decomposition	2
3	**Hypothesis Testing**	
3.1	Understanding Hypothesis Testing, Null and Alternate Hypothesis	1
3.2	Non-directional Hypothesis, Directional Hypothesis Critical Value Method, P-Value Method,	2
3.3	Types of Errors-Type1 Error,Type2 Error,	1
3.4	Types of Hypothesis Test Z Test, Chi-Square,	2
4	**Exploratory Data Analysis**	
4.1	Data Collection –Public and Private Data	1
4.2	Data Cleaning-Fixing Rows and Columns	1
4.3	Missing Values	1
4.4	Standardizing values	1
4.5	Invalid values, filtering data	1
4.6	Data Integration, Data Reduction, Data Transformation	1

5	**Machine Learning and Python for Data Science**	
5.1	Python Data structures-List, Tuple, Set,	1
5.2	Dictionary, Pandas, Numpy, Matplotlib	2
5.3	Machine Learning-Supervised Machine Learning, Unsupervised Machine Learning	1
5.4	Regression, Classification	1
5.5	Naïve-Bayes	1

Reference Books

1. Python Data Science Handbook. Essential Tools for Working with Data, Author(s): Jake VanderPlas, Publisher: O'Reilly Media, Year: 2016
2. Practical Statistics for Data Scientists: 50 Essential Concepts, Author(s): Peter Bruce, Andrew Bruce, Publisher: O'Reilly Media, Year: 2017
3. Practical Linear Algebra for Data Science, by Mike X Cohen, Released September 2022, Publisher(s): O'Reilly Media, Inc.
4. Data Science from Scratch ‘by Joel Grus, Released, April 2015, Publisher(s): O'Reilly Media, Inc.
5. Hands-On Exploratory Data Analysis with Python, by Suresh Kumar Mukhiya, Usman Ahmed, Released March 2020,Publisher(s): Packt Publishing

Modal Question Paper **Total Pages:**

Reg No.: _ Name:___

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

THIRD SEMESTER M.TECH DEGREE EXAMINATION, MARCH 2024

Course Code: 243AGE003

Course Name: DATA SCIENCE FOR ENGINEERS

Max. Marks: 60 Duration: 2.5 Hours

Answer any five full questions, each carries 12 marks.

1 a) It is observed that 50% of mails are spam. There is software that filters 5 spam mail before reaching the inbox. It accuracy for detecting a spam mail is 99% and chances of tagging a non-spam mail as spam mail is 5%. If a certain mail is tagged as spam finds the probability that it is not a spam mail.

b) Depict the relevance of measures of central tendency in data wrangling with a suitable example 7

2. a) Calculate the inverse of the Matrix 4

2	4	-6
7	3	5
1	-2	4

b)Find all Eigenvalues and Corresponding Eigenvectors for the matrix if 8

3. a) A statistician wants to test the hypothesis H0: $\mu = 120$ using the alternative hypothesis Hα: $\mu > 120$ and assuming that $\alpha = 0.05$. For that, he took the sample values as $n = 40$, $\sigma = 32.17$ and $\bar{x} = 105.37$. Determine the conclusion for this hypothesis? 5

b) Hypothesis testing is an integral part of statistical inference, list out the various types of hypothesis testing and also mentions their significances in data science. 7

4. a) Brief in detail directional and non-directional hypothesis 6

b) Differentiate null and alternate hypothesis and also elaborate on type 1 6 and type 2 errors

5. a) Explain the concepts of Tuple, List and Directory in python with example 6

b) Elucidate reinforcement learning and application in real world. 6

6. a) What is Feature Engineering , demonstrate with an example 6

b) Describe in detail different steps involved in data preprocessing. 6

7. a) Illustrate supervised learning model with linear regression model 5

b) Predict the probability for the given feature vector if an accident will happen or not? 7

2 0 -3

2 0 -5

0 3 0

Weather condition: rain, Road condition: good, Traffic condition: normal, Engine problem: no, the task is to predict using Naïve Bayes classification.

SNo.	Weather condition	Road condition	Traffic condition	Engine problem	Accident
1	Rain	bad	high	no	yes
2	snow	average	normal	yes	yes
3	clear	bad	light	no	no
4	clear	good	light	yes	yes
5	snow	good	normal	no	no
6	rain	average	light	no	no
7	rain	good	normal	no	no
8	snow	bad	high	no	yes
9	clear	good	high	yes	no
10	clear	bad	high	yes	yes

243AGE004	DESIGN THINKING	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble:

This course offers an introductory exploration of fundamental engineering concepts and techniques, the design process, analytical thinking and creativity, as well as the fundamentals and development of engineering drawings, along with their application in engineering problems.

Course Outcomes:

After the completion of the course the student will be able to

CO 1	Identify and frame design challenges effectively.
CO 2	Generate creative ideas through brainstorming and ideation
CO 3	Iterate on designs based on user insights
CO 4	Apply Design Thinking to real-world problems and projects.

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1				2		2	2
CO 2	2		2	2			2
CO 3		2		2		2	2
CO 4	2		2	3	2		2

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	40
Analyse	30
Evaluate	30
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: AUDIT COURSES

Continuous Internal Evaluation Pattern: 40 marks

Course based task : 15 marks
Seminar/Quiz : 15 marks Test
paper, 1 no. : 10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern: 60 marks

The examination will be conducted by the respective College. The examination willbe for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question paper

		SET1	**Total Pages:**
Reg No.:_______________		Name:_________________________	
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY THIRD SEMESTER M.TECH DEGREE EXAMINATION, MARCH 2024			
Course Code: 243AGE004			
Course Name: DESIGN THINKING			
Max. Marks: 60		Duration: 2.5 Hours	
Answer any five full questions, each carries 12 marks.			

1 a)	How can a multidisciplinary team collaborate effectively to implement design principles?	7
1 b)	What are the key differences between human-centred design and other design methodologies?	5
2 a)	How do you measure the success of a design project in terms of user satisfaction and impact?	7
2 b)	How does the iterative nature of the design process contribute to better outcomes	5

3 a)	What are the fundamental principles of effective brainstorming, and how do they differ from traditional problem-solving approaches?	7
3 b)	What are some key principles of ergonomic design, and how do they contribute to the usability and comfort of products?	5
4 a)	Enumerate some examples of successful and unsuccessful market testing scenarios, and what lessons can be learned from these experiences to improve future product or service launches?	7
4b)	What is the primary purpose of creating prototypes in the design and development process?	5
5	What strategies and methodologies can designers use to embrace agility and respond quickly to changing user needs and market dynamics?	12
6	Illustrate any four examples of successful bio-mimicry applications in various industries.	12
7	What ethical considerations should designers keep in mind when designing for diverse user groups?	12
****		

Syllabus:

Module 1

Design process: Traditional design, Design Thinking Approach, Introduction to Design Thinking, History and evolution of Design Thinking, Role of design thinking in the human-centred design process. Design space, Design Thinking in a Team Environment, Team formation.

Module 2

Design Thinking Stages: Empathize, Define, Ideate, Prototype and Test. The importance of empathy, Building a user-centred mindset. Problem statement formulation, User needs and pain points, establishing target specifications, Setting the final specifications.

Module 3

Generating Ideas, Brainstorming techniques, Application of Aesthetics and Ergonomics in Design. Bio-mimicry, Conceptualization, Visual thinking, Drawing/Sketching, Presenting ideas.

Module 4

Use of prototyping, Types of prototypes, Rapid prototyping techniques, User testing and feedback collection, Iterative prototyping, testing to gauge risk and market interest

Module 5

Entrepreneurship/business ideas, Patents and Intellectual Property, Agility in design, Ethical considerations in design. Overcoming common implementation challenges

Corse Plan SyllabusandCorsePlan (For 3credit courses, thec ontent can be for 40 hrs and for2credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30hours).

No	Topic	No. of lectures
1	**Design process:**	
1.1	Design process: Traditional design, Design Thinking Approach, Introduction to Design Thinking, History and evolution of Design Thinking.	3
1.2	Role of design thinking in the human-centred design process. Design space,	2
1.3	Design Thinking in a Team Environment, Team formation.	2
2	**Design Thinking Stages:**	
2.1	Design Thinking Stages: Empathize, Define, Ideate, Prototype and Test.	2
2.2	The importance of empathy, Building a user-centred mindset.	2
2.3	Problem statement formulation, User needs and pain points, establishing target specifications, Setting the final specifications.	3
3	**Ideation**	
3.1	Generating Ideas, Brainstorming techniques.	2
3.2	Application of Aesthetics and Ergonomics in Design. Bio-mimicry.	3
3.3	Conceptualization, Visual thinking, Drawing/Sketching, Presenting ideas.	2

4	**Prototyping and testing**	
4.1	Use of prototyping, Types of prototypes, Rapid prototyping techniques.	3
4.2	User testing and feedback collection, Iterative prototyping, testing to gauge risk and market interest	2
5	**IPR in design**	
5.1	Entrepreneurship/business ideas, Patents and Intellectual Property.	2
5.2	Agility in design, Ethical considerations in design. Overcoming common implementation challenges	2

Reference Books

1. Christoph Meinel, Larry Leifer and Hasso Plattner- “Design Thinking: Understand – Improve – Apply”, Springer Berlin, Heidelberg, 2011.
2. Thomas Lockwood and Edgar Papke – “Design Thinking: Integrating Innovation, Customer Experience, and Brand Value", Allworth Press, 2009.
3. Pavan Soni – “Design Your Thinking”, Penguin Random House India Private Limited, 2020.
4. Andrew Pressman- “Design Thinking : A Guide to Creative Problem Solving for Everyone”, Taylor & Francis, 2018.
5. N Siva Prasad, “Design Thinking Techniques an Approaches” Ane Books Pvt. Ltd.,2023

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
243AGE005	**FUNCTIONAL PROGRAMMING IN HASKELL**	**AUDIT COURSE**	**3**	**0**	**0**	**-**

Preamble: This course introduces a functional programming approach in problem solving. Salient features of functional programming like recursion, pattern matching, higher order functions etc. and the implementation in Haskell are discussed.

Course Outcomes:

After the completion of the course the student will be able to

CO 1	Understand the functional programming paradigm which is based on the mathematics of lambda calculus.
CO 2	Develop Haskell programs using functions, guards and recursive functions
CO 3	Apply the concept of tuples, lists and strings in Haskell programming
CO 4	Apply the concept of algebraic data types, abstract data types, modules, recursive data types and user defined data types in Haskell programming
CO 5	Develop Haskell programs with files for reading input and storing output

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1					3		
CO 2	2			2	3		
CO 3	2			2	3		
CO 4	2			2	3		
CO 5	2			2	3		

Assessment Pattern

Bloom's Category	End Semester Examination
Apply	40%
Analyse	40%
Evaluate	20%
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration

100	40	60	2.5 hours

Continuous Internal Evaluation: 40 marks

Course based task : 15 marks

Seminar/Quiz : 15 marks

Test paper, 1 no. : 10 marks

Test paper shall include minimum 80% of the syllabus**.**

End Semester Examination: 60 marks

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question paper

		Total Pages:
Reg No.:______________		Name:______________________

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
THIRD SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2023

Course Code: 243AGE005

Course Name: Functional Programming in Haskell

Max. Marks: 60	Duration: 2.5 Hours

Answer any five full questions, each carries 12 marks.

1 a.	Explain the basic differences between imperative style programming and functional style programming.	3
1 b.	Analyse each of the following lambda expressions to clarify its structure. If the expression is a function, identify the bound variable and the body expression, and then analyse the body expression. If the expression is an application, identify the function and argument expressions, and then analyse the function and argument expressions: i) λa.(a λb.(b a)) ii) λx.λy.λz.((z x) (z y)) iii) (λf.λg.(λh.(g h) f) λp.λq.p)	9

2 a.	Design a recursive function to find 2^n where n is a natural number.	4

2 b.	Explain various forms of function definitions in Haskell with the help of examples.	8
3 a.	Explain any three list operations along with function definitions and examples.	6
3 b.	Write a program to duplicate only even numbers among the elements of a list using a Haskell function by (i) Recursion (ii) List Comprehension and explain. Example : λ> dupli [1, 2, 3] ANS: [2,2]	6
4	Write Recursive definitions along with an explanation for the below arithmetic operations. Illustrate the recursive flow with the help of a diagram. i. add x y ii. mult x y iii. div x y	12
5	Write the Haskell code to split a list into two lists such that the elements with odd index are in one list while the elements with even index are in the other list.	12
6 a	Give the type definition of a binary tree along with explanation of two functions on binary trees.	6
6 b	Define a queue data type in Haskell along with any two operations on it with examples.	6
7 a.	Explain the basic steps of reading from files and writing to files in Haskell.	4
7 b.	Write a Haskell program to read from the file "input.txt", display the contents on the screen and write the contents to another file "output.txt".	8

Syllabus and Corse Plan (For 3 credit courses, the content can be for 40 hrs and for 2 credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30 hours).

Module 1 (5 Hrs)

Introduction to Functional Programming: Programming language paradigms, imperative style programming, comparison of programming paradigms.
Functional programming, Functions - Mathematical concepts and terminology, Lambda calculus, Function definitions, programs as functions, Functional programming Languages. Haskell basics, GHCi interpreter.

Module 2 (6 Hrs)

Programming in Haskell: Expressions and evaluation, Lazy evaluation, let expressions, scopes.

Basic data types in Haskell, operators, infix operators, associativity and precedence, Arithmetic functions.

types, definitions, currying and uncurrying, type abstraction.

Function definitions, pattern matching, guards, anonymous functions, higher order functions.

Recursion, Programming exercises.

Module 3 (7 Hrs)

Data types: tuples and lists: Tuples , Lists: building lists, decomposing lists, functions on lists, built-in functions on lists, primitive and general recursion over lists, infinite lists.

Strings: functions on strings.

Polymorphism and overloading, conditional polymorphism

Module 4 (6 Hrs)

Type classes, Algebraic data types, Modules, Recursive data types.

User defined data types, Records, Stacks, Queues, Binary trees, Constructors, Destructors.

Module 5 (6 Hrs)

Functor, Applicative functor, Monad

Programming with actions: Functions vs actions, Basics of input / output, the do notation, interacting with the command line and lazy I/O, File I/O.

No	Topic	No. of Lectures
1	**Introduction to Functional Programming**	
1.1	Programming language paradigms, imperative style programming, comparison of programming paradigms	1
1.2	Functional programming, Functions - Mathematical concepts and terminology	1
1.3	Lambda calculus	1
1.4	Function definitions, programs as functions, Functional programming Languages	1
1.5	Haskell basics, GHCi interpreter	1
2	**Haskell basics**	
2.1	Expressions and evaluation, Lazy evaluation	1
2.2	let expressions, scopes, Basic data types in Haskell	1
2.3	operators, infix operators, associativity and precedence, Arithmetic functions	1

2.4	types, definitions, currying and uncurrying, type abstraction.	1
2.5	Function definitions, pattern matching, Guards	1
2.6	anonymous functions, higher order functions, Recursion	1
3	**Data types: tuples and lists**	
3.1	Tuples , Lists: building lists, decomposing lists	1
3.2	functions on lists, built-in functions on lists	1
3.3	primitive and general recursion over lists	1
3.4	infinite lists	1
3.5	Strings: functions on strings	1
3.6	Polymorphism and overloading	1
3.7	conditional polymorphism	1
4	**User defined data types**	
4.1	Type classes, Algebraic data types, Modules	1
4.2	Recursive data types	1
4.3	User defined data types, Records	1
4.4	Stacks, Queues	1
4.5	Binary trees	1
4.6	Constructors, Destructors	1
5	**Programming with actions**	
5.1	Functor, Applicative functor,	1
5.2	Monad	1
5.3	Programming with actions: Functions vs actions, Basics of input / output, the do notation	1
5.4	interacting with the command line and lazy I/O	1
5.5	File I/O	2

Reference Books

[1] Richard Bird, “Introduction to functional programming using Haskell', second edition, Prentice hall series in computer science

[2] Bryan O'Sullivan, Don Stewart, and John Goerzen, “Real World Haskell”

[3] Richard Bird, “Thinking Functionally with Haskell”, Cambridge University Press, 2014

[4] Simon Thompson, “Haskell: The Craft of Functional Programming”, Addison-Wesley, 3rd Edition, 2011

[5] H. Conrad Cunningham, “Notes on Functional Programming with Haskell”, 2014

[6] Graham Hutton, “Programming in Haskell”, Cambridge University Press, 2nd Edition, 2016

[7] Alejandro Serrano Mena, “Practical Haskell: A Real-World Guide to Functional Programming”, 3rd Edition, Apress, 2022

[8] Miran Lipovaca, “Learn You a Haskell for Great Good!: A Beginner's Guide”, No Starch Press, 2011

243AGE010	REUSE AND RECYCLE TECHNOLOGY	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: "Reuse and Recycle Technology" typically focuses on sustainable practices and technologies aimed at reducing waste, conserving resources, and promoting environmental responsibility.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Explain the principles and technologies behind waste reduction, resource conservation, and sustainable practices
CO 2	Describe and Analyze waste generation and management.
CO 3	Apply the knowledge of various reuse strategies and their application in different industries and Analyze various recycling technologies
CO 4	Appraise the methods of E-waste management and Eco friendly packaging
CO 5	Comprehend Environmental Regulations and Policies**,** Understand the importance of environmental regulations and policies in addressing environmental challenges

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1			3			
CO 2				3		
CO 3				3		
CO 4					3	
CO 5			3			

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	60%
Apply	20%
Analyse	20%

Mark distribution

Total Mark s	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks

Course based task : 15 marks
Seminar/Quiz : 15 marks Test
paper, 1 no. : 10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern: 60 marks

The examination will be conducted by the respective College. The examination willbe for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question paper

AUDIT COURSE

223AGE010 - REUSE AND RECYCLE TECHNOLOGY

Answer any five full questions, each carries 12 marks.

1.	(a) What are the 3 pillars of sustainability? *(b)* What is sustainable waste management? What makes sustainable waste management so important?	5 7
2.	(a)How do the three categories of municipal solid waste differ? *(b)* Discuss the municipal waste collection and management?	5 7
3.	(a)Explain the major differences between Reuse and Recycle? *(b)* Give an overview of recycling technologies used for any two materials. Discuss the Process involved.	5 7
4.	(a)What are the common source of E-waste *(b)* What are the challenges and opportunities in E-waste management	5 7
5.	(a)What is the case law for waste recycling in India *(b)* Discuss sustainable packaging and its environmental impacts	5 7
6.	Explain the various environmental regulations in India for addressing Environmental challenges	12
7.	a) Give examples of water reuse technologies in circular economy b) How can we reduce e-waste with sustainable solutions	5 7

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	**Introduction to Sustainability**, Understanding sustainability and its importance, The three pillars of sustainability: Environmental, Social, and Economic. Biodiversity conservation, Climate change and mitigation Sustainable resource management.	**6**	**20**
II	**Waste Management,** Definition and classification of waste, Waste Generation and Composition**,** Waste Collection and Transportation, Waste Segregation and Sorting. Waste Disposal Methods Historical perspectives on waste management, The three Rs: Reduce, Reuse, and Recycle.	**6**	**20**
III	**Recycling and Reuse:** Importance of reuse, Application of reuse in various industries, Challenges and opportunities in reuse, Overview of recycling technologies, Circular economy, Sorting and processing of recyclable materials, Advanced recycling methods. Emerging technologies in recycling.	**6**	**20**
IV	**E-waste Recycling,** Challenges and environmental impact of electronic waste, E-waste recycling methods and regulations, Sustainable electronics design, **Sustainable Packaging,** Packaging materials and their environmental impact, Eco-friendly packaging alternatives, Packaging design for sustainability	**6**	**20**
V	**Environmental Regulations and Policies,** Understand the importance of environmental regulations and policies in addressing environmental challenges, National and international waste and recycling regulations, Compliance and enforcement, Industry standards and certifications	**6**	**20**

Course Plan

No	Topic	No. of Lectures
1	**Introduction to Sustainability (6)**	
1.1	Understanding sustainability and its importance	1
1.2	The three pillars of sustainability: Environmental, Social, and Economic.	3
1.3	Biodiversity conservation, Climate change and mitigation	1
1.4	Sustainable resource management	1
2	**Waste Management (6)**	
2.1	Definition and classification of waste	1
2.2	Waste Generation and Composition	1
2.3	Waste Collection and Transportation.	1
2.4	Waste Segregation and Sorting.	1
2.5	Waste Disposal Methods	1
2.6	Historical perspectives on waste management, The three Rs: Reduce, Reuse, and Recycle.	1
3	**Recycling and Reuse (6)**	
3.1	Importance of reuse, Examples of reuse in various industries.	1
3.2	Challenges and opportunities in reuse	1
3.3	Overview of recycling technologies, Sorting and processing of recyclable materials	2
3.4	Advanced recycling methods	1
3.5	Emerging technologies in recycling.	1
4	**E-waste Recycling (6)**	
4.1	Challenges and environmental impact of electronic waste	1
4.2	E-waste recycling methods and regulations	1
4.3	Sustainable electronics design	1
4.4	Packaging materials and their environmental impact	1
4.5	Eco-friendly packaging alternatives	1
4.6	Packaging design for sustainability	1
5	**Environmental Regulations and Policies (6)**	
5.1	Importance of environmental regulations and policies in addressing environmental challenges	2
5.2	National and international waste and recycling regulations	2
5.3	Industry standards and certifications, Compliance and enforcement	2

Reference Books

1. Sustainable Engineering: Concepts, Design and Case Studies, David T. Allen, Pearson Publication.
2. A Comprehensive Book on Solid Waste Management with Application, Dr. H.S. Bhatia , Misha Books, 2019
3. "Cradle to Cradle: Remaking the Way We Make Things" by William McDonough and Michael Braungart.
4. "Recycling of Plastic Materials" edited by Vijay Kumar Thakur
5. E-waste: Implications, Regulations and Management in India and Current Global Best Practices, Rakesh Johri, TERI
6. "Sustainable Packaging", Subramanian Senthilkannan Muthu , Springer Nature.
7. **I**ndian Environmental Law: Key Concepts and Principles " Orient Black swan Private Limited, New Delhi.

243AGE012	EXPERT SYSTEMS	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: The course aims to provide an understanding of the basic concepts of Artificial Intelligence (AI) and Expert Systems. The course also covers the knowledge representation in expert systems, classes of expert systems, applications of expert systems.

Course Outcomes: After the completion of the course the student will be able to:

CO 1	Explain the concepts of Artificial Intelligence and different ways of knowledge representations.
CO 2	Explain the components of expert systems, development stages of expert systems and tools available for expert system design.
CO 3	Apply the concept of knowledge representation in expert systems
CO 4	Differentiate the classes of expert systems and examine properties of existing systems

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO7
CO 1	1		2	1	2	2	
CO 2	1		1	3	2	2	
CO 3	1		1	2	2	2	
CO 4	2		2	2	3	2	

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	60%
Apply	20%
Analyse	20%

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks
Course based task (Project/Assignments/Simulations/Case studies): 15 marks
Seminar/Quiz : 15 marks
Test paper, 1 no. : 10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern:60 marks
The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 mark.

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY THIRD SEMESTER M.TECH DEGREE EXAMINATION, MARCH 2024		
Course Code: 243AGE012		
Course Name: EXPERT SYSTEMS		
Max. Marks: 60		Duration: 2.5 Hours
Answer any five full questions, each carries 12 marks.		
1	a) What are the types of AI? Explain with examples .	6
	b) What do you mean by knowledge in AI and explain the different ways of knowledge representation used in AI?	6
2.	a) Write note on semantic network.	6
	b) What are Predicates? Explain its syntax and semantics.	6
3.	a) Write notes on different tools available for expert system design.	6
	b). What are the different stages in the development of an expert system?	6
4.	a) Illustrate Conceptual Dependencies with an example.	6
	b) Illustrate with an example the Structured Knowledge representation of an Expert System.	6
5.	a) What do you mean by Frame based Expert System? Explain	6
	b)Explain the architecture of MYCIN	6
6.	a)Explain Fuzzy based expert systems	6
	b) Explain the neural network based expert systems	6

7.	a) Explain any two applications of expert systems?	6
	b)What are the limitations of expert system ? Explain	6

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	Overview of Artificial Intelligence (AI): Definition & Importance of AI. Knowledge general concepts: Definition and Importance of knowledge, Knowledge-Based Systems, Knowledge organization, Knowledge Manipulation and acquisition. Knowledge Representation: Introduction, Syntax and Semantics- Propositional logic and predicate logic.	6	20
II	Basic concepts of expert systems-Introduction to expert systems, Components of expert systems. Features of Expert System, Stages in the development of expert system, Types of tools available for expert system design	6	20
III	Knowledge representation in expert systems: Structured Knowledge representation: Graphs, Frames and related structures, Associative networks, Conceptual dependencies, Examples of structured knowledge representation.	6	20
IV	Classes of expert systems: Rule-based expert systems, Example- MYCIN, Frame-based expert system, terminologies, IF-THEN structure. Fuzzy and Neural network based expert systems(basic concepts)	7	20
V	Currents trends in expert systems, Advantages and limitations of expert systems, Applications of expert systems.	5	20

Course Plan

No	Topics	No. of Lectures
1	**Overview of Artificial Intelligence& Knowledge general concepts**	
1.1	Definition & Importance of AI	1
1.2	Definition and Importance of Knowledge,	1
1.3	Knowledge-Based Systems, Knowledge Organization	1
1.4	Knowledge Manipulation and acquisition	1
1.5	Knowledge Representation: Introduction, Syntax and Semantics	1
1.6	Propositional logic and predicate logic	1
2	**Basic concepts of expert systems**	
2.1	Introduction to Expert System, Components of expert systems	2
2.2	Features of Expert System, Stages in the development of expert system	2
2.3	Types of tools available for expert system design	2
3	**Knowledge representation in expert systems**	
3.1	Structured Knowledge representation	1
3.2	Graphs, Frames and Related Structures	2
3.3	Associative Networks, Conceptual Dependencies	2
3.4	Examples of structured knowledge representation	1
4	**Classes of expert systems**	
4.1	A rule-based expert system -Introduction	1
4.2	MYCIN	1
4.3	IF-THEN structure	1
4.4	Frame-based expert system	2
4.5	Fuzzy based expert systems	1
4.6	Neural network based expert systems	1
5	**Currents trends and applications of expert systems**	
5.1	Currents trends of expert systems	2
5.2	Advantages and limitations of expert systems	1
5.3	Applications of expert systems	2

Reference Books

1. E. Rich & K. Knight - Artificial Intelligence, 2/e, TMH, New Delhi, 2005.
2. P.H. Winston - Artificial Intelligence, 3/e, Pearson Edition, New Delhi, 2006.
3. D.W. Rolston - Principles of AI & Expert System Development, TMH, New Delhi
4. Kevin Night and Elaine Rich, Nair B., “Artificial Intelligence (SIE) “, McGraw Hill – 2010
5. Dan W Patterson, ‘Introduction to Artificial intelligence and Expert systems’, Prentice Hall of India Pvt. Ltd,2007
6. Russel (Stuart), ‘Artificial Intelligence- Modern approach, Pearson Education series in AI’, 3rd Edition, 2009.
7. I. Gupta, G. Nagpal · Artificial Intelligence and Expert Systems, Mercury Learning and Information -2020

243AGE011	SYSTEM MODELLING	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: Study of this course provides the learners a clear understanding of fundamental concepts in simulation and modelling. This course covers the different statistical models, importance of data collection and various types of simulations. The course helps the learners to find varied applications in engineering, medicine and bio-technology.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Analyse the requirement and find appropriate tool for simulation.
CO 2	Differentiate the different statistical models.
CO 3	Discuss the different techniques for generating random numbers.
CO 4	Analyse the different methods for selecting the different input models..
CO 5	Discuss the different measures of performance and their estimation

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	2		1	1	2	
CO 2	2		1	1	1	
CO 3	1					
CO 4	1		1	1		
CO 5	2		1	1	1	

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	60%
Apply	20%
Analyse	20%

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Course based task (Project/Assignments/Simulations/Case studies): 15 marks
Seminar/Quiz: 15 marks

Test paper, 1 no.: 10 marks

Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern:

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question paper

AUDIT COURSE

243AGE001 – SYSTEM MODELLING

Answer any five questions Each carries 12 marks

PART A

1. a. Discuss the advantages and disadvantages of simulation. (5marks)

   b. What are the areas of applications of simulation (7 marks)

2. a.A bus arrives every 20 minutes at a specified stop beginning at 6:40 A.M. and continuing until 8:40 A.M. A certain passenger does not know the schedule, but arrives randomly (uniformly distributed) between 7:00A.M. and 7:30 A.M. every morning. What is the probability that the passenger waits more than 5 minutes for a bus? (5 marks)

   b. A production process manufactures computer chips on the average at 2% nonconforming. Every day, a random sample of size 50 is taken from the process. If the sample contains more than two nonconforming chips, the process will be stopped. Compute the probability that the process is stopped by the sampling scheme. (7 marks)

3. a.Discuss the different types of tests for random numbers. (5 marks)

   b. Generate random numbers using multiplicative congruential method with X0 = 5, a 11, and m = 64. (7 marks)

4. a. What are the different methods of data collection. (4marks)

   b. Records pertaining to the monthly number of job-related injuries at an underground coalrnine were being studied by a federal agency. The values for the past 100 months were as follows:

Injuries per Month	Frequency of Occurrence
0	35
1	40
2	13
3	6

4	4
5	1
6	1

(a) Apply the chi-square test to these data to test the hypothesis that the underlying distribution is Poisson. Use the level of significance α== 0.05.
(b) Apply the chi-square test to these data to test the hypothesis that the distribution is Poisson with mean 1.0. Again let α = 0.05. ·
c) What are the differences between parts (a) and (b), and when might each case arise? (8 marks)

5. a.What is the difference between validation and verification.(5 marks)
   b. Discuss the different measures of performance and their estimation(7 marks)
6. a. Discuss the different methods of parameter estimation(5 marks)
   b. With an example, describe the Poisson process.(7 marks)
7. a. Distinguish between discrete and continuous systems(5 marks)
   b.What are the different components of a simulation system(7 marks)

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	When simulation is the appropriate tool. Advantages and disadvantages of Simulation; Areas of application, Systems and system environment; Components of a system; Discrete and continuous systems, Model of a system; Types of Models, Discrete-Event System Simulation, Steps of a simulation study.	6	20
II	Review of terminology and concepts, Useful statistical models, Discrete distributions. Continuous distributions, Poisson process, Empirical distributions. (basic idea only)	6	20
III	Properties of random numbers; Generation of pseudo-random numbers, Techniques for generating random numbers, Tests for Random Numbers	6	20
IV	Data Collection; Identifying the distribution with data, Parameter estimation, Goodness of Fit Tests, Fitting a non-stationary Poisson process, Selecting input models without data, Multivariate and Time-Series input models.	6	20

V	Measures of performance and their estimation, Output analysis for terminating simulations, Output analysis for steady-state simulations, Verification, calibration and validation	**6**	**20**

Course Plan

No	Topic	No. of Lectures
1	**Introduction**	
1.1	When simulation is the appropriate tool	1
1.2	Advantages and disadvantages of Simulation;	1
1.3	Areas of application, Systems and system environment;	1
1.4	Components of a system; Discrete and continuous systems,	1
1.5	Model of a system; Types of Models,	1
1.6	Discrete-Event System Simulation ,Steps of a simulation study	1
2	**Statistical Models in Simulation**	
2.1	Review of terminology and concepts, Empirical distributions. (basic idea only)	1
2.2	Useful statistical models,	1
2.3	Discrete distributions.	1
2.4	Continuous distributions,.	1
2.5	Poisson process	1
2.6	Empirical distributions	1
3	**Random Number Generation**	
3.1	Properties of random numbers;	1
3.2	Generation of pseudo-random numbers,	
3.3	Techniques for generating random numbers	1
3.4	Techniques for generating random numbers(cont)	1
3.5	Tests for Random Numbers	1
3.6	Tests for Random Numbers(cont)	1
4	**Input Modelling**	
4.1	Data Collection;	1
4.2	Identifying the distribution with data.	1
4.3	Parameter estimation, Goodness of Fit Tests	1
4.4	Fitting a non-stationary Poisson process	1
4.5	Selecting input models without data,	1
4.6	Multivariate and Time-Series input models	1
5	**Measures of Performance and their Estimation**	
5.1	Measures of performance and their estimation	1
5.2	Measures of performance and their estimation(cont)	1

5.3	Output analysis for terminating simulations	1
5.4	Output analysis for steady-state simulations	1
5.5	Verification, calibration and validation	1
5.6	Verification, calibration and validation(cont)	1

Textbooks:

1. Jerry Banks, John S. Carson II, Barry L. Nelson, David M. Nicol: Discrete-Event System Simulation, 5th Edition, Pearson Education, 2010.

Reference Books:

1. Lawrence M. Leemis, Stephen K. Park: Discrete – Event Simulation: A First Course, Pearson Education, 2006.

2. Averill M. Law: Simulation Modeling and Analysis, 4 th Edition, Tata McGraw-Hill, 2007

3. System Modelling and Response by Ernest O. Doebelin

4. Averill M Law, "Simulation Modeling and Analysis",McGraw-Hill Inc,2007
Geoffrey Gorden, "System Simulation",Prentice Hall of India,1992.

243AGE009	Principles of Automation	CATEGORY	L	T	P	CREDIT
		CREDIT COURSE	3	0	0	0

Preamble:

This course deals in detail with the various aspects of automation such as sensors, actuators, controllers, mechanical and electrical elements and their integration for automating new and existing manufacturing and process industries and applications. This course will be beneficial to students in designing automation schemes for industries and to design automated systems

Course Outcomes: After the completion of the course the student will be able to

CO 1	Explain the fundamentals of sensor systems and to choose a suitable sensor system for the given application based on the evaluation of the constraints.
CO 2	Explain the fundamentals of signal conditions and to design a suitable signal conditioning scheme for given application.
CO 3	Describe the characteristics of various actuator systems and to decide the right type of actuator for the given application.
CO 4	Describe the importance of an industrial robot and fundamentals of numerical control in automation.
CO 5	Explain the fundamentals of controllers used in industrial automation and to construct simple automation schemes by ladder logic programs.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	2		2	2	2		
CO 2	2		2	2	2		
CO 3	2		2	2	2		
CO 4	2		2	2	2		
CO 5	2		2				

Assessment Pattern

Bloom's Category	End Semester Examination
Understand	70 %
Apply	30 %

Mark distribution

Total Marks	CIE	ESE	ESE Duration

100	40	60	2.5 hours

Continuous Internal Evaluation Pattern: 40 marks
Course based task (Project/Assignments/Simulations/Case studies): 15 marks
Seminar/Quiz: 15 marks
Test paper, 1 no.: 10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination Pattern:**60 marks**
The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question Paper
243AGE009 Principles of Automation

Time 2.5 Hrs | Marks 60

Answer any five questions Each carries 12 marks

1. (a) Differentiate the static and dynamic characteristics of a temperature sensor and explain how it affects the selection of a suitable temperature sensor. (6 marks)
   (b) Explain the working of a strain-gauge. (6marks)

2. (a) Explain why anti-aliasing filters are used in analog to digital converters. (3 marks)
   (b) Design a first order low pass filter with a cutoff frequency of 2 kHz. (9 marks)

3. (a) What are the factors to consider while deciding choosing between hydraulic, pneumatic or electrical actuation systems for an automation scheme? (4 marks)
   (b) Explain the working of a three-way pressure reducing valve. (4 marks)
   (c) Explain the working of solenoids. In what applications would you use a Solenoid valve. (4 marks)

4. (a) Explain the principle of the Touch sensor and also mention how they are used in robots. (5 marks)
   (b) Explain the basic terminologies in robotic system and also explain the components of robotic system. (7 marks)

5. (a)With neat schematic explain the architecture of the PLC. (6 marks)
   (b) Explain the use of an up-down counter in PLC with a suitable example. (6 marks)

6. (a) Write short note on SCADA. What is difference PLC and SCADA? (3 marks)

(b)Construct a ladder logic for controlling a process tank as per the logic given below;

The tank should be filled by a valve V1 when low level float switch L1 is ON and an external input S1 is received.
V1 should be closed when the liquid level reaches a high-level float switch L2.
An agitator motor should be turned on after a delay of 5sec after L2 is triggered.
After agitating for 30mins, contents of the tank should be emptied by opening another valve V2.
The temperature should be maintained at 70℃ using a thermostat T1 and Heater H
(9 marks)

7. (a) Explain the levels of Automation. (6 marks)
(b) Explain the working of Flow sensor (6 marks)

Syllabus and Course Plan

No	Topics	No. of Lectures
1	**Introduction to Industrial Automation**	
1.1	Basic Elements of an Automated System, Levels of Automation	2
1.2	Hardware components for Automation: Sensors, classification, Static and dynamic behaviour of sensors.	2
1.3	Basic working principle of different sensors: Proximity sensors, Temperature sensors, flow sensors, Pressure sensors, Force sensors. Position sensors	4
2	**Signal conditioning**	
2.1	Need for signal conditioning, Types of signal conditioning.	2
2.2	Signal conditioning using operational amplifier-Amplifier (Inverting and Non-inverting) and Filter circuits (Basic concepts). Design of first order low pass filter.	2
2.3	Signal conditioning for data acquisition systems, anti-aliasing filters, Analog–Digital Conversions, Analog-to-Digital Converters (ADC)-Steps in analog-to-digital conversion, Successive Approximation Method, Digital-to-Analog Converters (DAC)- Steps in digital to analog conversion, Zero-order and first order data hold circuits	4
3	**Actuators**	
3.1	Types of actuators- mechanical, electrical, pneumatic and hydraulic actuators. (Basic working principle)	2
3.2	Mechanical systems for motion conversion, transmission systems	3

3.3	Solenoids, Electric and stepper motors control.	3
4	**Robotics and Automated Manufacturing Systems**	
4.1	Robot Anatomy and Related Attributes: Joints and Links, Common Robot Configurations, Joint Drive Systems, Sensors in Robotics (Basic concepts)	3
4.2	Robot Control Systems, Applications of Industrial Robots- Material handling	4
4.3	Fundamentals of Numerical control (NC) Technology	1
5	**Discrete Control and Programmable Logic Controllers**	
5.1	Discrete Process Control: Logic and Sequence control	2
5.2	Ladder Logic Diagrams, Programmable Logic Controllers: Components of the PLC, PLC Operating Cycle, Programming the PLC (Basic concepts only)	4
5.3	Introduction to Distributed control system (DCS) and Supervisory Control and Data Acquisition Systems (SCADA)	2

Reference Books

1. Mikell Groover, Automation, Production Systems, and Computer-Integrated Manufacturing, 5th Edition, Pearson, 2019.
2. Yoram Koren, “Computer Control of Manufacturing Systems”,TataMcGraw Hill Edition2005.
3. S. R. Deb; Sankha Deb. Robotics Technology and Flexible Automation, Second Edition McGraw-Hill Education: New York, 2010.
4. W. Bolton, “Mechatronics: Electronic Control Systems in Mechanical and Electrical Engineering” - PrenticeHall- 2013 - 5th Edition.
5. Doebelin, E.O. and Manic, D.N., “Measurement Systems: Applications and Design”, 7th Edition, McGraw Hill, 2019.
6. Krishna Kant, Computer Based Industrial Control-, EEE-PHI,2nd edition,2010.
7. Nathan Ida, Sensors, Actuators, and Their Interfaces- A multidisciplinary introduction, 2nd Edition, IET Digital Library, 2020.
8. Salivahanan, S., and VS Kanchana Bhaaskaran. Linear integrated circuits. McGraw-Hill Education, 2nd edition, 2014.
9. Petruzella, Frank D. Programmable logic controllers. Tata McGraw-Hill Education, 2005
10. Chapman and Hall, “Standard Handbook of Industrial Automation”, Onsidine DM C & Onsidine GDC”, NJ, 1986

243AGE002	FORENSIC ENGINEERING	CATEGORY	L	T	P	CREDIT
		Audit Course	3	0	0	-

Preamble: This course explores various aspects of Forensic Engineering and different methods , tools and procedures used by Engineers to investigate and analyze . The students will learn to develop their awareness in Forensic Engineering .

Pre-requisite: Nil

Course Outcomes:

After the completion of the course the student will be able to

CO 1	Identify the fundamental aspects of forensic Engineering
CO 2	Apply forensic Engineering in Practical work flow and Investigation
CO 3	Apply methods and analysis in Forensic Investigation
CO 4	Develop practical strategies and standards of Investigation
CO 5	Create an awareness in criminal cases and create Engineering expertise in court room on forensic Engineering

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO7
CO 1	2	2	3	3	3	3	
CO 2	2	2	3	3	3	3	1
CO 3	3	3	3	3	3	3	1
CO 4	3	3	3	3	3	3	1
CO 5	3	3	3	3	3	3	

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation	End Semester Examination
Apply	40 %	60 %
Analyse	40 %	40 %
Evaluate	20 %	

Mark distribution

Total Marks	CIE	ESE	ESE Duration

100	40	60	2.5 hours

Continuous Internal Evaluation: 40 marks

Course based task :15marks
Seminar/Quizz :15marks
Test paper :10 marks
Test paper shall include minimum 80% of the syllabus.

End Semester Examination: 60 marks

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

Model Question paper

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

THIRD SEMESTER M. TECH DEGREE EXAMINATION

Course Code: 243AG002

Course Name: FORENSIC ENGINEERING

Max. Marks: 60 **Duration: 2.5 Hours**

PART A

Answer any 5 questions, each question carries 12 marks. **Marks**

1. (a) What are the uses of forensic engineering in legal laws ? (7)
   (b) Discuss the professional responsibility of a forensic Engineer . (5)

2. (a) What are the steps in preliminary on site Investigation ? (7)
   (b) With suitable examples, explain photo cataloguing? (5)

3.

4. (a) Discuss STEP method . (7)
   (b) Explain root cause Analysis (5)

5. (a) Detail about EDAX Method. (7)
   (b) Enlist the uses of NDT in forensic Analysis with example (5)

6. (a) Differentiate NFPA & FMV Standards (7)

(b) Briefly discuss the term Email Phishing ? (5)

7. Define the responsibility and duty of a forensic expert in the court. (12)

8. Explain Forensic Engineering workflow with examples (12)

Syllabus and Course Plan

Module No	Topic	No. of Lectures (Hours)
1	**Module 01: Introduction to Forensic Engineering (6 Hours)**	
1.1	Forensic Engineering-Definition, Investigation Pyramid, Eyewitness Information, Role in Legal System	2
1.2	Scientific Method-Applying scientific methods in Forensic Engineering-Engineer as expert Witness-Scientific methods and legal system	2
1.3	Qualification of Forensic Engineer-Technical- Knowledge- Oral-written-Communication- other skills-Personality Characteristics	1
1.4	Ethics and professional responsibilities.	1
2	**Module 02: Forensic Engineering Workflow and Investigation Methods (6 Hours)**	
2.1	Forensic Engineering Workflow-Team &planning-preliminary onsite investigation. Sampling-selection of sample-collection- packing-sealing of samples.	2
2.2	Source and type of evidence - Paper documentation- digital documentation-electronic data. Physical Evidence-Collection of photograph-cataloguing -Recognizing the Evidence-organizing-Evidence Analysis -Reporting	2
2.3	Investigation Methods- Cause and Causal mechanism analysis-Time and event sequence-STEP method. Human Factors, Human errors -Analysis of Operative Instruction and working Procedures	2
3	**Module 03: Physical Product Failure & Analytical Methods (6 Hours)**	
3.1	Introduction to typical Forensic Engineering Tool box-NDT, Crack detection and human eye -Hardness testing- and Destructive testing Methods with case studies	2
3.2	Indirect stress strain Analysis-Brittle lacquer technique, Contact Radiography-Metallography-EDAX method	1

3.3	Forensic Optical Microscopy-Examination-Magnification-USB Microscopy -Wifi Enabled microscopy -Reflected microscopy	2
3.4	Novel Tools and System -Contour Method-Flash Thermography-Thermographic signal reconstruction (TSR)-Electromagnetically induced acoustic Emission (EMAE)-Pulsed Eddy Current (PEA)-Theory only	1
4	**Module 04: Cyber Forensic , Civil ,Electrical Accidents & Standards (6 Hours)**	
4.1	Basics of Digital & Cyber forensics: Technical concepts; labs and tools; collecting evidence Operating System Forensic basics with - Windows, Linux -Mobile Forensic-Anti forensics-Malware- Web attack forensics with Email Crimes-Cyber Laws	3
4.2	Different types of Forensic accident investigations- Civil Engineering-Structural- Road accidents -Fire accidents - Water related accidents-Electrical accidents and Investigation methods	2
4.3	Protocol for forensic Investigations-Standard guides-scope significance - use -procedures- reports. Standards – ASTM standards -FMV Standards - SAE Standards -Relevant Standards -NFPA Standards -International Standards	1

5	**Module 05: Engineer in the Court room& Criminal Cases (6 Hours)**	
5.1	Role of an Engineering Expert-Report-pre trial meetings-Alternative dispute resolution-Single joint expert. Engineer in the court room	2
5.2	Criminal Cases-Introduction-Counterfeit coins-fraudulent road accidents-Fraudulent Insurance claims.	2
5.3	Cyber Crimes and Cases- SIM Swapping -ATM Cloning-Microsoft Internal Spam- Intellectual property cases.	2

Reference Books

1. Colin R Gagg, *Forensic EngineeringThe Art &Craft of a failure detective* , Taylor & Francis Publishing, 2020
2. Luca Fiorentini ,Luca Marmo *Principles of Forensic Engineering Applied to Industrial Accidents* , Wiley, 2019
3. Harold Franck, Darren Franck , *Forensic Engineering Fundamentals* ,Taylor & Francis publishing 2013
4. Randall K Noon , *Forensic Engineering Investigation*, CRC press limited , 2001
5. Stephen E Petty , *Forensic Engineering: Damage assessment for residential and commercial structures* CRC press 2nd edition , 2017
6. Joshua B Kardon , *Guideliness for forensic Engineering practice* , ASCE, 2012
7. Richard W. Mclay and Robert N. Anderson, *Engineering standards for forensic*

Applications , Academic Press; 1st edition 2018

8. Max M Houck ,*Forensic Engineering (Advanced forensic Science),* Academic press 1st edition 2017
9. Niranjan Reddy - Practical Cyber Forensics. *An Incident-based Approach to Forensic Investigations-Apress (2019)*
10. Peter Rhys Lewis, Ken Reynolds, Colin Gagg - *Forensic Materials Engineering Case Studies-CRC Press (2003) (1)*

INTERNSHIP

A student shall opt for carrying out the Internship at an Industry/Research Organization or at another institute of higher learning and repute (Academia). The organization for Internship shall be selected/decided by the students on their own with prior approval from the faculty advisor/respective PG Programme Coordinator/Guide/Supervisor. Every student shall be assigned an internship Supervisor/Guide at the beginning of the Internship. The training shall be related to their specialisation after the second semester for a minimum duration of six to eight weeks. On completion of the course, the student is expected to be able to develop skills in facing and solving the problems experiencing in the related field.

Objectives

- Exposure to the industrial environment, which cannot be simulated in the classroom and hence creating competent professionals for the industry.
- Provide possible opportunities to learn understand and sharpen the real time technical / managerial skills required at the job.
- Exposure to the current technological developments relevant to the subject area of training.
- Create conducive conditions with quest for knowledge and its applicability on the job.
- Understand the social, environmental, economic and administrative considerations that influence the working environment.
- Expose students to the engineer's responsibilities and ethics.

Benefits of Internship Benefits to Students

- An opportunity to get hired by the Industry/ organization.
- Practical experience in an organizational setting & Industry environment.
- Excellent opportunity to see how the theoretical aspects learned in classes are integrated into the practical world. On-floor experience provides much more professional experience which is often worth more than classroom

teaching.

- Helps them decide if the industry and the profession is the best career option to pursue.
- Opportunity to learn new skills and supplement knowledge.
- Opportunity to practice communication and teamwork skills.
- Opportunity to learn strategies like time management, multi-tasking etc in an industrial setup.
- Makes a valuable addition to their resume.
- Enhances their candidacy for higher education/placement.
- Creating network and social circle and developing relationships with industry people.
- Provides opportunity to evaluate the organization before committing to a full time position.

Benefits to the Institute

- Build industry academia relations.
- Makes the placement process easier.
- Improve institutional credibility & branding.
- Helps in retention of the students.
- Curriculum revision can be made based on feedback from Industry/ students.
- Improvement in teaching learning process.

Benefits to the Industry

- Availability of ready to contribute candidates for employment.
- Year round source of highly motivated pre-professionals.
- Students bring new perspectives to problem solving.
- Visibility of the organization is increased on campus.

- Quality candidate’s availability for temporary or seasonal positions and projects.
- Freedom for industrial staff to pursue more creative projects.
- Availability of flexible, cost-effective workforce not requiring a long-term employer commitment.
- Proven, cost-effective way to recruit and evaluate potential employees.
- Enhancement of employer’s image in the community by contributing to the educational enterprise.

Types of Internships

- Industry Internship with/without Stipend
- Govt / PSU Internship (BARC/Railway/ISRO etc)
- Internship with prominent education/research Institutes
- Internship with Incubation centres /Start-ups

Guidelines

- All the students need to go for internship for minimum duration of 6 to 8 weeks.
- Students can take mini projects, assignments, case studies by discussing it with concerned authority from industry and can work on it during internship.
- All students should compulsorily follow the rules and regulations as laid by industry.
- Every student should take prior permissions from concerned industrial authority if they want to use any drawings, photographs or any other document from industry.
- Student should follow all ethical practices and SOP of industry.
- Students have to take necessary health and safety precautions as laid by the industry.
- Student should contact his /her Guide/Supervisor from college on weekly basis to communicate the progress.
- Each student has to maintain a diary/log book
- After completion of internship, students are required to submit
  - Report of work done
  - Internship certificate copy
  - Feedback from employer / internship mentor
  - Stipend proof (in case of paid internship).

Total Marks 100: The marks awarded for the Internship will be on the basis of (i) Evaluation done by the Industry (ii) Students diary (iii) Internship Report and (iv) Comprehensive Viva Voce.

Continuous Internal Evaluation: 50 marks

Student's diary - 25 Marks Evaluation done by the industry -

25 Marks

Student's Diary/ Daily Log: The main purpose of writing daily diary is to cultivatethe habit of documenting and to encourage the students to search for details. It develops the students' thought process and reasoning abilities. The students should record in the daily training diary the day to day account of the observations, impressions, information gathered and suggestions given, if any. It should contain the sketches & drawings related to the observations made by the students. The daily training diary should be signed after every day by the supervisor/ in charge of the section where the student has been working. The diary should also be shown to the Faculty Mentor visiting the industry from time to time and got ratified on the day of his visit. Student's diary will be evaluated on the basis of the following criteria:

- Regularity in maintenance of the diary
- Adequacy & quality of information recorded
- Drawings, design, sketches and data recorded
- Thought process and recording techniques used
- Organization of the information.

The format of student's diary

Name of the Organization/Section:

Name and Address of the Section Head:

Name and Address of the Supervisor:

Name and address of the student:

Internship Duration: From .... To ....

Brief description about the nature of internship:

Day	Brief write up about the Activities carried out: Such as design, sketches, result observed, issues identified, data recorded, etc.
1	
2	
3	

Signature of Industry Supervisor   *Signature of Section Head/HR Manager Office Seal*

Attendance Sheet

Name of the Organization/Section:

Name and Address of the Section Head:

Name and Address of the Supervisor:

Name and address of the student:

Internship Duration: From .... To ....

Month & Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	...
Month & Year																					
Month & Year																					

Signature of Industry Supervisor *Signature of Section Head/HR Manager Office Seal*

Note:

- Student's Diary shall be submitted by the students along with attendance record and an evaluation sheet duly signed and stamped by the industry to the Institute immediately after the completion of the training.
- Attendance Sheet should remain affixed in daily training diary. Do not remove or tear it off.
- Student shall sign in the attendance column. Do not mark 'P'.
- Holidays should be marked in red ink in the attendance column. Absent should be marked as 'A' in red ink.

Evaluation done by the Industry (Marks 25)

Format for Supervisor Evaluation of Intern

Student Name :____________________Date:____________________ Supervisor Name :____________

Designation:______________________ Company/Organization :_____________

Internship Address: ____ Dates of Internship: From______To ________________________________

Please evaluate intern by indicating the frequency with which you observed the following parameters:

Parameters **Marks**	Needs improvement (0 – 0.25 mark)	Satisfactory (0.25 – 0.50 mark)	Good (0.75 mark)	Excellent (1 mark)
Behavior				
Performs in a dependable Manner				
Cooperates with coworkers and supervisor				
Shows interest in work				
Learns quickly				
Shows initiative				
Produces high quality work				
Accepts responsibility				
Accepts criticism				
Demonstrates organizational skills				
Uses technical knowledge and expertise				
Shows good judgment				
Demonstrates creativity/originality				
Analyzes problems effectively				
Is self-reliant				
Communicates well				
Writes effectively				
Has a professional attitude				
Gives a professional appearance				
Is punctual				
Uses time effectively				

Overall performance of student

Intern (Tick one) : Needs improvement (0 - 0.50 mark) / Satisfactory (0.50 – 1.0 mark) / Good (1.5 mark) / Excellent (2.0 mark)

Additional comments, if any (2 marks) :

Signature of Industry Supervisor *Signature of Section Head/HR Manager Office Seal*

End Semester Evaluation (External Evaluation): 50 Marks

Internship Report	-	25 Marks
Viva Voce	-	25 Marks

Internship Report: After completion of the internship, the student should prepare a comprehensive report to indicate what he has observed and learnt in the training period and should be submitted to the faculty Supervisor. The student may contact Industrial Supervisor/ Faculty Mentor for assigning special topics and problems and should prepare the final report on the assigned topics. Daily diary will also help to a great extent in writing the industrial report since much of the information has already been incorporated by the student into the daily diary. The training report should be signed by the Internship Supervisor, Programme Coordinator and Faculty Mentor.

The Internship report (25 Marks) will be evaluated on the basis of following criteria:

- Originality
- Adequacy and purposeful write-up
- Organization, format, drawings, sketches, style, language etc.
- Variety and relevance of learning experience
- Practical applications, relationships with basic theory and concepts taught in the course

Viva Voce (25 Marks) will be done by a committee comprising Faculty Supervisor, PG Programme Coordinator and an external expert (from Industry or research/academic Institute). This committee will be evaluating the internship report also.

RESEARCH PROJECT/DISSERTATION

Research Project: Students choosing track 2 shall carry out the research project in their parent Institution only under the guidance of a supervisor assigned by the DLAC.

Dissertation: All categories of students in track 1 are to carry out the dissertation in the Institute they are studying or can work either in any CSIR/Industrial R&D organization/any other reputed Institute which have facilities for dissertation work in the area proposed.

Mark Distribution:

Phase 1: Total marks: 100, only CIA

SEMESTER-IV SYLLABUS

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
244PBT100	DISSERTATION PHASE II	Project Work	0	0	24	16

All categories of students in track 1 are to carry out the DISSERTATION PHASE II in the institute they are studying or in any Industrial/ R&D organization/any other reputed institute which have facilities for dissertation work in the area proposed. DISSERTATION PHASE II shall not compulsorily continuation of DISSERTATION PHASE I. The student has to publish a research article in a conference or a reputed journal before appearing for the end-semester examination. The eligibility criteria for registering to the end semester examination are attendance in the course and no pending disciplinary action. The minimum attendance for appearing for the end semester examination is 75%. Students who do not meet these eligibility criteria are ineligible (identified by FE grade) to appear for the ESE. Students, who have completed a course but could not appear for the end semester examination, shall be awarded 'AB' Grade, provided they meet other eligibility criteria The pass minimum for the course is 45% for ESE and 50% for (CIA and ESE) put together.

Continuous Internal Assessment (CIA) Total Marks: 100

The evaluation committee comprises 1- Project Coordinator(s) 2- A Senior faculty member 3- Supervisor of the student

Pattern (CIA)

Zeroth evaluation by the Evaluation Committee	-
Interim evaluation by the Evaluation Committee	30 marks
Final evaluation by the Evaluation Committee	50 marks
Project progress evaluation by supervisor	20 marks

Evaluation by the supervisor

The guide/supervisor shall monitor the progress being carried out by the student on a regular basis. In case it is found that progress is unsatisfactory it shall be reported to the Department Evaluation Committee for necessary action.

Student's Diary/ Log book: The main purpose of writing diary/log book is to cultivate the habit of documenting and to encourage the students to search for details. The activity diary shall be signed after every week by the supervisor.

End Semester Evaluation (ESE) Total Marks: 100

The evaluation committee
comprises 1- Project
Coordinator(s)
2- An external expert (from Industry or research/academic
institute) 3- Supervisor of the student

Pattern (ESE)

1. **Innovation and Originality (10 marks):**
Assessment of the uniqueness and innovation demonstrated in the project work. Original contributions, if any, to the field or problem area.
2. **Implementation and Execution (20 marks):**
Evaluation of the actual implementation or execution of the project,
including: Quality of work done
Demonstrated skills and techniques
applied Adherence to project timelines
and milestones
3. **Project Documentation (25 marks):**
Comprehensive project report evaluation
including: Introduction and problem
statement
Literature review
Methodology and
approach Results and
analysis
Conclusion and
recommendations
References and citations
Details of the
publications
Plagiarism
certificate

The Plagiarism level in the project report shall be less than 25%.

4. **Presentation and Defence (40 marks):**
Oral presentation of the project to a panel of examiners,
including: Clarity and effectiveness of the
presentation
Ability to explain the project objectives, methodologies, and
findings Handling questions and providing satisfactory answers
during the defence

5. **Publication of the work either in a conference or in a journal (5 marks)**

6. SYLLABUS:

Dissertation outside the Institute: For doing dissertation outside the Institution, the following conditions are to be met:

i. They have completed successfully the course work prescribed in the approved curriculum up to the second semester.

ii. The student has to get prior approval from the DLAC and CLAC.

iii. Facilities required for doing the dissertation shall be available in the Organization/Industry (A certificate stating the facilities available in the proposed organization and the time period for which the facilities shall be made available to the student, issued by a competent authority from the Organization/Industry shall be submitted by the student along with the application).

DETAIL S	HOURS
` 1. Literature study/survey of published literature on the assigned topic 2. Topic Selection and Proposal 3. Formulation of objectives 4. Research and Planning 5. Formulation of work plan and task allocation. 6. Execution 7. Documentation and Reporting 8. Project Showcase reflecting on the project experience and lessons learned	200

iv. They should have an external as well as an internal supervisor. The internal supervisor should belong to the parent institution and the external supervisor should be Scientists or Engineers from the Institution/Industry/ R&D organization with which the student is associated for doing the dissertation work. The external supervisor shall be with a minimum post graduate degree in the related area.

v. The student has to furnish his /her monthly progress as well as attendance report signed by the external supervisor and submit the same to the concerned internal supervisor.

vi. The external supervisor is to be preferably present during all the stages of evaluation of the dissertation.