SEMESTER VI

CST 302	COMPILER DESIGN	Category	L	T	P	Credit	Year of Introduction
		PCC	3	1	0	4	2019

Preamble:

The purpose of this course is to create awareness among students about the phases of a compiler and the techniques for designing a compiler. This course covers the fundamental concepts of different phases of compilation such as lexical analysis, syntax analysis, semantic analysis, intermediate code generation, code optimization and code generation. Students can apply this knowledge in design and development of compilers.

Prerequisite: Sound knowledge in Data Structures, Formal Languages & Automata Theory.

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

CO1	Explain the phases in compilation process(lexical analysis, syntax analysis, semantic analysis, intermediate code generation, code optimization and code generation) and model a lexical analyzer **(Cognitive Knowledge Level: Apply)**
CO2	Model language syntax using Context Free Grammar and develop parse tree representation using leftmost and rightmost derivations (**Cognitive Knowledge Level: Apply)**
CO3	Compare different types of parsers(Bottom-up and Top-down) and construct parser for a given grammar **(Cognitive Knowledge Level: Apply)**
CO4	Build Syntax Directed Translation for a context free grammar, compare various storage allocation strategies and classify intermediate representations **(Cognitive Knowledge Level: Apply)**
CO5	Illustrate code optimization and code generation techniques in compilation **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓	✓							✓
CO2	✓	✓	✓	✓	✓							✓
CO3	✓	✓	✓	✓	✓							✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination Marks**
	Test 1 (Marks)	**Test 2 (Marks)**	
Remember	**20**	**20**	**20**
Understand	**40**	**40**	**40**
Apply	**40**	**40**	**40**
Analyze			

Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : **10 marks**

Continuous Assessment - Test : **25 marks**

Continuous Assessment - Assignment : **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module - 1 (Introduction to compilers and lexical analysis)

Analysis of the source program - Analysis and synthesis phases, Phases of a compiler. Compiler writing tools. Bootstrapping. Lexical Analysis - Role of Lexical Analyser, Input Buffering, Specification of Tokens, Recognition of Tokens.

Module - 2 (Introduction to Syntax Analysis)

Role of the Syntax Analyser – Syntax error handling. Review of Context Free Grammars - Derivation and Parse Trees, Eliminating Ambiguity. Basic parsing approaches - Eliminating left recursion, left factoring. Top-Down Parsing - Recursive Descent parsing, Predictive Parsing, LL(1) Grammars.

Module - 3 (Bottom-Up Parsing)

Handle Pruning. Shift Reduce parsing. Operator precedence parsing (Concept only). LR parsing - Constructing SLR, LALR and canonical LR parsing tables.

Module - 4 (Syntax directed translation and Intermediate code generation)

Syntax directed translation - Syntax directed definitions, S-attributed definitions, L-attributed definitions, Bottom-up evaluation of S-attributed definitions. Run-Time Environments - Source Language issues, Storage organization, Storage-allocation strategies. Intermediate Code Generation - Intermediate languages, Graphical representations, Three-Address code, Quadruples, Triples.

Module 5 – (Code Optimization and Generation)

Code Optimization - Principal sources of optimization, Machine dependent and machine independent optimizations, Local and global optimizations. Code generation - Issues in the design of a code generator, Target Language, A simple code generator.

Text Books

1. Aho A.V., Ravi Sethi and D. Ullman. Compilers – Principles Techniques and Tools, Addison Wesley, 2006.

Reference Books

1. D.M.Dhamdhere, System Programming and Operating Systems, Tata McGraw Hill & Company, 1996.
2. Kenneth C. Louden, Compiler Construction – Principles and Practice, Cengage Learning Indian Edition, 2006.

3. Tremblay and Sorenson, The Theory and Practice of Compiler Writing, Tata McGraw Hill & Company,1984.

Sample Course Level Assessment Questions

Course Outcome 1 (CO1):

1) Explain the phases of a compiler with a neat diagram.

2) Define a token. Identify the tokens in the expression a := b + 10.

Course Outcome 2 (CO2):

1) Illustrate the process of eliminating ambiguity, left recursion and left factoring the grammar.

2) Is the following grammar ambiguous? If so eliminate ambiguity.

E→E + E | E*E |(E) | id

Course Outcome 3 (CO3):

1. What are the different parsing conflicts in the SLR parsing table?
2. Design a recursive descent parser for the grammar

   E→E + T | T

   T→T*F | F

   F→(E) | id
3. Construct canonical LR(0) collection of items for the grammar below.

   S → L = R

   S → R

   L → * R

   L → id

   R → L

   Also identify a shift reduce conflict in the LR(0) collection constructed above.

Course Outcome 4 (CO4):

1. Write the quadruple and triple representation of the following intermediate code

   R1 = C * D

   R2 = B + R1

   A = R2

   B[0] = A

2. Differentiate S-attributed Syntax Directed Translation(SDT) and L-attributed SDT. Write S - attributed SDT for a simple desktop calculator

Course Outcome 5 (CO5):

1. List out the examples of function preserving transformations.
2. What are the actions performed by a simple code generator for a typical three-address statement of the form x: = y op z.

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION , MONTH & YEAR

Course Code: CST 302

Course Name: Compiler Design

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Specify the analysis and synthesis parts of compilation.
2. Define the terms token, lexemes and patterns with examples.
3. Is the grammar S --> S | (S) S / Ɛ ambiguous? Justify your answer.
4. What is left recursive grammar? Give an example. What are the steps in removing left recursion?
5. Compare different bottom-up parsing techniques.
6. What are the possible actions of a shift reduce parser.

7. Differentiate synthesized and inherited attributes with examples.

8. Translate a[i] = b * c – b * d, to quadruple.

9. What is the role of peephole optimization in the compilation process

10. What are the issues in the design of a code generator

(10x3=30)

Part B

Answer any one question from each module. Each question carries 14 Marks)

11. (a) Explain the different phases of a compiler with a running example. **(9)**

(b) List and explain any three compiler construction tools. **(5)**

OR

12. (a) What is a regular definition? Give the regular definition of an unsigned integer **(7)**

(b) Express the role of transition diagrams in recognition of tokens. **(7)**

13. (a) What is Recursive Descent parsing? List the challenges in designing such a parser? **(4)**

(b) Consider the following grammar **(10)**

E-→E or T | T

T→ T and F |F

F→not F | (E) | true | false

(i) Remove left recursion from the grammar.

(ii) Construct a predictive parsing table.

(iii) Justify the statement "The grammar is LL (1)".

OR

14. (a) What is Recursive Descent parsing? List the problems in designing such a parser **(4)**

(b) Design a recursive descent parser for the grammar S→cAd, A→ab/ b **(5)**

Find the FIRST and FOLLOW of the non-terminals S, A and B in the grammar **(5)**

S→aABe
A→Abc | b
B→d

15. (a) Construct the LR(0) set of items and their GOTO function for the grammar **(10)**
S→S S + | S S * | a

(b) Is the grammar SLR? Justify your answer **(4)**

OR

16. (a) Identify LR(1) items for the grammar
S→ CC **(7)**
C→ cC | d

(b) Construct LALR table for the above grammar **(7)**

17. (a) Design a Syntax Directed Translator(SDT) for the arithmetic expression (4 * 7 + 19) * 2 and draw an annotated parse tree for the same. **(8)**

(b) Consider the grammar with following translation rules and E as the start symbol **(6)**

E → E1 # T {E.value=E1.value x T.value ;}
| T{E.value=T.value ;}
T → T1 & F{ T.value=T1.value + F.value ;}
| F{T.value= F.value ; }
F → num { F.value=num. lvalue ;}

Compute E.value for the root of the parse tree for the expression
2#3 & 5# 6 &7

OR

18. (a) Write Syntax Directed Translator (SDT) and parse tree for infix to postfix translation of an expression. **(8)**

    (b) Explain the storage allocation strategies. **(6)**

19. (a) Describe the principal sources of optimization **(7)**

    (b) Illustrate the optimization of basic blocks with examples. **(7)**

OR

20. (a) Write the Code Generation Algorithm and explain the *getreg* function **(6)**

    (b) Generate target code sequence for the following statement **(8)**

d := (a-b)+(a-c)+(a-c).

Teaching Plan

No	Contents	No. of Lecture Hours
Module - 1(Introduction to Compilers and lexical analyzer) (8 hours)		
1.1	Introduction to compilers, Analysis of the source program	1 hour
1.2	Phases of the compiler – Analysis Phases	1 hour
1.3	Phases of the Compiler - Synthesis Phases	1 hour
1.4	Symbol Table Manager and Error Handler	1 hour
1.5	Compiler writing tools, bootstrapping	1 hour
1.6	The role of Lexical Analyzer , Input Buffering	1 hour
1.7	Specification of Tokens	1 hour
1.8	Recognition of Tokens	1 hour

Module – 2 (Introduction to Syntax Analysis) (10 hours)		
2.1	Role of the Syntax Analyser, Syntax error handling	1 hour
2.2	Review of Context Free Grammars	1 hour
2.3	Parse Trees and Derivations	1 hour
2.4	Grammar transformations, Eliminating ambiguity	1 hour
2.5	Eliminating left recursion	1 hour
2.6	Left factoring the grammar	1 hour
2.7	Recursive Descent parsing	1 hour
2.8	First and Follow	1 hour
2.9	Predictive Parsing table constructor	1 hour
2.10	LL(1) Grammars	1 hour
Module - 3 (Bottom up parsing) (9 hours)		
3.1	Bottom-up parsing - Handle Pruning	1 hour
3.2	Shift Reduce parsing	1 hour
3.3	Operator precedence parsing (Concept only)	1 hour
3.4	LR parsing , SLR Grammar, items	1 hour
3.5	Augmented Grammar, Canonical collection of LR(0) items	1 hour
3.6	SLR Parser Table Construction	1 hour
3.7	Constructing Canonical LR Parsing Tables	1 hour
3.8	Constructing LALR Parsing Tables	1 hour
3.9	LALR parser	1 hour
Module - 4 (Syntax Directed Translation and Intermediate code Generation) (9 hours)		
4.1	Syntax directed definitions	1 hour
4.2	S- attributed definitions, L- attributed definitions	1 hour
4.3	Bottom- up evaluation of S- attributed definitions.	1 hour
4.4	Source Language issues	1 hour
4.5	Storage organization	1 hour

4.6	Storage- allocation strategies	1 hour
4.7	Intermediate languages , Graphical representations	1 hour
4.8	Three-Address code	1 hour
4.9	Quadruples, Triples	1 hour
Module - 5 (Code Optimization and Generation) (9 hours)		
5.1	Principal sources of optimization	1 hour
5.2	Machine dependent optimizations	1 hour
5.3	Machine independent optimizations	1 hour
5.4	Local optimizations	1 hour
5.5	Global optimizations	1 hour
5.6	Issues in the design of a code generator – Lecture 1	1 hour
5.7	Issues in the design of a code generator – Lecture 2	1 hour
5.8	Target Language	1 hour
5.9	Design of a simple code generator.	1 hour

CST 304	COMPUTER GRAPHICS AND IMAGE PROCESSING	Category	L	T	P	Credit	Year of Introduction
		PCC	3	1	0	4	2019

Preamble:
The purpose of this course is to make awareness about strong theoretical relationships between computer graphics and image processing. This course helps the learner to understand three-dimensional environment representation in a computer, transformation of 2D/3D objects, basic mathematical techniques and algorithms used to build useful applications, imaging, and image processing techniques. The study of computer graphics and image processing develops the ability to create image processing frameworks for different domains and develops algorithms for emerging display technologies.

Prerequisite: A sound knowledge of Mathematics and a programming language.

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

CO#	CO
CO1	Describe the working principles of graphics devices**(Cognitive Knowledge level: Understand)**
CO2	Illustrate line drawing, circle drawing and polygon filling algorithms(Cognitive Knowledge level: Apply)
CO3	Demonstrate geometric representations, transformations on 2D & 3D objects, clipping algorithms and projection algorithms(Cognitive Knowledge level: Apply)
CO4	Summarize visible surface detection methods(Cognitive Knowledge level: Understand)
CO5	Summarize the concepts of digital image representation, processing and demonstrate pixel relationships(Cognitive Knowledge level: Apply)
CO6	Solve image enhancement and segmentation problems using spatial domain techniques(Cognitive Knowledge level: Apply)

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓											✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓								✓
CO4	✓		✓									✓
CO5	✓	✓	✓	✓								
CO6	✓	✓	✓	✓		✓						✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	**30**	**30**	**30**
Understand	**30**	**30**	**30**

Apply	**40**	**40**	**40**
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance **10 marks**

Continuous Assessment Tests (Average of SeriesTests1&2) **25 marks**

Continuous AssessmentAssignment **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one full question. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module – 1(Basics of Computer graphics and Algorithms)

Basics of Computer Graphics and its applications. Video Display devices- Refresh Cathode Ray Tubes, Random Scan Displays and systems, Raster scan displays and systems. Line drawing algorithms- DDA, Bresenham's algorithm. Circle drawing algorithms- Midpoint Circle generation algorithm, Bresenham's algorithm.

Module - 2(Filled Area Primitives and transformations)

Filled Area Primitives- Scan line polygon filling, Boundary filling and flood filling. Two dimensional transformations-Translation, Rotation, Scaling, Reflection and Shearing, Composite transformations, Matrix representations and homogeneous coordinates. Basic 3D transformations.

Module - 3 (Clipping and Projections)

Window to viewport transformation. Cohen Sutherland Line clipping algorithm. Sutherland Hodgeman Polygon clipping algorithm. Three dimensional viewing pipeline. Projections- Parallel and Perspective projections. Visible surface detection algorithms- Depth buffer algorithm, Scan line algorithm.

Module - 4 (Fundamentals of Digital Image Processing)

Introduction to Image processing and applications. Image as 2D data. Image representation in Gray scale, Binary and Colour images. Fundamental steps in image processing. Components of image processing system. Coordinate conventions. Sampling and quantization. Spatial and Gray Level Resolution. Basic relationship between pixels– neighbourhood, adjacency, connectivity. Fundamentals of spatial domain-convolution operation.

Module - 5 (Image Enhancement in Spatial Domain and Image Segmentation)

Basic gray level transformation functions - Log transformations, Power-Law transformations, Contrast stretching. Histogram equalization. Basics of spatial filtering - Smoothing spatial filter-Linear and nonlinear filters, and Sharpening spatial filters-Gradient and Laplacian.

Fundamentals of Image Segmentation. Thresholding - Basics of Intensity thresholding and Global Thresholding. Region based Approach - Region Growing, Region Splitting and Merging. Edge Detection - Edge Operators- Sobel and Prewitt.

Text Book

1. Donald Hearn and M. Pauline Baker, Computer Graphics, PHI, 2e, 1996
2. Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing. Pearson, 4e, 2017

References

1) William M. Newman and Robert F. Sproull, Principles of Interactive Computer Graphics. McGraw Hill, 2001

2) Zhigang Xiang and Roy Plastock, Computer Graphics (Schaum's outline Series), McGraw Hill, 2019.
3) David F. Rogers , Procedural Elements for Computer Graphics, Tata McGraw Hill,2001.
4) M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis, and Machine Vision, Thomson India Edition, 4e, 2017.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Compare the working principle of raster scan systems and random scan systems.
2. How much time is spent scanning across each row of pixels during screen refresh on a raster system with resolution of 1280*1024 and a refresh rate of 60 frames per second?

Course Outcome 2 (CO2):

1. Rasterize the line with end points(2,3) and (5,8) using Bresenham's line drawing algorithm.
2. Explain how the 4-connected area filling approach differs from 8- connected area filling in boundary filling algorithm

Course Outcome 3 (CO3):

1. Rotate a triangle ABC 45 degree counter clockwise about the pivot point (10,3), where the position vector of the coordinate ABC is given as A(4,1), B(5,2) and C(4,3).
2. Given a clipping window A(20,20), B(60,20), C(60,40) and D(20,40). Using Cohen Sutherland algorithm, find the visible portion of the line segment joining the points P(40,80) and Q(120,30)

Course Outcome 4 (CO4):

1. Explain scan line algorithm for detecting visible surfaces in an object.

Course Outcome 5 (CO5):

1. Give an image representation model and describe how the representation changes in grayscale, binary and colour images.
2. Consider an image segment shown below.

3 1 2 1 (q)

2 2 0 2

1 2 1 1

(p) 1 0 1 2

(a) Let V={0,1} and compute the length of the shortest 4-,8- and m- path between p and q. If a particular path does not exist between these two points , explain why?

(b) Repeat for V={1,2}.

3. The spatial resolution of an image is given by 128 X 128.What is its storage requirements if it is represented by 64 gray levels?

Course Outcome 6 (CO6):

1. A skilled medical technician is charged with the job of inspecting a certain class of monochrome images generated by electronic microscope. To facilitate the inspection, the technician uses image processing aids. However when he examines the images he finds the following problems.
   (a) Presence of bright isolated dots that are not of interest.
   (b) Lack of sharpness
   (c) Poor contrast

   Identify the sequence of preprocessing steps that the technician may use to overcome the above mentioned problems and explain it.

2. A 4x4, 4 bits/pixel original image is given by

10	12	8	9
10	12	12	14
12	13	10	9
14	12	10	12

   (a) Apply histogram equalisation to the image by rounding the resulting image pixels to integers
   (b) Sketch the histogram of the original image and the histogram-equalised image.

3. You have Sobel operator and Laplacian operator for edge detection. Which operator will you select for edge detection in the case of noisy image? Explain.**(Assignment)**

Model Question Paper

QP CODE:

Reg No: _____

Name:_____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST 304

Course Name: Computer Graphics and Image Processing

Max. Marks : 100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Justify the approach of using integer arithmetic in Bresenham's line drawing algorithm.

2. Consider a raster system with a resolution of 1024*1024. What is the size of the raster needed to store 4 bits per pixel? How much storage is needed if 8 bits per pixel are to be stored?

3. Show that two successive reflections about either of the coordinate axes is equivalent to a single rotation about the coordinate origin.

4. Determine a sequence of basic transformations that are equivalent to the x-direction shearing matrix.

5. Find the window to viewport normalization transformation with window lower left corner at (1,1) and upper right corner at (2,6).

6. Find the orthographic projection of a unit cube onto the x=0, y=0 and z=0 plane.

7. Define Sampling and Quantization of an image.

8. Give any three applications of digital image processing.

9. A captured image appears very dark because of wrong lens aperture setting. Describe an enhancement technique which is appropriate to enhance such an image.

10. Suggest an approach of thresholding that should be used in case of uniform illumination. **(10x3=30)**

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Write Midpoint circle drawing algorithm and use it to plot a circle with radius=20 and center is (50,30). **(10)**

(b) Draw the architecture of raster scan display systems and explain its working principle. **(4)**

OR

12. (a) Derive the initial decision parameter of Bresenham's line drawing algorithm and use the algorithm to rasterize a line with endpoints (2,2) and (10,10). **(10)**

(b) Explain the working principle of color CRT monitors with suitable illustrations. **(4)**

13. (a) Compare boundary fill algorithm and flood fill algorithm. **(5)**

(b) Reflect a triangle ABC about the line 3x-4y+8=0. The position vector of the coordinate ABC is given as A(4,1), B(5,2) and C(4,3). **(9)**

OR

14. (a) Explain the need of using vanishing points in projections. **(4)**

(b) Explain Cohen-Sutherland line clipping algorithm. Use the algorithm to clip line P1(70, 20) and P2(100,10) against a window lower left hand corner (50,10) and upper right hand corner (80,40). **(10)**

15. (a) Describe Sutherland Hodegman polygon clipping algorithm and what are its **(7)**

limitations.

(b) Explain how visible surfaces can be detected using depth buffer algorithm. **(7)**

OR

16. (a) Describe Sutherland Hodegman polygon clipping algorithm and what are its limitations. **(7)**

(b) Explain how visible surfaces can be detected using depth buffer algorithm. **(7)**

17. (a) Explain the components of an image processing system with suitable diagram **(9)**

(b) Define Resolution of an image. Explain the spatial and gray level resolution of an image with an example. **(5)**

OR

18. (a) Define 4-adjacency, 8 adjacency and m-adjacency. Consider the image segment shown. **(7)**

$$\begin{array}{cccc} 4 & 2 & 3 & 2 \ (q) \\ 3 & 3 & 1 & 3 \\ 2 & 3 & 2 & 2 \\ 2 & 1 & 2 & 3 \ (p) \end{array}$$

Let V={1,2} and compute the length of the shortest 4- ,8- and m- path between p and q. If a particular path does not exist between these two points, explain why?

(b) Using any one application, explain the steps involved in image processing. **(7)**

19. (a) A 5x5 image patch is shown below. Compute the value of the marked pixel if it is smoothened by a 3x3 average filter and median filter. **(4)**

$$f(m,n)=\begin{pmatrix} 0 & 1 & 2 & 3 & 2 \\ 5 & 6 & 7 & 8 & 4 \\ 4 & 3 & \textcircled{2} & 1 & 2 \\ 8 & 7 & 6 & 5 & 3 \\ 1 & 5 & 3 & 7 & 6 \end{pmatrix}$$

(b) Define Image segmentation and describe in detail method of edge and region based segmentation technique. **(10)**

OR

20. (a) Distinguish between smoothing and sharpening filters in terms of **(10)**

(i) Functionality
(ii) Types
(iii) Applications
(iv) Mask Coefficients

(b) Describe how an image is segmented using split and merge technique in association with the region adjacency graph. **(8)**

Teaching Plan

No	Contents	No of Lecture Hrs (45 hrs)
Module – 1 (Basics of Computer Graphics and Algorithms) (9 hrs)		
1.1	Basics of Computer Graphics and applications	1 hour
1.2	Refresh Cathode Ray Tubes	1 hour
1.3	Random Scan Displays and systems	1 hour
1.4	Raster scan displays and systems	1 hour
1.5	DDA Line drawing Algorithm	1 hour
1.6	Bresenham's line drawing algorithm	1 hour
1.7	Midpoint Circle generation algorithm	1 hour
1.8	Bresenham's Circle generation algorithm	1 hour
1.9	Illustration of line drawing and circle drawing algorithms	1 hour
Module - 2 (Filled Area Primitives and transformations) (9 hrs)		
2.1	Scan line polygon filling	1 hour
2.2	Boundary filling and flood filling	1 hour
2.3	Basic 2D transformations-Translation	1 hour

2.4	Basic 2D transformations- Rotation and Scaling	1 hour
2.5	Reflection and Shearing	1 hour
2.6	Composite transformations	1 hour
2.7	Matrix representations and homogeneous coordinates	1 hour
2.8	Basic 3D transformation-Translation and scaling	1 hour
2.9	Basic 3D transformation-Rotation	1 hour
Module - 3 (Clipping and Projections) (8 hrs)		
3.1	Window to viewport transformation	1 hour
3.2	Cohen Sutherland Line clipping algorithm	1 hour
3.3	Sutherland Hodgeman Polygon clipping algorithm	1 hour
3.4	Practice problems on Clipping algorithms	1 hour
3.5	Three dimensional viewing pipeline, Projections-Parallel projections	1 hour
3.6	Projections- Perspective projections	1 hour
3.7	Visible surface detection algorithms- Depth buffer algorithm	1 hour
3.8	Scan line visible surface detection algorithm	1 hour
Module - 4 (Fundamentals of Digital Image Processing) (8 hrs)		
4.1	Introduction to Image processing-Image as a 2D data, Image representation-Gray scale, Binary and Colour images.	1 hour
4.2	Fundamental steps in image processing and applications	1 hour
4.3	Components of image processing system	1 hour
4.4	Coordinate conventions, Sampling and quantization, Spatial and Gray Level Resolution	1 hour
4.5	Basic relationship between pixels – neighbourhood, adjacency, connectivity	1 hour
4.6	Illustration of basic relationship between pixels– neighbourhood,	1 hour

	adjacency, connectivity	
4.7	Fundamentals of spatial domain - Convolution operation	1 hour
4.8	Illustration of Convolution operation	1 hour
Module - 5 (Image Enhancement in spatial domain and Image Segmentation) (11 hrs)		
5.1	Basic gray level transformation functions- Log transformations.	1 hour
5.2	Power-Law transformations, Contrast stretching	1 hour
5.3	Histogram equalization	1 hour
5.4	Illustration of Histogram equalization	1 hour
5.5	Basics of spatial filtering, Smoothing spatial filter- Linear and nonlinear filters	1 hour
5.6	Sharpening spatial filtering-Gradient filter mask	1 hour
5.7	Sharpening spatial filtering-Laplacian filter mask	1 hour
5.8	Fundamentals of Image Segmentation, Basics of Intensity thresholding, Basic Global Thresholding	1 hour
5.9	Region Based Approach- Region Growing, Region Splitting and Merging	1 hour1
5.10	Basics of Edge Detection	1 hour
5.11	Sobel and Prewitt edge detection masks	1 hour

CST 306	ALGORITHM ANALYSIS AND DESIGN	Category	L	T	P	Credit	Year of Introduction
		PCC	3	1	0	4	2019

Preamble:

The course introduces students to the design of computer algorithms, as well as analysis of algorithms. Algorithm design and analysis provide the theoretical backbone of computer science and are a must in the daily work of the successful programmer. The goal of this course is to provide a solid background in the design and analysis of the major classes of algorithms. At the end of the course students will be able to develop their own versions for a given computational task and to compare and contrast their performance.

Prerequisite:

Strong Foundation in Mathematics, Programming in C, Data Structures and Graph Theory.

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

CO#	CO
CO1	Analyze any given algorithm and express its time and space complexities in asymptotic notations. **(Cognitive Level: Apply)**
CO2	Derive recurrence equations and solve it using Iteration, Recurrence Tree, Substitution and Master's Method to compute time complexity of algorithms. **(Cognitive Level: Apply)**
CO3	Illustrate Graph traversal algorithms & applications and Advanced Data structures like AVL trees and Disjoint set operations. **(Cognitive Level: Apply)**
CO4	Demonstrate Divide-and-conquer, Greedy Strategy, Dynamic programming, Branch-and Bound and Backtracking algorithm design techniques (**Cognitive Level: Apply)**
CO5	Classify a problem as computationally tractable or intractable, and discuss strategies to address intractability **(Cognitive Level: Understand)**
CO6	Identify the suitable design strategy to solve a given problem. **(Cognitive Level: Analyze)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓								✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓								✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓										√
CO6	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40

Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance **10 marks**

Continuous Assessment Tests (Average of SeriesTests1&2) **25 marks**

Continuous Assessment Assignment **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module-1 (Introduction to Algorithm Analysis)

Characteristics of Algorithms, Criteria for Analysing Algorithms, Time and Space Complexity - Best, Worst and Average Case Complexities, Asymptotic Notations - Big-Oh (O), Big- Omega (Ω), Big-Theta (Θ), Little-oh (o) and Little- Omega (ω) and their properties. Classifying functions by their asymptotic growth rate, Time and Space Complexity Calculation of simple algorithms.

Analysis of Recursive Algorithms: Recurrence Equations, Solving Recurrence Equations – Iteration Method, Recursion Tree Method, Substitution method and Master's Theorem (Proof not required).

Module–2 (Advanced Data Structures and Graph Algorithms)

Self Balancing Tree - AVL Trees (Insertion and deletion operations with all rotations in detail, algorithms not expected); Disjoint Sets- Disjoint set operations, Union and find algorithms.

DFS and BFS traversals - Analysis, Strongly Connected Components of a Directed graph, Topological Sorting.

Module–3 (Divide & Conquer and Greedy Strategy)

The Control Abstraction of Divide and Conquer- 2-way Merge sort, Strassen's Algorithm for Matrix Multiplication-Analysis. The Control Abstraction of Greedy Strategy- Fractional Knapsack Problem, Minimum Cost Spanning Tree Computation- Kruskal's Algorithms - Analysis, Single Source Shortest Path Algorithm - Dijkstra's Algorithm-Analysis.

Module-4 (Dynamic Programming, Back Tracking and Branch & Bound))

The Control Abstraction- The Optimality Principle- Matrix Chain Multiplication-Analysis, All Pairs Shortest Path Algorithm - Floyd-Warshall Algorithm-Analysis. The Control Abstraction of Back Tracking – The N Queen's Problem. Branch and Bound Algorithm for Travelling Salesman Problem.

Module-5 (Introduction to Complexity Theory)

Tractable and Intractable Problems, Complexity Classes – P, NP, NP- Hard and NP-Complete Classes- NP Completeness proof of Clique Problem and Vertex Cover Problem- Approximation algorithms- Bin Packing, Graph Coloring. Randomized Algorithms (Definitions of Monte Carlo and Las Vegas algorithms), Randomized version of Quick Sort algorithm with analysis.

Text Books

1. T.H.Cormen, C.E.Leiserson, R.L.Rivest, C. Stein, Introduction to Algorithms, 2nd Edition, Prentice-Hall India (2001)
2. Ellis Horowitz, Sartaj Sahni, Sanguthevar Rajasekaran, "Fundamentals of Computer Algorithms", 2nd Edition, Orient Longman Universities Press (2008)

3. Sara Baase and Allen Van Gelder —Computer Algorithms, Introduction to Design and Analysis, 3rd Edition, Pearson Education (2009)

Reference Books

1. Jon Kleinberg, Eva Tardos, "Algorithm Design", First Edition, Pearson (2005)
2. Robert Sedgewick, Kevin Wayne, "Algorithms",4th Edition Pearson (2011)
3. GIlles Brassard, Paul Brately, "Fundamentals of Algorithmics", Pearson (1996)
4. Steven S. Skiena, "The Algorithm Design Manual", 2nd Edition, Springer(2008)

Course Level Assessment Questions

Course Outcome 1 (CO1):

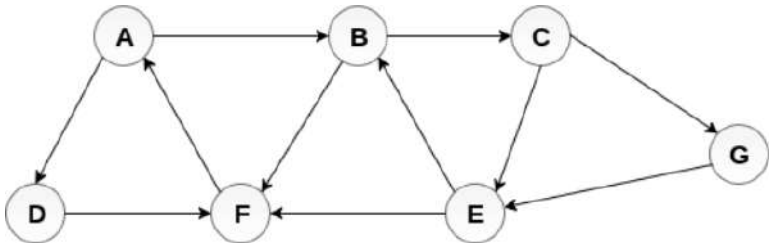
1. Is $2^{n+1} = O(2^n)$? Is $2^{2n} = O(2^n)$? Justify your answer.
2. What is the need of asymptotic analysis in calculating time complexity? What are the notations used for asymptotic analysis?
3. Calculate the time complexity for addition of two matrices.
4. Define time complexity and space complexity. Write an algorithm for adding n natural numbers and analyse the time and space requirements of the algorithm.

Course Outcome 2 (CO2):

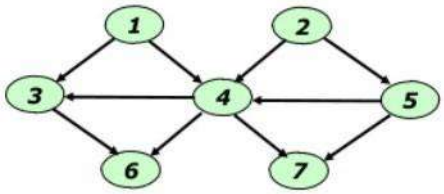
1. State Master's theorem for solving recurrences.
2. Solve the recurrence T(n) = 3T(n-2), using iteration method
3. State the conditions in recurrences where Master Theorem is not applicable.
4. Solve the following recurrence equations using Master's theorem.
   a) $T(n) = 8T(n/2) + 100\ n^2$
   b) $T(n) = 2T(n/2) + 10\ n$
5. Using Recursion Tree method, Solve T(n)= 2T(n/10)+ T(9n/10)+n. Assume constant time for small values of n.

Course Outcome 3 (CO3):

1. Explain the rotations performed for insertion in AVL tree with example.
2. Write down BFS algorithm and analyse the time complexity. Perform BFS traversal on the given graph starting from node A. If multiple node choices are available for next travel, choose the next node in alphabetical order.

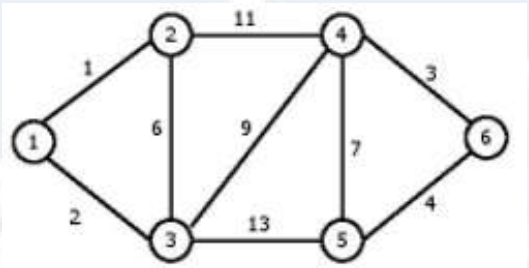


3. Find the minimum and maximum height of any AVL-tree with 7 nodes? Assume that the height of a tree with a single node is 0. (3)
4. Find any three topological orderings of the given graph.

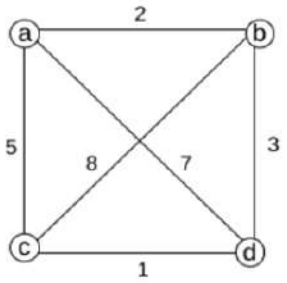


Course Outcome 4 (CO4):

1. Give the control abstraction for Divide and Conquer method.
2. Construct the minimum spanning tree for the given graph using Kruskal's algorithm. Analyse the complexity of the algorithm.



3. Compare Divide and Conquer and Dynamic programming methodologies
4. What is Principle of Optimality?
5. Define Travelling Salesman Problem (TSP). Apply branch and bound algorithm to solve TSP for the following graph, assuming the start city as 'a'. Draw the state space tree.



Course Outcome 5 (CO5):

1. Compare Tractable and Intractable Problems
2. With the help of suitable code sequence convince Vertex Cover Problem is an example of NP-Complete Problem

3. Explain Vertex Cover problem using an example. Suggest an algorithm for finding Vertex Cover of a graph.
4. Write short notes on approximation algorithms.
5. Compare Conventional quick sort algorithm and Randomized quicksort with the help of a suitable example?

Course Outcome 6 (CO6): (CO attainment through assignment only, not meant for examinations)

Choosing the best algorithm design strategy for a given problem after applying applicable design strategies – Sample Problems Given.

1. Finding the Smallest and Largest elements in an array of 'n' numbers
2. Fibonacci Sequence Generation.
3. Merge Sort
4. Travelling Sales Man Problem
5. 0/1 Knapsack Problem

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST 306

Course Name: Algorithm Analysis and Design

Max. Marks : 100 **Duration: 3 Hours**

PART A

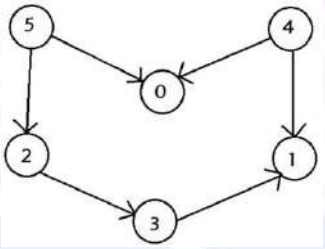
Answer All Questions. Each Question Carries 3 Marks

1. Define asymptotic notation? Arrange the following functions in increasing order of asymptotic growth rate.

   n^3, 2^n, $\log n^3$, 2^{100}, $n^2 \log n$, n^n, $\log n$, $n^{0.3}$, $2^{\log n}$

2. State Master's Theorem. Find the solution to the following recurrence equations using Master's theorem.
   a) $T(n) = 8T(n/2) + 100\ n^2$
   b) $T(n) = 2T(n/2) + 10\ n$

3. Find any two topological ordering of the DAG given below.



4. Show the UNION operation using linked list representation of disjoint sets.

5. Write the control abstraction of greedy strategy to solve a problem.

6. Write an algorithm based on divide-and-conquer strategy to search an element in a given list. Assume that the elements of list are in sorted order.

7. List the sequence of steps to be followed in Dynamic Programming approach.

8. Illustrate how optimal substructure property could be maintained in Floyd-Warshall algorithm.

9. Differentiate between P and NP problems.

10. Specify the relevance of approximation algorithms.

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Define Big O, Big Ω and Big Θ Notation and illustrate them graphically. **(7)**

    (b) Solve the following recurrence equation using recursion tree method **(7)**

    $T(n) = T(n/3) + T(2n/3) + n$, where $n>1$

    $T(n) = 1$, Otherwise

OR

12. (a) Explain the iteration method for solving recurrences and solve the following recurrence equation using iteration method. **(7)**

$$T(n) = 3T(n/3) + n; T(1) = 1$$

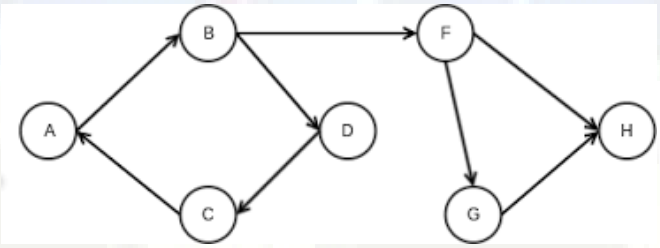
(b) Determine the time complexities of the following two functions fun1( ) and fun2( ). **(7)**

i)
```
int fun1(int n)
{
    if (n <= 1) return n;
    return 2*fun1(n-1);
}
```

ii)
```
int fun2 (int n)
{
    if (n <= 1) return n;
    return fun2 (n-1) + fun2 (n-1)
}
```

13. (a) Write DFS algorithm and analyse its time complexity. Illustrate the classification of edges in DFS traversal. **(7)**

(b) Find the strongly connected components of the digraph given below: **(7)**



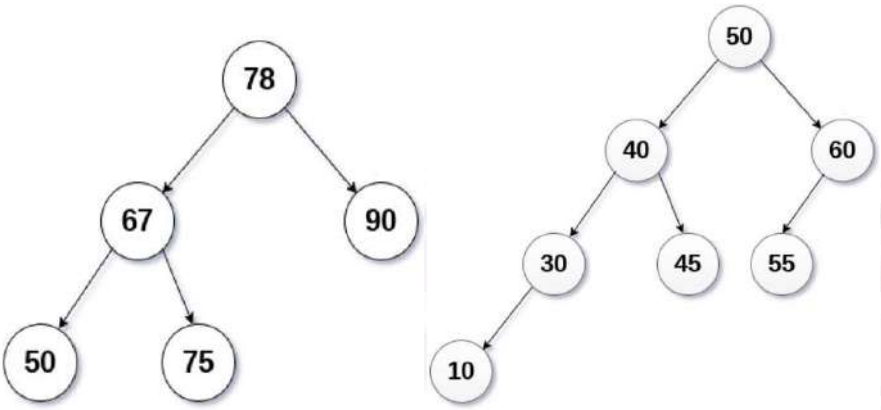
OR

14. (a) Illustrate the advantage of height balanced binary search trees over binary search trees? Explain various rotations in AVL trees with example. **(7)**

(b) Perform the following operations in the given AVL trees. **(7)**

i) Insert 70

ii) Delete 55



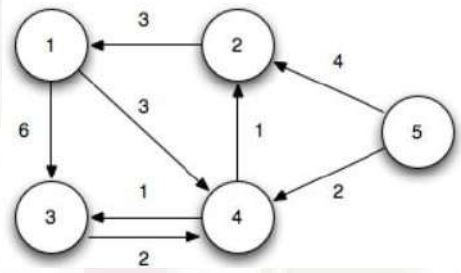
15. (a) State Fractional Knapsack Problem and write Greedy Algorithm for Fractional Knapsack Problem. **(7)**

(b) Find the optimal solution for the following Fractional Knapsack problem. Given the number of items(n) = 7, capacity of sack(m) = 15, W={2,3,5,7,1,4,1} and P = {10,5,15,7,6,18,3} **(7)**

OR

16. (a) Write and explain merge sort algorithm using divide and conquer strategy using the data {30, 19, 35, 3, 9, 46, 10}. Also analyse the time complexity. **(7)**

(b) Write the pseudo code for Dijkstra's algorithm. Compute the shortest distance from vertex 1 to all other vertices using Dijkstra's algorithm. **(7)**



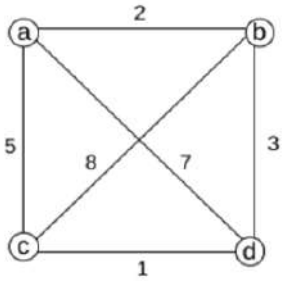
17. (a) Write Floyd-Warshall algorithm and analyse its complexity. **(5)**

(b) Write and explain the algorithm to find the optimal parenthesization of matrix chain product whose sequence of dimension is 4x10,10x3, 3x12,12x20. **(9)**

OR

18. (a) Explain the concept of Backtracking method using 4 Queens problem. **(7)**

(b) Define Travelling Salesman Problem (TSP). Apply branch and bound algorithm to solve TSP for the following graph, assuming the start city as 'a'. Draw the state space tree. **(7)**



19. (a) State bin packing problem? Explain the first fit decreasing strategy **(7)**

(b) Prove that the Clique problem is NP-Complete. **(7)**

OR

20. (a) Explain the need for randomized algorithms. Differentiate Las Vegas and Monte Carlo algorithms. **(6)**

(b) Explain randomized quicksort and analyse the expected running time of randomized quicksort with the help of a suitable example? **(9)**

Teaching Plan

No	Topic	No. of Hours (45 hrs)
Module -1 (Introduction to Algorithm Analysis) 9 hrs.		
1.1	Introduction to Algorithm Analysis: Characteristics of Algorithms.	1 hour
1.2	Criteria for Analysing Algorithms, Time and Space Complexity - Best, Worst and Average Case Complexities.	1 hour
1.3	Asymptotic Notations - Properties of Big-Oh (O), Big- Omega (Ω), Big-Theta (Θ), Little-Oh (o) and Little- Omega (ω).	1 hour
1.4	Illustration of Asymptotic Notations	1 hour

1.5	Classifying functions by their asymptotic growth rate	1 hour
1.6	Time and Space Complexity Calculation of algorithms/code segments.	1 hour
1.7	Analysis of Recursive Algorithms: Recurrence Equations, Solving Recurrence Equations – Iteration Method.	1 hour
1.8	Recursion Tree Method	1 hour
1.9	Substitution method and Master's Theorem and its Illustration.	1 hour
Module-2 (Advanced Data Structures and Graph Algorithms) 10 Hrs.		
2.1	Self Balancing Trees - Properties of AVL Trees, Rotations of AVL Trees	1 hour
2.2	AVL Trees Insertion and Illustration	1 hour
2.3	AVL Trees Deletion and Illustration	1 hour
2.4	Disjoint set operations.	1 hour
2.5	Union and find algorithms.	1 hour
2.6	Illustration of Union and find algorithms	1 hour
2.7	Graph Algorithms: BFS traversal, Analysis.	1 hour
2.8	DFS traversal, Analysis.	1 hour
2.9	Strongly connected components of a Directed graph.	1 hour
2.10	Topological Sorting.	1 hour
Module-3 (Divide & Conquer and Greedy Method) 8 Hrs		
3.1	Divide and Conquer: The Control Abstraction.	1 hour
3.2	2-way Merge Sort, Analysis.	1 hour
3.3	Strassen's Algorithm for Matrix Multiplication, Analysis	1 hour

3.4	Greedy Strategy: The Control Abstraction.	1 hour
3.5	Fractional Knapsack Problem.	1 hour
3.6	Minimum Cost Spanning Tree Computation- Kruskal's Algorithm, Analysis.	1 hour
3.7	Single Source Shortest Path Algorithm - Dijkstra's Algorithm	1 hour
3.8	Illustration of Dijkstra's Algorithm-Analysis.	1 hour
Module-4 (Dynamic Programming, Back Tracking and Branch and Bound) 8 Hrs.		
4.1	Dynamic Programming: The Control Abstraction, The Optimality Principle.	1 hour
4.2	Matrix Chain Multiplication-Analysis.	1 hour
4.3	Illustration of Matrix Chain Multiplication-Analysis.	1 hour
4.4	All Pairs Shortest Path Algorithm- Analysis and Illustration of Floyd-Warshall Algorithm.	1 hour
4.5	Back Tracking: The Control Abstraction .	1 hour
4.6	Back Tracking: The Control Abstraction – The N Queen's Problem.	1 hour
4.7	Branch and Bound:- Travelling salesman problem.	1 hour
4.8	Branch and Bound:- Travelling salesman problem.	1 hour
Module-5 (Introduction to Complexity Theory) 10 Hrs		
5.1	Introduction to Complexity Theory: Tractable and Intractable Problems.	1 hour
5.2	Complexity Classes – P, NP.	1 hour
5.3	NP- Hard and NP-Complete Problems.	1 hour
5.4	NP Completeness Proof of Clique Problem.	1 hour

5.5	NP Completeness Proof of Vertex Cover Problem.	1 hour
5.6	Approximation algorithms- Bin Packing Algorithm and Illustration.	1 hour
5.7	Graph Colouring Algorithm and Illustration.	1 hour
5.8	Randomized Algorithms (definitions of Monte Carlo and Las Vegas algorithms).	1 hour
5.9	Randomized Version of Quick Sort Algorithm with Analysis.	1 hour
5.10	Illustration of Randomized Version of Quick Sort Algorithm with Analysis.	1 hour

SEMESTER VI

PROGRAM ELECTIVE I

CST 312	FOUNDATIONS OF MACHINE LEARNING	Category	L	T	P	Credit	Year of Introduction
		PEC	2	1	0	3	2019

Preamble:

This course enables the learners to understand the mathematical foundations of Machine Learning concepts. This course covers Linear Algebra, Probability and Distributions. Concepts in this course help the learners to identify the inherent assumptions & limitations of the current methodologies and develop new Machine Learning solutions.

Prerequisite: A sound background in higher secondary school Mathematics.

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

CO 1	Illustrate operations and applications of linear equations, matrix algebra, vector spaces, eigen values & eigenvectors **(Cognitive Knowledge Level: Apply)**
CO 2	Illustrate the concepts of orthogonality & diagonalization. **(Cognitive Knowledge Level: Apply)**
CO 3	Solve computational problems using probability and random variables. **(Cognitive Knowledge Level: Apply)**
CO 4	Identify an appropriate probability distribution for a given discrete or continuous random variable and use its properties. (**Cognitive Knowledge Level: Apply)**
CO 5	Illustrate moment generating function, law of large numbers and central limit theorems (**Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO 1	✓	✓	✓	✓								✓
CO 2	✓	✓	✓	✓								✓
CO 3	✓	✓	✓	✓								✓
CO 4	✓	✓	✓	✓								✓

CO 5	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination**
	1	**2**	
Remember	**30%**	**30%**	**30%**
Understand	**30%**	**30%**	**30%**
Apply	**40%**	**40%**	**40%**
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance **: 10 marks**

Continuous Assessment Tests **: 25 marks**

Continuous Assessment Assignment **: 15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks

First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contain 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which student should answer anyone. Each question can have maximum 2 sub-divisions and carry 14 marks.

Syllabus

Module 1 (LINEAR ALGEBRA)

Systems of Linear Equations – Matrices, Solving Systems of Linear Equations. Vector Spaces - Linear Independence, Basis and Rank, Linear Mappings.

Module 2 (LINEAR ALGEBRA)

Norms - Inner Products, Lengths and Distances, Angles and Orthogonality. Orthonormal Basis, Orthogonal Complement, Orthogonal Projections. Matrix Decompositions - Eigenvalues and Eigenvectors, Eigen decomposition and Diagonalization.

Module 3 (PROBABILITY AND DISTRIBUTIONS)

Probability Space - Sample Spaces, Probability Measures, Computing Probabilities, Conditional Probability, Baye's Rule, Independence. Random Variables - Discrete Random Variables (Bernoulli Random Variables, Binomial Distribution, Geometric and Poisson Distribution, Continuous Random Variables (Exponential Density, Gamma Density, Normal Distribution, Beta Density)

Module 4 (RANDOM VARIABLES)

Functions of a Random Variable. Joint Distributions - Independent Random Variables, Conditional Distributions, Functions of Jointly Distributed Random Variables.

Expected Values - Expected Value of a Random Variable, Expectations of Functions of Random Variables, Expectations of Linear Combinations of Random Variables, Variance and Standard Deviation, Covariance and Correlation, Conditional Expectation

Module 5 (LIMIT THEOREMS)

Moment-Generating Function. Limit Theorems(Proof not expected) - Law of Large Numbers, Convergence in Distribution and the Central Limit Theorem. Distributions derived from the Normal Distribution - Chi-square, t, and F Distributions, Sample Mean and the Sample Variance.

Text book:

1. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, Mathematics for Machine Learning, Cambridge University Press (freely available at https://mml – book.github.io)

2. John A. Rice, Mathematical Statistics and Data Analysis, University of California, Berkeley, Third edition, published by Cengage.

Reference books:

1. Gilbert Strang, Linear Algebra and Its Applications, 4th Edition,
2. Axler, Sheldon, Linear Algebra Done Right, 2015 Springer
3. Stephen Boyd and Lieven Vandenberghe, Introduction to Applied Linear Algebra, 2018 published by Cambridge University Press

Sample Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Find the set ***S*** of all solutions in ***x*** of the following inhomogeneous linear systems ***Ax = b***, where ***A*** and ***b*** are defined as follows:

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -3 & 0 \\ 2 & -1 & 0 & 1 & -1 \\ -1 & 2 & 0 & -2 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 6 \\ 5 \\ -1 \end{bmatrix}$$

2. Determine the inverses of the following matrix if possible

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

3. Are the following sets of vectors linearl independent?

$$x_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 3 \\ -3 \\ 8 \end{bmatrix}$$

4. A set of ***n*** linearly independent vectors in $\boldsymbol{R^n}$ forms a basis. Does the set of vectors **(2, 4,−3) (0, 1, 1)** , **(0, 1,−1**) form a basis for $\boldsymbol{R^3}$? Explain your reasons.

Course Outcome 2 (CO2):

1. Determine which of the following sets are orthogonal sets.

$$\left\{\begin{bmatrix}3\\1\\1\end{bmatrix},\begin{bmatrix}-1\\2\\1\end{bmatrix},\begin{bmatrix}-1/2\\-2\\7/2\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\-1\\1\end{bmatrix},\begin{bmatrix}2\\1\\-1\end{bmatrix},\begin{bmatrix}3\\0\\-3\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}3\\-2\\1\\3\end{bmatrix},\begin{bmatrix}-1\\3\\-3\\4\end{bmatrix},\begin{bmatrix}3\\8\\7\\0\end{bmatrix}\right\}$$

2. Find the characteristic equation, eigenvalues, and eigenspaces corresponding to each eigenvalue of the following matrix.

$$\begin{bmatrix}2&0&4\\0&3&0\\0&1&2\end{bmatrix}$$

3. Diagonalize the following matrix, if possible

$$\begin{bmatrix}3&0&0&0\\0&2&0&0\\0&0&2&0\\1&0&0&3\end{bmatrix}$$

Course Outcome 2 (CO3):

1. Let ***J*** and ***T*** be independent events, where ***P(J)=0.4*** and ***P(T)=0.7***.
   - ***i.*** Find ***P(J∩T)***
   - ***ii.*** Find ***P(J∪T)***
   - ***iii.*** Find ***P(J∩T')***

2. Let ***A*** and ***B*** be events such that ***P(A)=0.45*** , ***P(B)=0.35*** and ***P(A∪B)=0.5***. Find ***P(A|B)***.

3. A random variable **R** has the probability distribution as shown in the following table:

r	1	2	3	4	5
P(R=r)	0.2	a	b	0.25	0.15

i. Given that ***E(R)=2.85***, find ***a*** and ***b***.
ii. Find ***P(R>2)***.

4. A biased coin (with probability of obtaining a head equal to $p > 0$) is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

5. Two players A and B are competing at a quiz game involving a series of questions. On any individual question, the probabilities that A and B give the correct answer are ***p*** and ***q*** respectively, for all questions, with outcomes for different questions being independent. The game finishes when a player wins by answering a question correctly. Compute the probability that A wins if
   i. A answers the first question,
   ii. B answers the first question.

6. A coin for which ***P(heads) = p*** is tossed until two successive tails are obtained. Find the probability that the experiment is completed on the n^{th} toss.

Course Outcome- 3 (CO4):

1. An urn contains **p** black balls, **q** white balls, and **r** red balls; and **n** balls are chosen without replacement.
   a. Find the joint distribution of the numbers of black, white, and red balls in the sample.
   b. Find the joint distribution of the numbers of black and white balls in the sample.
   c. Find the marginal distribution of the number of white balls in the sample.

2. Suppose that two components have independent exponentially distributed lifetimes, T_1 and T_2, with parameters α and β, respectively. Find (a) $P(T_1 > T_2)$ and (b) $P(T_1 > 2T_2)$.

3. Let Z_1 and Z_2 be independent random variables each having the standard normal distribution. Define the random variables X and Y by $X = Z_1 + 3Z_2$ and $Y = Z_1 + Z_2$. Argue that the joint distribution of (X, Y) is a bivariate normal distribution. What are the parameters of this distribution?

4. Given a continuous random variable x, with cumulative distribution function $F_x(x)$, show that the random variable $y = F_x(x)$ is uniformly distributed.
5. You roll a fair dice twice. Let the random variable X be the product of the outcomes of the two rolls. What is the probability mass function of X? What are the expected values and the standard deviation of X?
6. Let X be a continuous random variable with the density function $f(x) = 2x, 0 \leq x \leq 1$
   a. Find $E(X)$.
   b. Find $E(X^2)$ and $Var(X)$.

Course Outcome 5 (CO5):

1. Find the moment-generating function of a Bernoulli random variable, and use it to find the mean, variance, and third moment.
2. Use moment-generating functions to show that if **X** and **Y** are independent, then $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$.
3. Suppose that you bet Rs 5 on each of a sequence of 50 independent fair games. Use the central limit theorem to approximate the probability that you will lose more than Rs 75.
4. Suppose that the number of insurance claims, N, filed in a year is Poisson distributed with $E(N) = 10,000$. Use the normal approximation to the Poisson to approximate $P(N > 10,200)$.

Model Question paper

QP Code : **Total Pages: 4**

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION (ELECTIVE), MONTH and YEAR

Course Code: CST 312

Course Name: FOUNDATIONS OF MACHINE LEARNING

Max. Marks: 100 Duration: 3 Hours

PART A

Answer all questions, each carries 3 marks.

1 Show that with the usual operation of scalar multiplication but with addition on reals given by $x \# y = 2(x + y)$ is not a vector space.

2 Are the following vectors linearly independent? Justify your answer.

$$x_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 3 \\ -3 \\ 8 \end{bmatrix}$$

3 Find the eigenvalues of the following matrix in terms of k. Can you find an eigenvector corresponding to each of the eigenvalues?

$$\begin{bmatrix} 1 & k \\ 2 & 1 \end{bmatrix}$$

4 Find a unit vector in $\mathbf{R}^2$ that is orthogonal to **(-1 , 2)**.

5 The first three digits of a telephone number are 452. If all the sequences of the remaining four digits are equally likely, what is the probability that a randomly selected telephone number contains seven distinct digits?

6 Show that if two events ***A*** and ***B*** are independent, then ***A*** and ***B'*** are independent.

7 Prove that ***X*** and ***Y*** are independent if and only if $f_{X|Y}(x|y) = f_X(x)$ for all ***x*** and ***y***.

8 If ***X*** is a discrete uniform random variable, i.e., $P(X = k) = 1/n$ for $k = 1, 2, \ldots, n$, find ***E(X)*** and ***Var(X)***.

9 Compare the Poisson cdf and the normal approximation for (a) $\lambda = 10$, (b) $\lambda = 20$, and (c) $\lambda = 40$.

10 State law of large numbers. **10 x 3 = 30**

PART B

Answer any one Question from each module. Each question carries 14 Marks

11 a) Find all solutions to the system of linear equations **(8)**

$$\begin{aligned} -4x + 5z &= -2 \\ -3x - 3y + 5z &= 3 \\ -x + 2y + 2z &= -1 \end{aligned}$$

b) Consider the transformation $T(x, y) = (x + y, x + 2y, 2x + 3y)$. Obtain **ker T** and use this to calculate the nullity. Also find the transformation matrix for **T**. **(6)**

OR

12 a) Consider the following linear mapping **(8)**

$$\Phi : \mathbb{R}^3 \to \mathbb{R}^4$$

$$\Phi\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 + x_2 + x_3 \\ x_1 - 3x_2 \\ 2x_1 + 3x_2 + x_3 \end{bmatrix}$$

i. Find the transformation matrix T.

ii. Determine rank(T).

iii. Compute the kernel and image of the mapping and find their dimension

b) Prove that all vectors orthogonal to $[2, -3, 1]^T$ forms a subspace W of R^3. What is $dim\ (W)$ and why? **(6)**

13 a) Find an orthonormal basis of R^3 consisting of eigenvectors for the following matrix **(8)**

$$\begin{bmatrix} 1 & 0 & -2 \\ 0 & 5 & 0 \\ -2 & 0 & 4 \end{bmatrix}$$

b) Find a 3 × 3 orthogonal matrix S and a 3 × 3 diagonal matrix D such that $A = SDS^T$ **(6)**

OR

14 a) Find an orthogonal basis for the subspace of R^4 spanned by $\{ w_1 = (1, 1, 3, 2), w_2 = (1, -2, 0, -1), w_3 = (0, 2, 1, 2) \}$. **(8)**

b) Find the characteristic equation, eigenvalues, and eigenspaces corresponding t each eigenvalue of the following matrix **(6)**

$$\begin{bmatrix} 2 & 0 & 4 \\ 0 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

15 a) Three players play 10 independent rounds of a game, and each player has probability 1/3 of winning each round. Find the joint distribution of the numbers of games won by each of the three players. **(7)**

b) An experiment consists of throwing a fair coin four times. Find the probability mass function and the cumulative distribution function of the following random variables: **(7)**

i. the number of heads before the first tail

ii. the number of heads following the first tail

iii. the number of heads minus the number of tails

iv. the number of tails times the number of heads.

OR

16 a) A factory runs three shifts. On a given day, 1% of the items produced by the first shift are defective, 2% of the second shift's items are defective, and 5% of the third shift's items are defective. If the shifts all have the same productivity, what percentage of the items produced in a day are defective? If an item is defective, what is the probability that it was produced by the third shift? **(8)**

b) Show that if ***A*** and ***B*** are two independent events, then $P(A \cup B) = P(A) + P(B) - P(A)P(B)$ **(6)**

17 a) Find the joint density of $X + Y$ and X/Y, where X and Y are independent exponential random variables with parameter λ. Show that $X + Y$ and X/Y are independent. **(8)**

b) Let X be a discrete random variable that takes on values 0, 1, 2 with probabilities 1/2 , 3/8 , 1/8 , respectively. **(6)**

i. Find **E(X)** and **Var(X)**.

ii. Let $Y = X^2$. Find the probability mass function of Y and use it to find **E(Y)**.

18 a) A random square has a side length that is a uniform [0, 1] random variable. Find the expected area of the square. **(7)**

b) Let X be a continuous random variable with probability density function on $0 <= x <= 1$ defined by $f(x) = 3x^2$. Find the pdf of $Y = X^2$. **(7)**

19 a) Using the fact that the mean of the chi-squared distribution is $(n-1)$, prove that $E(S^2) = \sigma^2$. **(7)**

b) i. Random samples of size 36 are taken from an infinite population whose mean is 80 and standard deviation is 18. Find the mean and standard error of the **(7)**

sampling distribution.

ii. Why is the Central Limit Theorem so important to statistical analysis?

OR

20 a) A six-sided die is rolled 100 times. Using the normal approximation, find the probability that the face showing a six turns up between 15 and 20 times. Find the probability that the sum of the face values of the 100 trials is less than 300. **(8)**

b) Determine an interval **(a, b)** such that $P[a \leq t \leq b] = 0.80$, and that 10% of the area is on each side of a and b, assuming that the sample is of size 21. **(6)**

Teaching Plan		
No	**Topic**	**No. of Lectures (35)**
1	**Module-1 (LINEAR ALGEBRA) TB-1(Ch 2,3,4) (6 hours)**	
1.1	Systems of Linear Equations – Matrices, Solving Systems of Linear Equations.	1 hour
1.2	Vector Spaces, sub space	1 hour
1.3	Linear Independence,	1 hour
1.4	Basis and Rank	1 hour
1.5.	Linear Mappings- Kernel, Range	1 hour
1.6.	Linear Mappings- Rank, Nullity	
2	**Module-2 (LINEAR ALGEBRA) (6 hours)**	
2.1.	Norms, Inner Products, Lengths and Distances, Angles and Orthogonality,	1 hour
2.2	Orthonormal Basis, Orthogonal Complement,	1 hour
2.3	Orthogonal Projections	1 hour
2.4.	Eigenvalues and Eigenvectors	1 hour
2.5.	Eigen decomposition	1 hour
2.6.	Eigen Diagonalization	1 hour
3.	**Module-3 (PROBABILITY AND DISTRIBUTIONS) TB-2(Ch 1,2) (9 hours)**	

3.1	Sample Spaces, Probability Measures, Computing Probabilities	1 hour
3.2	Conditional Probability,	1 hour
3.3	Baye's Rule	1 hour
3.4	Independence of events	1 hour
3.5	Discrete Random Variables -Bernoulli Random Variables, Binomial Distribution	1 hour
3.6	Discrete Random Variables -Geometric Distribution	1 hour
3.7	Discrete Random Variables -Poisson Distribution	1 hour
3.8	Continuous Random Variables - Exponential Density, Gamma Density,	1 hour
3.9	Continuous Random Variables - Normal Distribution, Beta Density	1 hour
4.	**Module-4 (RANDOM VARIABLES) TB-2 (Ch 3, 4, 5, 6) (9 hours)**	
4.1	Functions of a Random Variable	1 hour
4.2	Joint Distributions - Independent Random Variables	1 hour
4.3	Conditional Distributions	1 hour
4.4	Functions of Jointly Distributed Random Variables	1 hour
4.5	Expected Value of a Random Variable,	1 hour
4.6	Expectations of Functions of Random Variables,	1 hour
4.7	Expectations of Linear Combinations of Random Variables	1 hour
4.6	Variance and Standard Deviation	1 hour
4.9	Covariance and Correlation	1 hour

5	**Module-5 (LIMIT THEOREMS) (6 hours)**	
5.1	Conditional Expectation,	1 hour
5.2	Moment-Generating Function	1 hour
5.3	Limit Theorems(Proof not expected) - Law of Large Numbers,	1 hour
5.4	Convergence in Distribution and the Central Limit Theorem.	1 hour
5.5	Distributions derived from the Normal Distribution - Chi-square and, and F Distributions,	1 hour
5.6	Distributions derived from the Normal Distribution - Sample Mean and the Sample Variance.	1 hour

CXT 322	DATA MINING	Category	L	T	P	Credit	Year of Introduction
		PEC	2	1	0	3	2021

Preamble: This course helps the learner to understand the concepts of data mining and data warehousing. It covers the key processes of data mining, data preprocessing techniques, fundamentals and advanced concepts of classification, clustering, association rule mining, web mining and text mining. It enables the learners to develop new data mining algorithms and apply the existing algorithms in real-world scenarios.

Prerequisite: NIL

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

CO1	Explain the key process of data mining and data warehousing in application domains. **(Cognitive Knowledge Level: Understand)**
CO2	Make use of appropriate preprocessing techniques to convert raw data into suitable format for practical data mining tasks **(Cognitive Knowledge Level: Apply)**
CO3	Use classification and clustering algorithms in various application domains. **(Cognitive Knowledge Level: Apply)**
CO4	Comprehend the use of association rule mining techniques. **(Cognitive Knowledge Level: Apply)**
CO5	Explain advanced data mining concepts and their applications in emerging domains **(Cognitive Knowledge Level: Understand)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓										✓
CO2	✓	✓	✓	✓	✓							✓
CO3	✓	✓	✓	✓	✓							✓

CO4	✓	✓	✓	✓	✓							✓
CO5	✓	✓										✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination Marks**
	Test 1 (Marks)	**Test 2 (Marks)**	
Remember	**20**	**20**	**20**
Understand	**30**	**30**	**30**
Apply	**50**	**50**	**50**
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : **10 marks**
Continuous Assessment - Test : **25 marks**
Continuous Assessment - Assignment : **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module – 1 (Introduction to Data Mining and Data Warehousing)

Data warehouse-Differences between Operational Database Systems and Data Warehouses, Multidimensional data model- Warehouse schema, OLAP Operations, Data Warehouse Architecture, Data Warehousing to Data Mining, Data Mining Concepts and Applications, Knowledge Discovery in Database Vs Data mining, Architecture of typical data mining system, Data Mining Functionalities, Data Mining Issues.

Module - 2 (Data Preprocessing)

Data Preprocessing-Need of data preprocessing, Data Cleaning- Missing values, Noisy data, Data Integration and Transformation, Data Reduction-Data cube aggregation, Attribute subset selection,

Dimensionality reduction, Numerosity reduction, Discretization and concept hierarchy generation.

Module - 3 (Advanced classification)

Classification Models: Introduction to Classification and Prediction, Issues regarding classification and prediction, Decision Tree- ID3, C4.5, Naive Bayes Classifier.

Rule based classification- 1R, Neural Networks-Back propagation. Support Vector Machines, Lazy Learners-K Nearest Neighbor Classifier. Accuracy and error measures evaluation. Prediction: -Linear Regression and Non-Linear Regression.

Module 4 (Association Rule Analysis)

Association Rules Mining: Concepts, Apriori and FP-Growth Algorithm. Cluster Analysis: Introduction, Concepts, Types of data in cluster analysis, Categorization of clustering methods. Partitioning method- K-Means and K-Medoid Clustering.

Module 5 (Advanced Data Mining Techniques)

Hierarchical Clustering method: BIRCH. Density-Based Clustering –DBSCAN and OPTICS. Advanced Data Mining Techniques- Introduction, Web Mining- Web Content Mining, Web Structure Mining, Web Usage Mining.Text Mining. Graph mining- Apriori based approach for mining frequent subgraphs. Social Network Analysis- characteristics of social networks. Link mining- Tasks and challenges. WEKA tool.

Text Books

1. Dunham M H, “Data Mining: Introductory and Advanced Topics”, Pearson Education, New Delhi, 2003.
2. Arun K Pujari, “Data Mining Techniques”, Universities Press Private Limited,2008.
3. Jaiwei Han and MichelineKamber, “Data Mining Concepts and Techniques”, Elsevier, 2006.

Reference Books

1. M Sudeep Elayidom, “Data Mining and Warehousing”, 1st Edition, 2015, Cengage Learning India Pvt. Ltd.
2. Mehmed Kantardzic, “Data Mining Concepts, Methods and Algorithms”, John Wiley and Sons, USA, 2003.
3. Pang-Ning Tan and Michael Steinbach, “Introduction to Data Mining”, Addison Wesley, 2006.

Sample Course Level Assessment Questions:

Course Outcome 1 (CO1):

1. (a) Explain the OLAP operations in a multidimensional model.
   (b) Compare the techniques used in ROLAP, MOLAP and HOLAP.
2. Explain the various data mining issues with respect to mining methodology, user interaction and diversity of data types.
3. 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) Draw star and snowflake schema diagrams for the data warehouse.
   b) Starting with the base cuboid [day; doctor; patient], what specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2004?

Course Outcome 2 (CO2):

1. Use the methods below to normalize the following group of data:100, 200, 300, 400, 550, 600, 680, 850, 1000
   (a) min-max normalization by setting min = 0 and max = 1.
   (b) z-score normalization.
   (c) Normalization by decimal scaling.

   Comment on which method you would prefer to use for the given data, giving reasons as to why.
2. Identify a suitable dataset from any available resources and apply different preprocessing steps that you have learned. Observe and analyze the output obtained.

Course Outcome 3 (CO3):

1. Illustrate the working of ID3 algorithm with the following example.

MOTOR	WHEEELS	DOORS	SIZE	TYPE	CLASS
NO	2	0	small	cycle	bicycle
NO	3	0	small	cycle	tricycle
YES	2	0	small	cycle	motorcycle
YES	4	2	small	automobile	Sports car
YES	4	3	medium	automobile	minivan
YES	4	4	medium	automobile	sedan
YES	4	4	large	automobile	sumo

2. Illustrate the working of K medoid algorithm for the given dataset. A1= (3,9), A2= (2,5), A3= (8,4), A4= (5,8), A5= (7,5), A6= (6,4), A7= (1,2), A8= (4,9).

3. Take a suitable dataset from available resources and apply all the classification and clustering algorithms that you have studied on original and preprocessed datasets. Analyze the performance variation in terms of different quality metrics. Give a detailed report based on the analysis.

Course Outcome 4 (CO4):

1. A database has five transactions. Let min sup = 60% and min conf = 80%.

TID	*items_bought*
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y }
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I ,E}

a) Find all frequent item sets using Apriori and FP-growth, respectively. Compare the efficiency of the two mining processes.
b) List all of the strong association rules (with support s and confidence c) matching the following meta rule, where X is a variable representing customers, and item denotes variables representing items (e.g., "A", "B", etc.)

$$\forall x \in transaction, buys(X, item_1) \wedge buys(X, item_2) \Rightarrow buys(X, item_3) \quad [s, c]$$

2. Identify and list some scenarios in which association rule mining can be used, and then use at least two appropriate association rule mining techniques in one of the two scenarios.

Course Outcome 5 (CO5):

1. Consider an e-mail database that stores a large number of electronic mail (e-mail) messages. It can be viewed as a semi structured database consisting mainly of text data.
   Discuss the following.
   a. How can such an e-mail database be structured so as to facilitate multidimensional search, such as by sender, by receiver, by subject, and by time?
   b. What can be mined from such an e-mail database?

c. Suppose you have roughly classified a set of your previous e-mail messages as junk, unimportant, normal, or important. Describe how a data mining system may take this as the training set to automatically classify new e-mail messages or unclassified ones.

2. Precision and recall are two essential quality measures of an information retrieval system.
   (a) Explain why it is the usual practice to trade one measure for the other.
   (b) Explain why the F-score is a good measure for this purpose.
   (c) Illustrate the methods that may effectively improve the F-score in an information retrieval system.
3. Explain DBSCAN algorithm with an example.

Model Question Paper

QP CODE:

Reg No: _____

Name:_____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CXT 322
Course Name: Data Mining

Max.Marks:100 **Duration: 3 Hours**

PART A
Answer All Questions. Each Question Carries 3 Marks

1. What is the purpose of data discretization in data mining? List out any four data discretization strategies.

2. How is a data warehouse different from a database?

3. Explain concept hierarchy with an example.

4. Explain heuristic methods of attribute subset selection techniques.

5. 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. Calculate precision and recall of the data.

6. Given two objects represented by the tuples (22,1,42,10) and (20,0, 36,8). Compute the Euclidean and Manhattan distance between the two objects.

7. What is the significance of CF (Clustering Feature) in BIRCH Algorithm?

8. How to compute the dissimilarity between objects described by binary variables?

9. How density based clustering varies from other methods?

10. Describe any two-text retrieval indexing techniques.

(10x3= 30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11 (a) Suppose a data warehouse consists of three measures: customer, account and branch and two measures count (number of customers in the branch) and balance. Draw the schema diagram using snowflake schema and tar schema. **(7)**

(b) Explain three- tier data warehouse architecture with a neat diagram. **(7)**

OR

12 (a) Explain various stages in knowledge discovery process with a neat diagram. **(7)**

(b) Illustrate different OLAP operations in multidimensional data model. **(7)**

13 (a) Describe various techniques for numerosity reduction in data mining. **(7)**

(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. **(7)**

(1) Use min-max normalization to transform the value 35 for age onto the range [0-1].
(2) Use z-score normalization to transform the value 35 for age, where the standard deviation of age is 12.94 years.
(3) Use normalization by decimal scaling to transform the value 35 for age.
(4) Use smoothing by bin means to smooth the above data, using a bin depth of 3.
Illustrate your steps..

OR

14 (a) Suppose a group of 12 sales price records has been sorted as follows: 5, 10, 11, 13, 15, 35, 50, 55, 72, 92, 204, 215. Sketch examples of each of the following sampling techniques: SRSWOR, SRSWR, cluster sampling, stratified sampling. Use samples of size 5 and the strata "youth," "middle aged," and "senior." **(8)**

(b) Partition the above data into three bins by each of the following methods: **(6)**
(i) equal-frequency (equi-depth) partitioning
(ii) equal-width partitioning

15 (a) The following table shows 'Car Theft Database'. Using Naive-Bayes algorithm to classify a **Red Domestic SUV as 'Stolen' or 'Not Stolen'.** **(8)**

N	Color	Type	Origin	Stolen?
1	red	sports	domestic	yes
2	red	sports	domestic	no
3	red	sports	domestic	yes
4	yellow	sports	domestic	no
5	yellow	sports	imported	yes
6	yellow	SUV	imported	no
7	yellow	SUV	imported	yes
8	yellow	SUV	domestic	no
9	red	SUV	imported	no
10	red	sports	domestic	yes

(b) How does back propagation algorithm work? **(6)**

OR

16 (a) For a sunburn dataset given below, find the first splitting attribute for the decision tree by using the ID3 algorithm. **(10)**

Name	**Hair**	**Height**	**Weight**	**Lotion**	**Class**
Sarah	Blonde	Average	Light	No	Sunburn
Dana	Blonde	Tall	Average	Yes	None
Alex	Brown	Tall	Average	Yes	None
Annie	Blonde	Short	Average	No	Sunburn
Emily	Red	Average	Heavy	No	Sunburn
Pete	Brown	Tall	Heavy	No	None
John	Brown	Average	Heavy	No	None
Katie	Blonde	Short	Light	Yes	None

(b) Describe the purpose of kernel function in SVM with a suitable example **(4)**

17 (a) Suppose that our task is to cluster given sample data (height, weight) into two clusters. Let Al and A2 be initial cluster centroids. Apply k-means algorithm to find a set of clusters. Use Manhattan distance function as dissimilarity measure. **(7)**

Id	Height	Weight
A1	185	72
A2	170	56
A3	168	60
A4	179	68
A5	182	72
A6	188	77

(b) A database has five transactions. Let min_sup =60% and min-con=50%. Find all frequent item sets using Apriori algorithm. **(7)**

TID	*items_bought*
T100	{M,O,N,K,E,Y}
T200	{D,O,N,K,E,Y}
T300	{M,A,K,E}
T400	{M,U,C,K,Y}
T500	{C,O,O,K,l,E}

OR

18 (a) A database has six transactions. Let min_sup be 60% and min_conf be 80%. **(8)**

TID	*items_bought*
T1	I1, I2, I3
T2	I2, I3, I4
T3	I4, I5
T4	I1, I2, I4
T5	I1, I2, I3, I5
T6	I1, I2, I3, I4

Find frequent itemsets using FP Growth algorithm and generate strong association rules from a three-item dataset.

(b) Explain K-Medoid Clustering algorithm. **(6)**

19 (a) Describe web content mining techniques. **(7)**

(b) Explain BIRCH Clustering Method. **(7)**

OR

20 (a) Define the following terms related to DBSCAN algorithm with suitable figures **(8)**

i) Core object.
ii) Directly density reached.
iii) Density reachable.
iv) Density connected.

(b) Describe different Text retrieval methods. Explain the relationship between text mining and information retrieval and information extraction**.** **(6)**

No	Contents	No. of lecture hours (37 Hrs)
Module 1(Introduction to Data Mining and Data Warehousing) (5 hours)		
1.1	Introduction to Data warehouse	1
1.2	Multidimensional data model, OLAP Operations	1
1.3	Data Warehouse Architecture, Data Warehousing to Data Mining	1
1.4	Data Mining Concepts and Applications, Knowledge Discovery in Database Vs Data mining	1
1.5	Architecture of typical data mining system, Data Mining Functionalities, Data Mining Issues	1
Module 2(Data Preprocessing) (6 hours)		
2.1	Data Preprocessing: Need of Data Preprocessing, Data Cleaning- Missing values, Noisy data.	1
2.2	Data integration	1
2.3	Data transformation	1
2.4	Data Reduction-Data cube aggregation, Attribute subset selection	1
2.5	Data Reduction-Dimensionality reduction	1
2.6	Numerosity reduction, Discretization and concept hierarchy generation	1
	Module 3(Advanced classification and Cluster analysis) (10 hours)	
3.1	Introduction to classification and prediction and its issues	1
3.2	Classification based on Decision tree ID3	1
3.3	Decision Tree- ID3, Decision trees C4.5	1
3.4	Naïve Bayes Classifier	1
3.5	Rule based classification - 1R	1
3.6	NN - Back propagation	1
3.7	Support vector machine	1
3.8	K-nearest neighbor classifier, Accuracy and error measures-evaluation	1
3.9	Linear Regression	1
3.10	Non-Linear Regression	1
	Module 4(Association Rule Analysis) (7 hours)	
4.1	Association Rules: Introduction, Methods to discover association rules	1
4.2	Apriori algorithm (Level-wise algorithm)	1

4.3	FP-tree Growth Algorithm	1
4.4	Improvements in Apriori algorithm	1
4.5	Clustering introduction, Types of data in cluster analysis	1
4.6	Partition Algorithm – K-Means	1
4.7	K-Medoid Algorithm	1
Module 5(Advanced Data Mining Techniques) (9 hours)		
5.1	Hierarchical clustering- BIRCH	1
5.2	Density based clustering-DBSCAN, OPTICS	1
5.3	Web mining - content, structure	1
5.4	Web usage mining - Pattern Discovery, Pattern Analysis	1
5.5	Text Mining-Text Data Analysis and information Retrieval, Basic measures for Text retrieval	1
5.6	Text Retrieval methods, Text Indexing Techniques, Query Processing Techniques	1
5.7	Graph mining	1
5.8	Social network analysis, link mining	1
5.9	Weka Tool	1

CST 342	AUTOMATED VERIFICATION	Category	L	T	P	CREDIT	YEAR OF INTRODUCTION
		PEC	2	1	0	3	2019

le: This course is intended to impart the basic theory and algorithm for an automatic verification process namely model checking. This course covers finite-state modelling of hardware/software, linear-time properties, classification of linear-time properties, Linear Temporal Logic (LTL) - a formal language for property specification, LTL model checking algorithm and model checking case studies. This course enables the learners to prove correctness of a hardware/software used in safety critical systems in domains such as avionics, health care and automotive.

Prerequisite: NIL

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

CO1	Illustrate an application for model checking. **(Cognitive Knowledge Level: Understand)**
CO2	Describe finite-state modelling for hardware and software. **(Cognitive Knowledge Level: Understand)**
CO3	Identify linear-time properties required to represent the requirements of a system. **(Cognitive Knowledge Level: Apply)**
CO4	Specify a given linear-time property in Linear Temporal Logic (LTL). **(Cognitive Knowledge Level: Apply)**
CO5	Perform LTL model checking using the tool Symbolic Analysis Laboratory (SAL). **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓	✓							✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓								✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓	✓	✓	✓	✓						✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	PO7	Environment and Sustainability
PO2	Problem Analysis	PO8	Ethics
PO3	Design/Development of solutions	PO9	Individual and team work
PO4	Conduct investigations of complex problems	PO10	Communication
PO5	Modern tool usage	PO11	Project Management and Finance
PO6	The Engineer and Society	PO12	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks
	Test 1 (Marks)	Test 2 (Marks)	
Remember	**30**	**30**	**30**
Understand	**30**	**30**	**30**
Apply	**40**	**40**	**40**
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Test : 25 marks

Continuous Assessment Assignment : 15 marks (Out 15, 10 marks shall be given for a model checking project to be implemented in SAL.)

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each

question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module - 1 (Introduction to Model Checking)

System Verification – Hardware and Software Verification, Model Checking, Characteristics of Model Checking.

Transition Systems – Transition System, Direct Predecessors and Successors, Terminal State, Deterministic Transition System.

Executions - Execution Fragment, Maximal and Initial Execution Fragment, Execution, Reachable States.

Module - 2 (Linear Time Properties)

Linear-Time (LT) Properties - Deadlock. Linear-Time Behavior - Paths and State Graph, Path Fragment, Maximal and Initial Path Fragment, Path. Traces - Trace and Trace Fragment, LT Properties - LT Property, Satisfaction Relation for LT Properties, Trace Equivalence and LT Properties. Safety Properties and Invariants - Invariants, Safety Properties, Trace Equivalence and Safety properties. Liveness Properties - Liveness Property, Safety vs. Liveness Properties. Fairness - Fairness, Unconditional, Weak and Strong Fairness, Fairness Strategies, Fairness and Safety. (Definition and examples only for all topics - no proof required).

Module - 3 (Regular Properties)

Regular Properties - Model Checking Regular Safety properties - Regular Safety property, Verifying Regular Safety Properties. Automata on Infinite Words - ω -Regular Languages and Properties, Nondeterministic Buchi Automata (NBA), Deterministic Buchi Automata (DBA),

Generalised Buchi Automata (Definitions only). Model Checking ω -Regular Properties - Persistence Properties and Product, Nested Depth-First Search (Only algorithms required).

Module - 4 (Linear Time Logic)

Linear Temporal Logic (LTL) - Syntax, Semantics, Equivalence of LTL Formulae, Weak Until, Release and Positive Normal Form, Fairness, Safety and Liveness in LTL (Definitions only). Automata Based LTL Model Checking (Algorithms and examples only).

Module - 5 (Model Checking in SAL)

Introduction - Introduction to the tool Symbolic Analysis Laboratory (SAL).

The Language of SAL - The expression language, The transition Language, The module language, SAL Contexts.

SAL Examples - Mutual Exclusion, Peterson's Protocol, Synchronous Bus Arbiter, Bounded Bakery protocol, Bakery Protocol, Traffic Signalling System.

Text Books

1. Christel Baier and Joost-Pieter Katoen, Principles of Model Checking, The MIT Press. (Modules 1 - 4)
2. Leonardo de Moura, Sam Owre and N. Shankar, The SAL Language Manual, SRI International (http://sal.csl.sri.com/doc/language-report.pdf, Chapters 1, 3, 4, 5, 6, 7) (Module 5)

Reference Materials

1. SAL Examples (http://sal.csl.sri.com/examples.shtml) (Module 5)

Sample Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Illustrate how model checking can make a system design reliable, based on a required set of properties/constraints.

Course Outcome 2 (CO2):

1. Consider a message delivery system. The sender *s* is trying to send a series of messages to the receiver *r* in such a way that the $(i+1)^{st}$ message is sent only after the i^{th} message is delivered. There is a possibility of error in sending a message and in that case, *s* keeps on

trying until it is able to send the message. Show a finite state transition system modeling this system.

Course Outcome 3 (CO3):

1. Consider a shared memory segment *s* protected using a mutex lock variable *m*. Two processes p_1 and p_2 are trying to access *s*. List the Linear Time properties of the system which will ensure safety, liveness and fairness.

Course Outcome 4 (CO4):

1. Show the LTL specifications of the safety, liveness and fairness properties listed for the assessment question given in CO3.

Course Outcome 5 (CO5):

1. Model the system mentioned in the question given in CO3 in SAL and verify that the system is correct with respect to the LTL properties shown as the answer for CO4.

Model Question paper

QP CODE: **PAGES: 3**

Reg No:_____ **Name :_____**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST342

Course Name: Automated Verification

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all questions. Each question carries 3 marks.

1. Define model checking. Show the schematic diagram of the model checking approach.

2. Show a transition system modeling a coffee/Tea vending machine.

3. Define invariant as a Linear Time (LT) property. Give an example

4. List any three Linear Time properties in the Mutual Exclusion problem of processes.

5. Illustrate the construction of a product automaton from two automata.

6. Differentiate between Deterministic Buchi Automaton and Non-deterministic Buchi Automaton.

7. Specify the following statements about traffic lights in Linear Temporal Logic (LTL).
   a. Once red, the light can not become green immediately.
   b. Once red, the light always becomes green eventually after being yellow for some time.

8. What is Positive Normal Form (PNF) in LTL? Give an example.

9. List any three applications of the tool Symbolic Analysis Laboratory (SAL).

10. What is a SAL context? Give an example.

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Explain in detail the various phases of the model checking process. **(8)**

    (b) Explain the strengths and weaknesses of model checking. **(6)**

OR

12. (a) Define and illustrate the following terms of a transition system. **(14)**
    a. Execution Fragment
    b. Maximal and Initial Execution Fragment
    c. Execution
    d. Reachable States

13. (a) With an example, explain the satisfaction relation for LT properties. **(7)**

(b) What is trace equivalence in Transition Systems? Give an example to show that if two transition systems satisfy the trace equivalence property, then they satisfy the same set of LT properties. **(7)**

OR

14. (a) Give the transition system for the fault tolerant variant of the dining philosophers problem. **(4)**

(b) With a suitable example, explain the algorithms to check whether a Transition System satisfies an invariant or not. **(10)**

15. (a) Explain Regular Safety Properties with a suitable example. **(7)**

(b) Illustrate an algorithm for verifying Regular Safety Properties. **(7)**

OR

16. (a) Explain ω-Regular Properties. **(4)**

(b) Illustrate how ω-Regular Properties are verified. **(10)**

17. (a) Explain the syntax of Linear Temporal Logic (LTL). **(7)**

(b) Explain the semantics of LTL. **(7)**

OR

18. (a) With an example, give the difference between until and weak until in LTL. **(4)**

(b) With a suitable example, explain automata based LTL model checking. **(10)**

19. (a) Explain Peterson's protocol. What are the LTL properties to be verified to ensure its correctness? **(8)**

(b) Write a SAL script for the verification of Peterson's protocol. **(6)**

OR

20. (a) Show the SAL model corresponding to Bakery protocol. **(8)**

(b) List any three Linear Time properties of this model and show their LTL **(6)**

Teaching Plan

Module 1 (Introduction to Model Checking)		**4 Hours**
1.1	**System Verification** – Hardware and Software Verification, Model Checking, Model Checking	1 Hour
1.2	**Transition Systems** – Transition System, Direct Predecessors and Successors, Terminal State, Deterministic Transition System	1 Hour
1.3	**Executions** - Execution Fragment, Maximal and Initial Execution Fragment	1 Hour
1.4	Execution, Reachable States	1 Hour
Module 2 (Linear Time Properties)		**8 Hours**
2.1	Linear-Time (LT) Properties - Deadlock	1 Hour
2.2	Linear-Time Behavior - Paths and State Graph, Path Fragment, Maximal and Initial Path Fragment, Path	1 Hour
2.3	Traces - Trace and Trace Fragment	1 Hour
2.4	LT Property, Satisfaction Relation for LT Properties, Trace Equivalence and LT Properties	1 Hour
2.5	Invariants	1 Hour
2.6	Safety Properties, Trace Equivalence and Safety properties	1 Hour
2.7	Liveness Property, Safety vs. Liveness Properties	1 Hour
2.8	Fairness, Unconditional, Weak and Strong Fairness, Fairness Strategies, Fairness and Safety	1 Hour
Module 3 (Regular Properties)		**9 Hours**
3.1	**Regular Properties** - Model Checking Regular Safety properties - Regular Safety property	1 Hour
3.2	Verifying Regular Safety Properties	1 Hour
3.3	Automata on Infinite Words - ω -Regular Languages and Properties	2 Hour

3.4	Nondeterministic Buchi Automata (NBA), Deterministic Buchi Automata (DBA), Generalised Buchi Automata	1 Hour
3.5	Model Checking ω-Regular Properties - Persistence Properties and Product - Lecture 1	1 Hour
3.6	Persistence Properties and Product - Lecture 2	1 Hour
3.7	Nested Depth-First Search (Lecture 1)	1 Hour
3.8	Nested Depth-First Search (Lecture 2)	1 Hour
Module 4 (Linear Time Logic)		**7 Hours**
4.1	**Linear Temporal Logic –** Linear Temporal Logic (LTL) - Syntax	1 Hour
4.2	Semantics - Lecture 1	1 Hour
4.3	Equivalence of LTL Formulae, Weak Until	1 Hour
4.4	Release and Positive Normal Form	1 Hour
4.5	Fairness, Safety and Liveness in LTL	1 Hour
4.6	Automata Based LTL Model Checking (Lecture 1)	1 Hour
4.7	Automata Based LTL Model Checking (Lecture 2)	1 Hour
Module 5 (Model Checking in SAL)		**7 Hours**
5.1	Introduction - Introduction to the tool Symbolic Analysis Laboratory (SAL).	1 Hour
5.2	The Language of SAL - The expression language, The transition Language	1 Hour
5.3	The module language, SAL Contexts.	1 Hour
5.4	SAL Examples - Mutual Exclusion	1 Hour
5.5	Peterson's Protocol, Synchronous Bus Arbiter	1 Hour
5.6	Bounded Bakery protocol, Bakery Protocol	1 Hour
5.7	Traffic Signalling System	1 Hour

CXT 332	MULTIMEDIA TECHNOLOGIES	Category	L	T	P	Credits	Year of Introduction
		PEC	2	1	0	3	2021

Preamble:

This course helps the learner to study the relevance and underlying infrastructure of multimedia systems. It also enables the students to apply contemporary theories of multimedia learning to the development of multimedia products.

Prerequisite: NIL

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

CO#	Course Outcomes
CO1	Describe the basic concepts of multimedia data representations, color models, audio and video signals and different compression techniques. **(Cognitive Knowledge Level: Understand)**
CO2	Apply the knowledge of various compression algorithms for developing multimedia applications. **(Cognitive Knowledge Level: Apply)**
CO3	Summarize the image compression standards, audio and video compression techniques. (**Cognitive Knowledge Level: Understand)**
CO4	Discuss the concepts of content-based image retrieval. **(Cognitive Knowledge Level: Understand)**
CO5	Describe the concept of cloud computing and its application in multimedia technologies. **(Cognitive Knowledge Level: Understand)**

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO 1	✓	✓										✓
CO 2	✓	✓	✓		✓							✓
CO 3	✓	✓										✓
CO 4	✓	✓										✓
CO 5	✓	✓										✓

Abstract POs Defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	50	50	50
Apply	20	20	20

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance **10 marks**

Continuous Assessment Tests (Average of Series Tests 1 & 2) **25 marks**

Continuous Assessment Assignment **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which students should answer any one. Each question can have a maximum 2 subdivisions and carries 14 marks.

Syllabus

Module 1

Multimedia Basics: Multimedia, Hypermedia, WWW, Internet, Multimedia Software, Editing and Authoring Tools.
Graphics and Image Data Representation— Graphics/Image Data Types, Popular File Formats. Color in Image - Color Science, Color Models in Images.
Concepts in Digital Video-Digital Video.

Module 2

Basics of Digital Audio—Digitization of Sound, Musical Instrument Digital Interface (MIDI).
Lossless Compression Algorithms— Introduction, Basics of Information Theory, Run-Length Coding, Variable-Length Coding, Dictionary-Based Coding, Arithmetic Coding, Lossless Image Compression.

Module 3

Lossy Compression Algorithms - Distortion Measures, The Rate- Distortion Theory, Quantization, Transform Coding, Wavelet-Based Coding, Wavelet Packets.
Image Compression Standards— JPEG, JPEG2000, JPEG-LS, Bi-level Image Compression Standards.
Audio Compression Techniques-— ADPCM in Speech Coding, Vocoders, Psychoacoustics, MPEG Audio.

Module 4

Basic Video Compression Techniques - Introduction to Video Compression, Video Compression Based on Motion Compensation, MPEG-1-Video Bitstream, MPEG-2- Supporting Interlaced Video, MPEG-4 - Overview, MPEG-7- Introduction.
Content-Based Retrieval in Digital Libraries- Image Retrieval, CBIRD: A Case Study, Quantifying Search Results, Querying on Videos.

Module 5

Cloud Computing for Multimedia Services - Cloud Computing Overview, Multimedia Cloud Computing, Cloud Assisted Media Sharing, Computation Offloading for Multimedia Services, Interactive Cloud Gaming.

Text Book

1. Ze-Nian Li and M. S. Drew, Fundamental *of Multimedia.*, Pearson Education,2004

References

1. K. R. Rao, Zoran S. Bojkovic, D. A. Milovanovic, Introduction *to Multimedia Communications.*, Wiley.
2. V. S. Subrahmanian,. Principles *of Multimedia Database Systems.*,

   Morgan Kaufmann Publishers.

3. R. Steinmetz and K. Nahrstedt,. Multimedia*: Computing, Communication & Applications.*, Pearson Education.
4. John F.Koegel Buford, .*Multimedia Systems.*, Pearson Education.
5. Prabhat K. Andheigh, Kiran Thakrar, *Multimedia Systems design.*, Prentice Hall PTR.
6. Jerry D. Gibson,. Multimedia *Communications: Directions and Innovations*., Elsevier Science.

Course level assessment questions

Course outcome 1 (CO1):

1. Discuss the relation between multimedia and hypermedia.
2. The Pitch Bend opcode in MIDI is followed by two data bytes specifying how the control is to be altered. How many bits of accuracy does this amount of data correspond to? Why?

Course outcome 2 (CO2):

1. Work out the details of the encoder and decoder for adaptive arithmetic coding when the input symbols are 01111.
2. What are the advantages and disadvantages of arithmetic coding as compared to Huffman coding?
3. Assume we have an unbounded source we wish to quantize using an M-bit midtread uniform quantizer. Derive an expression for the total distortion if the step size is 1.

Course outcome 3 (CO3):

1. Could we use wavelet-based compression in ordinary JPEG? How?
2. Draw block diagrams for an MPEG-2 encoder and decoder for (a) SNR and spatial hybrid scalability, (b) SNR and temporal hybrid scalability.
3. What is the compression ratio of MPEG audio if stereo audio sampled with 16 bits per sample at 48 kHz is reduced to a bitstream of 256 kbps?

Course outcome 4 (CO4):

1. What is the need of content-based image retrieval?
2. How can you evaluate the performance of image search engines?

Course outcome 5 (CO5):

1. Differentiate between public and private cloud computing.
2. How does cloud help in live media streaming service?
3. Explain the modules and their relations in multimedia cloud computing.
4. What are the different requirements for computation offloading?
5. Explain different cloud service models.

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CXT332

Course Name: Multimedia Technologies

Max Marks :100 **Duration: 3 Hrs.**

PART A

(Answer all Questions. Each question carries 3 Marks)

1. Differentiate between multimedia & hypermedia.
2. What are the different types of tweening?
3. Explain about MIDI.
4. Mention the models used in lossless compression schemes.
5. State the rate- distortion theory.
6. Explain the need of vocoders.
7. What is the need of content-based image retrieval?
8. Draw the architecture for the layers of MPEG-1 video bit streams.
9. Differentiate between public and private cloud computing.
10. How does cloud help in live media streaming service?

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. a) Illustrate the different multimedia software & tools. (8)
    b) Explain about various color models. (6)

OR

12. Explain in detail the different image data types? (14)
13. Explain
    a) Variable length coding (7)
    b) Arithmetic coding. (7)

OR

14. a) Describe vector quantization theory. Discuss its merits over scalar quantization. (8)
    b) Explain the different techniques for digitization of sound. (6)

15. a) Describe the various steps of the JPEG compression process. (7)

b) Enumerate the difference between JPEG and JBIG standards. (7)

OR

16. a) Explain Bi-level image compression standards. (8)

b) Differentiate between uniform and non-uniform scalar quantization. (6)

17. a) Illustrate Similarity- Based Retrieval with example. (8)

b) How can you evaluate the performance of image search engines? (6)

OR

18. Outline the key features of the following

a) MPEG-2 (7)

b) MPEG-4 (7)

19. a) Explain different cloud service models. (8)

b) What are the different requirements for computation offloading? (6)

OR

20. a) Explain the modules and their relations in multimedia cloud computing. (8)

b) What are the issues and challenges of cloud gaming? (6)

(5 x 14 = 70)

Teaching Plan

No	Contents	No of Lecture37 Hrs
	Module 1 (6 hours)	
1.1	Multimedia Basics - Multimedia, Hypermedia, WWW, Internet.	1
1.2	Multimedia Software, Editing and Authoring Tools.	1
1.3	Graphics/Image Data Types, Popular File Formats.	1
1.4	Color Science.	1
1.5	Color Models in Images.	1
1.6	Concepts in Digital Video.	1
	Module 2 (9 hours)	
2.1	Digitization of Sound.	1
2.2	Musical Instrument Digital Interface(MIDI).	1
2.3	Lossless Compression Algorithms— Introduction.	1
2.4	Basics of Information Theory.	1
2.5	Run-Length Coding.	1
2.6	Variable-Length Coding.	
2.7	Dictionary-Based Coding,	1
2.8	Arithmetic Coding	
2.9	Lossless Image Compression.	1

Module 3 (8 hours)		
3.1	Lossy Compression Algorithms - Distortion Measures, The Rate- Distortion Theory.	1
3.2	Quantization, Transform Coding.	1
3.3	Wavelet-Based Coding, Wavelet Packets.	1
3.4	Image Compression Standards— JPEG, JPEG2000.	1
3.5	JPEG-LS, Bi-level image compression standards.	1
3.6	Audio Compression Techniques-—ADPCM in Speech Coding	1
3.7	Vocoders, Psychoacoustics.	1
3.8	MPEG Audio.	1
Module 4 (8 hours)		
4.1	Basic Video Compression Techniques - Introduction to Video Compression.	1
4.2	Video Compression Based on Motion Compensation.	
4.3	MPEG-1-Video Bitstream,MPEG-2- Supporting interlaced video.	1
4.4	MPEG-4 - Overview.	1
4.5	MPEG-7- Introduction.	1
4.6	Image Retrieval.	
4.7	CBIRD: A Case Study.	1
4.8	Quantifying Search Results, Querying on Videos.	1

Module 5 (6 hours)		
5.1	Cloud Computing Overview.	1
5.2	Multimedia Cloud Computing.	1
5.3	Cloud Assisted Media Sharing.	1
5.4	Computation Offloading -Requirements, Service Partitioning for Video Coding	1
5.5	Case Study: Cloud Assisted Motion Estimation	1
5.6	Interactive Cloud Gaming	1

CXT 352	VISUAL DESIGN AND COMMUNICATION	Category	L	T	P	Credits	Year of Introduction
		PEC	2	1	0	3	2021

Preamble:

This course helps the learner to understand the basic concepts of visual design and communication. This course covers basic elements and principles of design, introduction to design projects, process of color and color theory, value and typeface design and design in visual communication. It enables the learners to perform visual design 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#	Course Outcomes
CO1	Summarize the basic ideas concerning the elements that make up a visual design and how these elements are organized. **(Cognitive Knowledge Level: Understand)**
CO2	Make use of digital tools to solve the steps involved in specific design problems **(Cognitive Knowledge Level: Apply)**
CO3	Describe the key concepts and applications of color and color theory in visual design **(Cognitive Knowledge Level: Understand)**
CO4	Explain value and typeface in design principles. **(Cognitive Knowledge Level: Understand)**
CO5	Describe the concepts of design in visual communication. **(Cognitive Knowledge Level: Understand)**
CO6	Utilize the basic understanding of design to digitally produce two-dimensional images **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✔	✔	✔									✔
CO2	✔	✔	✔		✔							✔
CO3	✔	✔	✔									✔
CO4	✔	✔	✔									✔
CO5	✔	✔	✔									✔
CO6	✔	✔	✔	✔	✔							✔

Abstract POs Defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	40	40	50
Apply	30	30	20

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance	**10 marks**
Continuous Assessment Tests (Average of Series Tests 1& 2)	**25 marks**
Continuous Assessment Assignment	**15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question

from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which students should answer any one. Each question can have a maximum 2 subdivisions and carries 14 marks.

Syllabus

Module – 1

Design Elements: Introduction, Elements of Design- Line, Shape, Negative Space, Volume, Value, Color, Texture.

Principles of Design: Gestalt - Unity or Harmony, Unity and Placement using Line, shape, Repetition using line, shape. Rhythm - Unity using value, Unity and continuity, Variety. Focal point and visual hierarchy, Balance.

Module - 2

Introduction to design projects: Conceptual process - research and define problem, preliminary solutions, organized list of options and thumbnail sketches, rough- refined preliminary visual design solutions and alternatives, composites or comps, the final design or finish. Production process- Computer applications used in design, software applications for graphics.

Basic methods of Abstraction: Simplification, Repetition, Line and Shape, Type combination.

Module - 3

Color and color theory: Background in color theory - Defining color, Color and unity, Color unity via Tonality, Color and variety, Warm and cool colors, Color discord, Color and design.

Basic problem defined - Conceptual process, mixing colors, mixing and adding colors to the swatches panel.

Module - 4

Value: Introduction, changing level of image, Creating the first drawing, Creating the simplified drawing, grouping sublayers, Changing the stacking order of groups in layers, Swatches panel.

Typeface Design: Introduction, Historic classification of type, Type styles, Visual and information Hierarchy, Type as an element of design summary.

Module - 5

Design in Visual Communication: Introduction, Communication and graphic design, Visual communication and visual design, Design in visual communication, Purpose of design in visual communication, Design products in visual communication for evaluation.

Text Book

1. Visual Design Fundamentals: A Digital Approach, Third Edition, Alan Hashimoto and Mike Clayton © 2009 Course Technology, a part of Cengage Learning.
2. Günay, M. (2021). Design in Visual Communication. Art and Design Review, 9, 109-122, Scientific research publishing.

References

1. Digital Texturing and Painting: Owen, Demers, New riders publishing ISBN: 0-7357-0918-1

Course Level Assessment

Questions Course Outcome 1 (CO1):

1. Explain focal point and visual hierarchy.
2. Describe the visual principle that a design is weighted equally.

Course Outcome 2 (CO2):

1. Explain the various methods of abstraction.
2. How to design software applications for graphics.

Course Outcome 3 (CO3):

1. Explain color unity in terms of Tonality.
2. Explain the different methods of mixing and adding colors to the swatches panel.

Course Outcome 4 (CO4):

1. How to create, delete and apply swatches.
2. Explain type as an element of design summary.

Course Outcome 5 (CO5):

1. Explain the importance of graphic design in visual communication.
2. Explain the different design communication products.

Course Outcome 6 (CO6):

1. Consider a technical fest in your institution. Design a poster using photoshop, based on the theme ‘Technology in society’.

Model Question Paper

QP CODE:

Reg No:_____

Name:_____ **PAGES: 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CXT 352

Course Name: Visual Design and Communication

Max. Marks :100 **Duration: 3 Hrs.**

PART A

(Answer all Questions. Each question carries 3 Marks)

1. Explain the concept behind focal point and visual hierarchy.
2. Explain any three basic elements of design.
3. Write a short note on Repetition using shape with example.
4. What do you mean by comps?
5. Explain the software applications for graphics.
6. Briefly describe the idea of abstraction.
7. Mention the impact of Bezier curves on visual design.
8. Explain Typeface design in visual design.
9. Differentiate between transparency and translucency.
10. What are sublayers? Also explain changing the stacking order of groups in layers.

Part B

(Answer any one question from each module. Each question carries 14 Marks)

Module I

11. a. Distinguish between rectilinear and curvilinear shapes. (6)
    b. Elaborate various elements of design in detail. (8)

OR

12. a. Explain Gestalt theory. What is its relevance in visual design? (8)
    b. Describe various types of rhythm with examples. (6)

Module II

13. a. Explain the steps in the conceptual process in design of a project. (8)
    b. Briefly illustrate the computer applications used in design. (6)

OR

14. a. Differentiate between bitmap graphics and vector graphics. (7)
    b. Explain the simplification method of abstraction. (7)

Module III

15. Narrate the background in color theory. (14)

OR

16. a. Write short notes on color unity via Tonality. (6)
    b. Describe mixing and adding colors to the swatches panel. (8)

Module IV

17. Explain Historic classification of type and Type styles. (14)

OR

18. Describe Visual and information Hierarchy. (14)

Module V

19.a. Distinguish Visual communication and Visual design. (8)
    b. Explain the need for design in visual communication. (6)

OR

20. Briefly describe: a) Graphic design.

    b) Design products in visual communication. (14)

(5 x 14 = 70)

Teaching Plan

No	Contents	No of Lecture Hrs. :35
Module I – (Design Elements) (6 hrs.)		
1.1	Design Elements: Introduction	1
1.2	Elements of Design- Line, Shape, Negative Space, Volume, Value, Color, Texture.	1
1.3	Principles of Design**:** Gestalt - Unity or Harmony	1
1.4	Unity and Placement using Line, shape, Repetition using line, shape.	1
1.5	Rhythm - Unity using value, Unity and continuity, Variety.	1
1.6	Focal point and visual hierarchy, Balance.	1
Module – 2 (Introduction to design projects) (8 hrs.)		
2.1	Introduction to design projects**:** Conceptual process - research and define problem, preliminary solutions	1
2.2	Organized list of options and thumbnail sketches	1
2.3	Rough- refined preliminary visual design solutions and alternatives	1
2.4	Composites or comps, the final design or finish	1
2.5	Production process- Computer applications used in design	1
2.6	Software applications for graphics.	1
2.7	Basic methods of Abstraction: Simplification, Repetition	1
2.8	Line and Shape, Type combination	1
Module - 3 (Color and color theory) (7 hrs.)		
3.1	Background in color theory - Defining color, Color and unity	1

3.2	Color unity via Tonality	1

3.3	Color and variety	1
3.4	Warm and cool colors, Color discord, Color and design	1
3.5	Basic problem defined - Conceptual process	1
3.6	Mixing colors	1
3.7	Mixing and adding colors to the swatches panel.	1
Module - 4 (Value and typeface design) (8 hrs.)		
4.1	Introduction, Changing level of image	1
4.2	Creating the first drawing, Creating the simplified drawing	1
4.3	Grouping sublayers, Changing the stacking order of groups in layers	1
4.4	Swatches panel	1
4.5	Typeface Design- Introduction	1
4.6	Historic classification of type	1
4.7	Type styles, Visual and information Hierarchy	1
4.8	Type as an element of design summary.	1
Module - 5 (Design in Visual Communication) (6 hrs.)		

5.1	Design in Visual Communication: Introduction	1
5.2	Communication and graphic design	1

5.3	Visual communication and visual design	1
5.4	Design in visual communication	1
5.5	Purpose of design in visual communication	1
5.6	Design products in visual communication	1

CXT 362	COMPUTER ARCHITECTURE	Category	L	T	P	Credits	Year of Introduction
		PEC	2	1	0	3	2021

Preamble:

This course helps the learner to understand the basic understanding of the parallel architecture and its operations and key features of high-performance computers. This course covers different parallel computer models, analyze the advanced processor technologies, compare different multiprocessor system interconnecting mechanisms, interpret the mechanisms for enforcing cache coherence, analyze different message passing mechanisms, analyze different pipelining techniques and appraise concepts of multithreaded architectures.

Prerequisite: Computer Organization and Architecture

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

CO#	Course Outcomes
CO1	Apply the basic concepts of various computer architectures, Amdahl's law for a fixed workload for evaluating the system performance. **(Cognitive Knowledge Level: Apply)**
CO2	Explain the concepts of various advanced processor technologies and conditions of parallelism. **(Cognitive Knowledge Level: Apply)**
CO3	Summarize different multiprocessor system interconnecting mechanisms and cache coherence mechanisms. **(Cognitive Knowledge Level: Understand)**
CO4	Articulate various message passing mechanisms and multithreaded architectures. **(Cognitive Knowledge Level: Understand)**
CO5	Apply the concepts of pipelining and pipeline design techniques to solve various problems. **(Cognitive Knowledge Level: Apply)**

Mapping of Course Outcomes with Program Outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✔	✔										✔
CO2	✔	✔	✔	✔								✔
CO3	✔	✔	✔									✔
CO4	✔	✔	✔									✔
CO5	✔	✔	✔	✔								✔

Abstract POs Defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination Marks (%)**
	Test 1 (%)	**Test 2 (%)**	
Remember	**30**	**30**	**30**
Understand	**40**	**40**	**40**
Apply	**30**	**30**	**30**

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
150	**50**	**100**	**3**

Continuous Internal Evaluation Pattern:

Attendance	**10 marks**
Continuous Assessment Tests (Average of Series Tests 1 & 2)	**25 marks**
Continuous Assessment Assignment	**15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which students should answer any one. Each question can have a maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module – 1 (Parallel Computer Models)

Parallel computer models − Evolution of Computer Architecture, System Attributes to performance, Amdahl's law for a fixed workload. Multiprocessors and Multicomputers, Multivector and SIMD computers.

Module - 2 (Advanced Processor Technologies**)**

Advanced Processor Technologies − Design Space of processors, Instruction Set Architectures, CISC Scalar Processors, RISC Scalar Processors, Superscalar and Vector processors, Conditions of parallelism.

Module - 3 (Multiprocessors System Interconnects, Cache Coherence Mechanisms**)**

Multiprocessors System Interconnects - Hierarchical bus systems, Crossbar switch and multiport memory, Multistage and combining networks. Cache Coherence Mechanisms, Cache Coherence Problem, Snoopy Bus Protocols.

Module - 4 (Message Passing Mechanisms and Multithreaded Architectures**)**

Message Passing Mechanisms - Message Routing schemes, Flow control Strategies, Dimension order routing. Multithreaded Architectures - Principles of multithreading, Multithreading Issues and Solutions, Multiple context processors.

Module - 5 (Pipelining Techniques and Pipeline Design**)**

Pipelining Techniques - Linear Pipeline processors and Nonlinear pipeline processors. Pipeline Design - Instruction pipeline design - Arithmetic pipeline design.

Text Book

1. K. Hwang and Naresh Jotwani, Advanced Computer Architecture, Parallelism, Scalability, Programmability, TMH, 2010.

References

1. H P Hayes, Computer Architecture and Organization, McGraw Hill, 1978.
2. K. Hwang & Briggs, Computer Architecture and Parallel Processing, McGraw Hill International, 1986.
3. M J Flynn, Computer Architecture: Pipelined and Parallel Processor Design, Narosa Publishing House, 2012.
4. M Sasikumar, D Shikkare and P Raviprakash, Introduction to Parallel Processing, PHI, 2014.
5. P M Kogge, The Architecture of Pipelined Computer, McGraw Hill, 1981.
6. P V S Rao, Computer System Architecture, PHI, 2009.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. A 400MHz processor was used to execute a program with 150000 floating point instructions with clock cycle count of 1. Determine the execution time and MIPS rate for this program.
2. A generalized multiprocessor system architecture combines features from the UMA, NUMA and COMA models. Justify the answer.

Course Outcome 2 (CO2):

1. Compare vector and super scalar processors.
2. Analyze the data dependences among the following statements and construct a dependency graph. Also detect the parallelism embedded in them.

   S1 : Load R1, M(100) / R1 ← Memory(100) /
   S2 : Move R2, R1 / R2 ← (R1) /
   S3 : Inc R1 / R1 ← (R1) + 1 /
   S4 : Add R2, R1 / R2 ← (R2) + (R1) /
   S5 : Store M(100), R1 / Memory(100) ← (R1) /

Course Outcome 3 (CO3):

1. Explain the three major operational characteristics of a multiprocessor interconnection network.
2. Differentiate write-invalidate and write-update coherence protocols for write through caches.

Course Outcome 4 (CO4):

1. Explain what are the various Message Routing schemes.
2. Explain what is a multithreaded architecture and also its principles.

Course Outcome 5 (CO5):

1. How Carry Save Adder (CSA) and Carry Propagate Adder (CPA) can be used in a fixed-point multiplication pipeline unit?
2. Compare linear and nonlinear pipelining processors.

Model Question Paper

QP CODE: **PAGES: 3**

Reg No: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CXT 362

Course Name: COMPUTER ARCHITECTURE

Max. Marks :100 **Duration: 3 Hrs.**

PART A

(Answer all Questions. Each question carries 3 Marks)

1. Explain the concept of Implicit parallelism.
2. State Amdahl's law. Suppose that we want to enhance the processor used for a server machine. The new processor is 15 times faster on computation in serving the applications than the original processor. Assuming the original processor is busy with computation 30% and waiting for I/O 70% time, what is the overall speedup gained by incorporating the enhancement?
3. Distinguish between scalar RISC and super-scalar RISC in terms of instruction issue, pipeline architecture and performance.
4. Explain Bernstein`s conditions for checking parallelism among a set of processes.
5. How does cache inconsistency occur in caches due to sharing writable data and process migration?
6. Describe a generalized multiprocessor system with a neat diagram.
7. Explain the various processor states in multiple context processors.
8. Explain different flow control strategies for resolving a collision between two packets.
9. What is dynamic instruction scheduling?
10. Explain internal data forwarding in pipelining.

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11 a. A 400MHz processor was used to execute a program with 150000 floating point instructions with clock cycle count of 1. Determine the execution time, CPI and MIPS rate for this program. (6)

b. Explain the two different models of the NUMA multiprocessor system. (8)

OR

12 a. Consider 2 programs A & B that solves a given problem. A is scheduled to run on a processor P1 operating at 1 GHz and B is scheduled to run on processor P2 running at 1.4 GHz. A has a total of 10000 instructions out of that 20% are branch instructions, 40% load and store and rest are ALU instructions. B is composed of 25% branch instructions. The number of load and store instructions in B is twice the count of ALU instructions. Total instruction count of B is 12000. In both P1 and P2 branch instructions have an average CPI of 5 and ALU instructions have an average CPI of 1.5. Both architectures differ in CPI for the load and store instructions, that is 2 for P1 and 3 for P2. Find out which mapping solves (A on P1 or B on P2) the problem faster and how much? (5)

b. Explain Flynn's classification of computer architecture with examples and diagrams. (9)

13 a. Consider the execution of the following code segment consisting of seven statements. Use Bernstein's conditions to detect the maximum parallelism embedded in this code. Justify the portions that can be executed in parallel and the remaining portions that must be executed sequentially. Rewrite the code using parallel constructs such as Cobegin and Coend. No variable substitution is allowed. All the statements can be executed in parallel if they are declared within the same block of a (Cobegin and Coend) pair.

S1: A=B+C
S2: C=D+E
S3: F=G+E
S4: C=A+F
S5: M=G+C
S6: A=L+E
S7: A=E+A (6)

b. With a neat diagram, explain about superscalar architecture. (8)

OR

14 a. Discuss pipelining in scalar, superscalar and vector processors with diagrams. (8)

b. Define the following terms and give examples: -

i) Flow dependence ii) Anti dependence iii) Output dependence (6)

15 a. Explain various snoopy bus protocols with examples and diagram. (9)

b. Explain the different levels of Hierarchical bus systems. (5)

OR

16 a. i. Draw a 16 input Omega network using 2x2 switches as building blocks.

ii. Show the switch settings for routing a message from node 1011 to node 0101 and from node 0111to node 1001 simultaneously. Does blocking exist in this case? (7)

b. With a neat sketch discuss on multiport memory. (7)

17 a. Explain the different context switching policies adopted by multithreaded architectures. (8)

b. Explain E-cube routing algorithm. Consider a 16-node hypercube network and show how to route a message from node (0111) to node (1101) using E-cube routing algorithm. All the intermediate nodes must be identified on the routing path. (6)

OR

18 a. Explain Multithreading issues and solutions. (6)

b. Discuss store and forward and wormhole routing schemes. Analyze and compare the communication latencies. (8)

19 a. Consider the following reservation table for a four-stage pipeline with a clock cycle τ=20ns.

	1	2	3	4	5	6
S1	x					x
S2		x		x		
S3			x			
S4				x	x	

i. What are the forbidden latencies and the initial collision vector?

ii. Draw the state transition diagram for scheduling the pipeline.

iii. Determine the MAL associated with the shortest greedy cycle.

iv. Determine the pipeline throughput corresponding to the MAL and given τ. (7)

b. Explain Arithmetic pipeline design with an example. (7)

OR

20 a. Explain in detail the effect of branching and various branch handling strategies. (9)

b. Explain the score boarding scheme employed by the CDC 6600 processor. (5)

Teaching Plan

No	Contents	No of Lecture Hrs.: 35
Module – 1 (Parallel Computer Models) (7 hrs.)		
1.1	Parallel computer models - Evolution of Computer Architecture	1
1.2	System Attributes to performance	1
1.3	System Attributes to performance	1
1.4	Amdahl's law for a fixed workload	1
1.5	Amdahl's law for a fixed workload	1
1.6	Multiprocessors and Multicomputers	1
1.7	Multivector and SIMD computers	1
Module – 2 (Advanced Processor Technologies) (7 hrs.)		
2.1	Advanced Processor Technologies - Design Space of processors	1
2.2	Instruction Set Architectures - CISC Scalar Processors	1
2.3	RISC Scalar Processors	1
2.4	Superscalar processors	1
2.5	Vector processors	1
2.6	Conditions of parallelism	1
2.7	Conditions of parallelism	1
Module – 3 (Multiprocessors System Interconnects, Cache Coherence Mechanisms) (7 hrs.)		
3.1	Multiprocessors system interconnects - Hierarchical bus systems	1
3.2	Crossbar switch and multiport memory	1
3.3	Multistage and combining networks	1
3.4	Multistage and combining networks	1
3.5	Cache Coherence Mechanisms- Cache Coherence Problem	1
3.6	Snoopy Bus Protocols	1
3.7	Snoopy Bus Protocols	1

Module - 4 (Message Passing Mechanisms and Multithreaded Architectures) (7 hrs.)		
4.1	Message Passing Mechanisms - Message Routing schemes	1
4.2	Flow control Strategies	1
4.3	Dimension order routing	1
4.4	Multithreaded architectures	1
4.5	Principles of Multithreading	1
4.6	Multithreading Issues and Solutions	1
4.7	Multiple context processors	1
Module - 5 (Pipelining Techniques and Pipeline Design) (7 hrs.)		
5.1	Pipelining techniques	1
5.2	Linear pipeline processors	1
5.3	Nonlinear pipeline processors	1
5.4	Nonlinear pipeline processors	1
5.5	Pipeline Design	1
5.6	Instruction pipeline design	1
5.7	Arithmetic pipeline design	1

CXT 308	COMPREHENSIVE COURSE WORK	Category	L	T	P	Credit	Year of Introduction
		PCC	1	0	0	1	2021

Preamble:

The objective of this Course work is to ensure the comprehensive knowledge of each student in the most fundamental core courses in the curriculum. Five core courses credited from Semesters 3, 4 and 5 are chosen for the detailed study in this course work. This course helps the learner to become competent in cracking GATE, placement tests and other competitive examinations

Prerequisite:

1. **Data Structures**
2. **Operating Systems**
3. **Computer Organization and Architecture**
4. **Web Programming**
5. **Virtual Reality**

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

CO1	Comprehend the concepts and applications of data structures (Cognitive Knowledge Level: **Understand)**
CO2 :	Comprehend the concepts, functions and algorithms in Operating System (Cognitive Knowledge Level: **Understand)**)
CO3:	Comprehend the organization and architecture of computer systems (Cognitive Knowledge Level: **Understand)**
CO4:	Comprehend the fundamental principles of Web Programming (Cognitive Knowledge Level: **Understand)**
CO5:	Comprehend the concepts in Virtual Reality (Cognitive Knowledge Level: **Understand)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓										✓
CO2	✓	✓										✓
CO3	✓	✓										✓
CO4	✓	✓										✓
CO5	✓	✓										✓

Assessment Pattern

Bloom's Category	End Semester Examination
Remember	**10**
Understand	**20**
Apply	**20**
Analyze	
Evaluate	
Create	

Mark distribution

Total Marks	CIE	ESE	ESE Duration
50	**0**	**50**	**1 hour**

End Semester Examination Pattern: Objective Questions with multiple choice (Four). Question paper include fifty questions of one mark each covering the five identified courses.

Syllabus

Full Syllabus of all five selected courses.

1. **Data Structures**
2. **Operating Systems**
3. **Computer Organization and Architecture**
4. **Web Programming**
5. **Virtual Reality**

Course Contents and Lecture Schedule

No	Topic	No. of Lectures (12)
1	**DATA STRUCTURES**	
1.1	Mock Test on Module 1, Module 2 and Module 3	1 hour
1.2	Mock Test on Module 4 and Module 5	1 hour
2	**OPERATING SYSTEMS**	
2.1	Mock Test on Module 1 and Module 2	1 hour
2.2	Mock Test on Module 3, Module 4 and Module 5	1 hour
2.3	Feedback and Remedial	1 hour
3	**COMPUTER ORGANIZATION AND ARCHITECTURE**	
3.1	Mock Test on Module 1, Module 2 and Module 3	1 hour
3.2	Mock Test on Module 4 and Module 5	1 hour
4	**WEB PROGRAMMING**	

4.1	Mock Test on Module 1, Module 2 and Module 3	1 hour
4.2	Mock Test on Module 4 and Module 5	1 hour
5	**VIRTUAL REALITY**	
5.1	Mock Test on Module 1, Module 2 and Module 3	1 hour
5.2	Mock Test on Module 4 and Module 5	1 hour
5.3	Feedback and Remedial	1 hour

Model Question Paper

QP CODE:

Reg No: _____

Name:_____ **PAGES: 10**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B. TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CXT 308

Course Name: Comprehensive Course Work

Max. Marks: 50 **Duration: 1 Hour**

Objective type questions with multiple choices. Mark one correct answer for each question. Each Question Carries 1 Mark

1. Consider the following sequence of operations on an empty stack.
   push(22); push(43); pop(); push(55); push(12); s=pop();

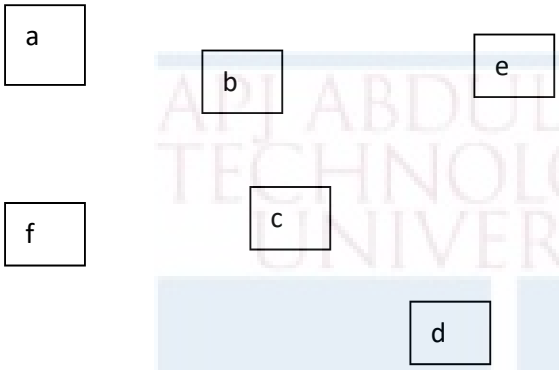
   Consider the following sequence of operations on an empty queue.
   enqueue(32);enqueue(27); dequeue(); enqueue(38); enqueue(12); q=dequeue();
   The value of s+q is _____
   (A) 44 (B) 54 (C) 39 (D) 70

2. The following postfix expression with single digit operands is evaluated using a stack:
   8 2 2 ^ / 4 3 * + 5 1 * -
   Note that ^ is the exponentiation operator. The top two elements of the stack after the first * is evaluated are:
   (A) 12,2 (B) 12,5 (C) 2,12 (D) 2,5

3. Construct a binary search tree by inserting 8, 6, 12, 3, 10, 9 one after another. To make the resulting tree as AVL tree which of the following is required?
   (A) One right rotation only
   (B) One left rotation followed by two right rotations
   (C) One left rotation and one right rotation
   (D) The resulting tree itself is AVL

4. In a complete 4-ary tree, every internal node has exactly 4 children or no child. The number of leaves in such a tree with 6 internal nodes is:
   (A) 20 (B) 18 (C) 19 (D) 17

5. Consider the following graph with the following sequences
   I. a b c f d e
   II. a b e d f c
   III. a b f c d e
   IV. a f c b e d



Which is Depth First Traversals of the above graph?
(A) I, II and IV only (B) I and IV only
(C) II, III and IV only (D) I, III and IV only

6. Consider a hash table of size seven, with starting index zero, and a hash function (2x + 5) mod7. Assuming the hash table is initially empty, which of the following is the contents of the table when the sequence 1, 4, 9, 6 is inserted into the table using closed hashing? Note that '_' denotes an empty location in the table.
   (A) 9, _, 1, 6, _, _, 4 (B) 1, _, 6, 9, _, _, 4
   (C) 4, _, 9, 6, _, _, 1 (D) 1, _, 9, 6, _, _, 4

7. Consider the following C program where TreeNode represents a node in a binary tree

```
struct TreeNode{
struct TreeNode *leftChild;
struct TreeNode *rightChild;
int element;
};
int CountNodes(struct TreeNode *t)
{
if((t==NULL)||((t->leftChild==NULL) && (t->rightChild==NULL)))
        return 0;
else
{
        return 1+CountNodes(t->leftChild)+CountNodes(t->rightChild)
}
}
```

The value returned by CountNodes when a pointer to the root of a binary tree is passed as its argument is

(A) number of nodes
(B) number of leaf nodes
(C) number of non-leaf nodes
(D) number of leaf nodes-number of non-leaf nodes

8. How many distinct binary search trees can be created out of 6 distinct keys?
   (A) 7 (B) 36 (C) 140 (D) 132

9. Suppose a disk has 400 cylinders, numbered from 0 to 399. At some time, the disk arm is at cylinder 58, and there is a queue of disk access requests for cylinder 66, 349, 201, 110, 38, 84, 226, 70, 86. If Shortest-Seek Time First (SSTF) is being used for scheduling the disk access, the request for cylinder 86 is serviced after servicing_____number of requests.
   (A) 1 (B) 2 (C)3 (D)4

10. If frame size is 4KB then a paging system with page table entry of 2 bytes can address _____bytes of physical memory.
    (A) 2^12 (B) 2^16 (C) 2^18 (D) 2^28

11. Calculate the internal fragmentation if page size is 4KB and process size is 103KB.
    (A) 3KB (B) 4KB (C) 1KB (D) 2KB

12. Which of the following scheduling policy is likely to improve interactiveness?
    (A) FCFS (B) Round Robin
    (C) Shortest Process Next (D) Priority Based Scheduling

13. Consider the following program

```
        Semaphore X=1, Y=0
Void A ( )                          Void B ( )
{                                   {
    While (1)                           While (1)
    {                                   {
    P(X);                               P(Y);
    Print'1';                           P(X);
    V(Y);                               Print'0';
    }                                   V(X);
}                                       }
                                    }
```

The possible output of the program:

(A) Any number of 0's followed by any number of 1's.
(B) Any number of 1's followed by any number of 0's.
(C) 0 followed by deadlock
(D) 1 followed by deadlock

14. In a system using single processor, a new process arrives at the rate of 12 processes per minute and each such process requires 5 seconds of service time. What is the percentage of CPU utilization?
(A) 41.66 (B) 100.00 (C) 240.00 (D) 60.00

15. A system has two processes and three identical resources. Each process needs a maximum of two resources. This could cause
(A) Deadlock is possible (B) Deadlock is not possible
(C) Starvation may be present (D) Thrashing

16. Which of the following is true with regard to Round Robin scheduling technique?
(A) Responds poorly to short process with small time quantum.
(B) Works like SJF for larger time quantum
(C) Does not use a prior knowledge of burst times of processes.
(D) Ensure that the ready queue is always of the same size.

17. The size of the physical address space of a 32-bit processor is 2^W words. The capacity of cache memory is 2^N words. The size of each cache block is 2^K words. For a M-way set-associative cache memory, the length (in number of bits) of the tag field is
(A) $W - N + \log_2 M$ (B) $W - N - \log_2 M$
(C) $W - N - K - \log_2 M$ (D) $W - N - K + \log_2 M$

18. A 64-bit processor can support a maximum memory of 8 GB, where the memory is word-addressable (one word is of 64 bits). The size of the address bus of the processor is at least
_____ bits.
(A) 30 (B) 31 (C) 32 (D) None

19. The stage delays in a 4-stage pipeline are 900, 450, 400 and 350 picoseconds. The first stage (with delay 900 picoseconds) is replaced with a functionally equivalent design involving two stages with respective delays 600 and 550 picoseconds. The throughput increase of the pipeline is_____percent.
(A) 38 (B) 30 (C) 58 (D) 50

20. Consider a direct mapped cache of size 256 Kilo words with block size 512 words. There are 6 bits in the tag. The number of bits in block (index) and word (offset) fields of physical address are is:
(A) block (index) field = 6 bits, word (offset) field = 9 bits

(B) block (index) field = 7 bits, word (offset) field = 8 bits
(C) block (index) field = 9 bits, word (offset) field = 9 bits
(D) block (index) field = 8 bits, word (offset) field = 8 bits

21. The memory unit of a computer has 1 Giga words of 64 bits each. The computer has instruction format, with 4 fields: an opcode field; a mode field to specify one of 12 addressing modes; a register address field to specify one of 48 registers; and a memory address field. If an instruction is 64 bits long, how large is the opcode field?
(A) 34 bits (B) 24 bits (C) 20 bits (D) 14 bits

22. A computer has 64-bit instructions and 28-bit address. Suppose there are 252 two-address instructions. How many 1-address instructions can be formulated?
(A) 2^24 (B) 2^26 (C) 2^28 (D) 2^30

23. Determine the number of clock cycles required to process 200 tasks in a six-segment pipeline. (Assume there were no stalls), each segment takes 1 cycle.
(A) 1200 cycles (B) 206 cycles (C) 207 cycles (D) 205 cycles

24. Match the following Lists:

P. DMA	1. Priority Interrupt
Q. Processor status Word	2. I/O Transfer
R. Daisy chaining	3. CPU
S. Handshaking	4. Asynchronous Data Transfer

(A) P-1, Q-3, R-4, S-2 (B) P-2, Q-3, R-1, S-4
(C) P-2, Q-1, R-3, S-4 (D) P-4, Q-3, R-1, S-2

25. What is the preferred way for adding a background color in HTML?
a.) <body background="yellow">
b.) <background>yellow</background>
c.) < body style="background-color:yellow">
d.) <background color="yellow">text<background>

26. Which of the following JavaScript cannot do?
a.) JavaScript can react to events
b.) JavaScript can manipulate HTML elements
c.) JavaScript can be used to validate data
d.) All of the Above

27. How can you make a list that lists the items with numbers?
a.)

b.) <list>

c.)

d.) <dl>

28. The latest HTML standard is

a) XML

b) HTML 4.0

c) HTML 5.0

d) HTML 6.0

29. Elements between and tags of HTML tables are ..... by default

a) left aligned

b) justified

c) Center aligned

d) right aligned

30. How do we write comments in HTML?

a) </ ..... >

b) <!……>

c) </……/>

d) < .....!>

31. Why were cookies designed?

a) for server-side programming

b) for client-side programming

c) both a and b

d) none

32. What are the types of lists available in HTML?

a) Ordered, Unordered lists

b) Bulleted, numbered lists

c) None

d) Named, Unnamed lists

33. typeof “null” in JavaScript is _____

a) number

b) string

c) object

d) undefined

34. Which method is not used for converting variables to number?

a) parseInt()

b) Number()

c) parseFloat()

d) valueOf()

35. HMD stands for?

A) Head Made Display
B) Head Mounted Display
C) Head Masked Display
D) Head Mounted Detection

36. An example of non-immersive VR device is

A) An iPad
B) An IMAX
C) A Screen Projection
D) A Virtual Reality Headset

37. What is one potential limitation of VR?

A. It can be isolating for the user
B. It can be expensive to set up and maintain
C. It requires a high level of technical knowledge to use
D. All of the above

38. Virtual reality provides

A. Sharp pictures
B. Individual audio
C. Preview of new films
D. Participatory experience

39. Which country government introduced the Industry 4.0 concept?

(a) Germany
(b) United States of America
(c) France
(d) Great Britain

40. ._____keep track of position.
a) Motion analyzers
b) Motion Trackers
c) HMD
d) SMD

41. How can we describe an array in the best possible way?

a. The Array shows a hierarchical structure.
b. Arrays are immutable.
c. Container that stores the elements of similar types
d. The Array is not a data structure

42. Which data structure is mainly used for implementing the recursive algorithm?

a. Queue
b. Stack
c. Binary tree
d. Linked list

43. What is a batch operating system?

a. Multiple individual tasks
b. Similar types of tasks are grouped together
c. Tasks operating at different systems
d. All of the above

44. Which one of the following isn't considered a real-time operating system?

A. PSOS
B. linuxRT
C. VRTX
D. Windows

45. With the help of_____we reduce the memory access time:
A. SDRAM
B. Cache
C. Heaps
D. Higher capacity RAMs

46. What is used to increase the apparent size of physical memory?
A. Disks
B. Hard-disk
C. Virtual memory

D. Secondary memory

47. Which of these memories would have the lowest access time in a system:

a. Main Memory

b. Magnetic Disk

c. Registers

e. Cache

48. In the Principle of locality, there is a justification of the use of:

a. DMA

b. Cache memory

c. Disk

d. Interrupts

49. Quick sort is also known as .....
A. insertion sort
B. tree sort
C. shell sort
D. partition and exchange sort

50. To represent hierarchical relationship between elements, which data structure is suitable?
A) Circular queue
B) Priority
C) Tree
D) Graph

ANSWER KEY: -

QNo	Ans. Key	QNo	Ans. Key	QNo	Ans. Key	QNo	Ans. Key	QNo	Ans. Key
1	(C)	11	(C)	21	(B)	31	(A)	41	(C)
2	(A)	12	(B)	22	(D)	32	(A)	42	(B)
3	(A)	13	(D)	23	(D)	33	(C)	43	(B)
4	(C)	14	(B)	24	(B)	34	(D)	44	(D)
5	(A)	15	(B)	25	(C)	35	(B)	45	(B)
6	(D)	16	(C)	26	(D)	36	(A)	46	(C)
7	(C)	17	(A)	27	(C)	37	(D)	47	(C)
8	(D)	18	(A)	28	(C)	38	(D)	48	(B)
9	(C)	19	(D)	29	(A)	39	(A)	49	(D)
10	(D)	20	(C)	30	(B)	40	(B)	50	(C)

CXL 332	CASE LAB	CATEGORY	L	T	P	CREDIT	YEAR OF INTRODUCTION
		PCC	0	0	3	2	2021

Preamble: Lab is aimed to provide hands-on experience with various aspects of Software Engineering and UML including requirements identification, DFD, behavioral and structural design using UML diagrams, implementation, and so on.

Prerequisite: Basic understanding of Computer Programming, Object Oriented Modelling and Design.

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

CO1	Prepare Software Requirement Specification document, Design document, Test cases and Software configuration and Risk management related document. **(Cognitive Knowledge Level: Apply)**
CO2	Develop function-oriented software design using appropriate open-source tools. **(Cognitive Knowledge Level: Apply)**
CO3	Develop object-oriented software design using appropriate open-source tools. **(Cognitive Knowledge Level: Apply)**
CO4	Develop Cost Estimation models using appropriate open source-tools. **(Cognitive Knowledge Level: Apply)**
CO5	Apply an Openproj tool to track the progress of the project. **(CognitiveKnowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO 5	PO 6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓				✓		✓		✓
CO2	✓	✓	✓	✓	✓			✓		✓		✓
CO3	✓	✓	✓	✓	✓			✓		✓		✓
CO4	✓	✓	✓	✓	✓			✓		✓		✓
CO5	✓	✓	✓	✓	✓			✓		✓		✓

Abstract POs defined by National Board of Accreditation			
PO#	Broad PO	PO#	Broad PO
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Test (Internal Exam)** ***Percentage***	**End Semester Examination** ***Percentage***
Remember	**20**	**20**
Understand	**20**	**20**
Apply	**60**	**60**
Analyse		
Evaluate		
Create		

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
150	**75**	**75**	**3** hours

Continuous Internal Evaluation Pattern:

Attendance	: **15 marks**
Continuous Evaluation in Lab	: **30 marks**
Continuous Assessment Test	: **15 marks**
Viva-voce	: **15 marks**

Internal Examination Pattern: The marks will be distributed as Algorithm 30 marks, Program 20 marks, Output 20 marks and Viva 30 marks. Total 100 marks which will be converted out of 15 while calculating Internal Evaluation marks.

End Semester Examination Pattern: The marks will be distributed as Algorithm 30 marks, Program 20 marks, Output 20 marks and Viva 30 marks. Total 100 marks will be converted out of 75 for End Semester Examination.

Fair Lab Record:

All Students attending the Computer Aided Software Engineering Lab should have a Fair Record. The fair record should be produced in the University Lab Examination. Every experiment conducted in the lab should be noted in the fair record. For every experiment in the fair record the right-hand page should contain Experiment Heading, Experiment Number, Date of Experiment, Aim of Experiment, Details of Experiment including algorithm and Result of Experiment. The left-hand page should contain a printout of the code used for the experiment and sample output obtained for a set of input.

Syllabus

1. Familiarization of System Requirement Specification (SRS) and related analysis documents as per the IEEE standards.
2. Demonstration of UML Tools.
3. Creation of design documents representing the complete design of the software system.
4. Application of COCOMO and Function Point (FP) model for the actual project that has been chosen.
5. Familiarization of CASE workbenches.

For university examinations students can use the references of the library. Questions should be created in such a way that it explores the students' understanding of the concept.

Operating system to use in lab: Linux
Software to use in lab: Any Open-Source Software (e.g.: LaTeX, PlantUML, etc.,)

PRACTICE QUESTIONS

List of Exercises/Experiments:

For any given case/ problem statement do the following:

1. Design an SRS document in line with the IEEE recommended standards.
2. Draw the use case diagram and specify the role of each of the actors. Also state the precondition, post condition and function of each use case.
3. Draw the activity diagram.
4. Identify the classes. Classify them as weak and strong classes and draw the class diagram.
5. Draw the sequence diagram for any two scenarios.
6. Design and Draw the collaboration diagram.
7. Design and Draw the state chart diagram.
8. Draw the component diagram.
9. Demonstrate forward engineering in java. (Model to code conversion)
10. Demonstrate reverse engineering in java. (Code to Model conversion)
11. Develop and draw the deployment diagram.
12. Develop programme/Tool to calculate effort and cost estimation using COCOMO model.

Note: *Students can be given a group micro project, so that they learn to work in a team environment. They can also be trained on project management tools.*

Reference Books

1. Bennett S., McRobb S. & Farmer R., Object Oriented Systems Analysis and Design using UML, Tata McGraw-Hill, second edition, 2004.
2. J. Greenspan and B. Bulger, MySQL/PHP Database Applications, M&T Books, 2008

CXD 334	MINI PROJECT	Category	L	T	P	Credit	Year of Introduction
		PCC	0	0	3	2	2021

Preamble:

The objective of this course is to apply the fundamental concepts of Computer Science and Engineering/Design principles for the effective development of an application/research project. This course helps the learners to practice the different steps to be followed in the software development process such as literature review and problem identification, preparation of Software Requirement Specification &Software Design Document (SDD), testing, development and deployment. Mini project enables the students to boost their skills, widen the horizon of thinking and their ability to resolve-real life problems.

Pre-requisite:

A sound knowledge in any programming language and fundamental concepts of Software Engineering.

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

CO#	CO
CO1	Identify technically and economically feasible problems. **(Cognitive KnowledgeLevel: Apply).**
CO2	Identify and survey the relevant literature for getting exposed to related solutions and get familiarized with software development processes. **(Cognitive Knowledge Level: Apply).**
CO3	Perform requirement analysis, identify design methodologies and develop adaptable & reusable solutions of minimal complexity by using modern tools & advanced programming techniques. (**Cognitive Knowledge Level: Apply).**
CO4	Prepare technical report and deliver presentation. **(Cognitive Knowledge Level:Apply).**
CO5	Apply engineering and management principles to achieve the goal of the project. **(Cognitive Knowledge Level: Apply).**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
CO2	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
CO3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**75**	**75**	

Continuous Internal Evaluation Pattern:

Attendance **10 marks**

Project Guide **15 marks**

Project Report **10 marks**

Evaluation by the Committee (will be evaluating the level of completion and demonstration of functionality/specifications, presentation, oral examination, work knowledge and involvement): **40 marks**

Student Groups with 3 or 4 members should identify a topic of interest in consultation with a Faculty/Advisor. Review the literature and gather information pertaining to the chosen topic. State the objectives and develop a methodology to achieve the objectives. Carryout the design/fabrication or develop codes/programs to achieve the objectives by strictly following steps specified in the teaching plan. Innovative design concepts, performance, scalability, reliability considerations, aesthetics/ergonomic, user experience and security aspects taken care of in the project shall be given due weight.

The progress of the mini project is evaluated based on a minimum of two reviews. The review committee may be constituted by the Head of the Department comprising HoD or a senior faculty member, Mini Project coordinator and project guide. The internal evaluation shall be made based on the progress/outcome of the project, reports and a viva-voce examination, conducted internally by a 3-member committee. A project report is required at the end of the semester. The project has to be demonstrated for its full design specifications.

End Semester Examination Pattern:

The marks will be distributed as

Presentation: **30 marks**

Demo : **20 marks**

Viva : **25 marks.**

Total : **75 marks**.

TEACHING PLAN

Students are expected to follow the following steps.

1. Review of Literature and Identification of a problem
2. Create an abstract with a problem statement, solution approach, technology stack, team, etc. and get department approval. Register Online course/ Collect study materials.
3. Create Software Requirements Specification (SRS Document)
4. Create Software Design Document (SDD). This may include designs like,
    a. System Architecture Design
    b. Application Architecture Design
    c. GUI Design (Mockups)
    d. API Design
    e. Database Design
    f. Technology Stack
5. Create Test Plan, Test Scenarios and Test Cases (Test Case Document) & Traceability Matrix
6. Create a Project Plan (with Modules, Tasks, Resources, Time schedule) [May use any project management tool for this] – Choose and follow agile or waterfall models.
7. Development
    a. Set coding standards
    b. Environment Setup
    c. Source Code Control Setup (Like Subversion(SVN), Git)
    d. Development
    e. Unit Testing
    f. Integration Testing
    g. Testing /Quality Assurance(QA)
        i. Functional Testing
        ii. Load Testing
        iii. Report Bugs
    h. Resolve Bugs & Retest

8. Deployment (of software from local development environment to a production environment)
9. Test Run & Get Results
10. Prepare Project Report

Guidelines for the Report preparation

A bonafide report on the mini project shall be submitted within one week after the final presentation. Minimum number of pages should be 40.

- Use Times New Roman font for the entire report – Chapter/Section Title – Times New Roman18, Bold; Heading 2 – Times New Roman16, Bold; Heading 3 – Times New Roman14, Bold; Body- Times New Roman 12, Normal.
- Line Spacing – Between Heading 2 – 3 lines, between lines in paragraph 1.5 lines.
- Alignments – Chapter/Section Title – Center, Heading 2 & 3 should be Left Aligned. Ensure that all body text is paragraph justified.
- Figures & Tables – Ensure that all Figures and Tables are suitably numbered and given proper names/headings. Write figure title under the figure and table title above the table.

- **Suggestive order of documentation:**

  i. Top Cover

  ii. Title page

  iii. Certification page

  iv. Acknowledgement

  v. Abstract

  vi. Table of Contents

  vii. List of Figures and Tables

  viii. Chapters

  ix. Appendices, if any

  x. References/Bibliography

SEMESTER VI

MINOR

CST 382	INTRODUCTION TO SOFTWARE TESTING	Category	L	T	P	Credits	Year of Introduction
		VAC	3	1	0	4	2019

Preamble:

This is a course in theoretical computer science that includes test cases for white-box, black-box, and grey-box approaches. This course describes the various techniques for test case design used to test software artifacts, including requirements, design, and code. The course includes different techniques for test case design based on graphs, programming language syntaxes and inputs. The course also covers symbolic execution using PEX tool.

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

CO1	List a range of different software testing techniques and be able to apply specific unit testing method to the projects using Junit**. (Cognitive Knowledge Level: Understand)**
CO2	Explain mutation testing method for a given piece of code to identify hidden defects that can't be detected using other testing methods. **(Cognitive Knowledge Level: Understand)**
CO3	Explain graph coverage criteria in terms of control flow graph and data flow graph for a given program. **(Cognitive Knowledge Level: Understand**)
CO4	Demonstrate the importance of black-box approaches in terms of domain and functional testing. **(Cognitive Knowledge Level: Understand)**
CO5	Illustrate the use of PEX tool with symbolic execution. **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO 9	PO10	PO11	PO12
CO1	✔	✔	✔									✔
CO2	✔	✔	✔	✔						✔		✔

CO3	✓	✓	✓	✓						✓		✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓	✓	✓						✓		✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	PO7	Environment and Sustainability
PO2	Problem Analysis	PO8	Ethics
PO3	Design/Development of solutions	PO9	Individual and team work
PO4	Conduct investigations of complex problems	PO10	Communication
PO5	Modern tool usage	PO11	Project Management and Finance
PO6	The Engineer and Society	PO12	Life long learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination Marks**
	Test 1 (Marks)	**Test 2 (Marks)**	
Remember	**30**	**30**	**30**
Understand	**40**	**40**	**40**
Apply	**30**	**30**	**30**
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : **10 marks**

Continuous Assessment - Test : **25 marks**

Continuous Assessment - Assignment : **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module - 1 (Introduction to Software Testing)

Some Popular Errors – Ariane 5, Therac 25, Intel Pentium Bug. What is Software testing? Why should it be tested? Software Quality, Role of Testing. Testing Process - Level 0 thinking, Level 1 thinking, Level 2 thinking, Level 3 thinking, Level 4 thinking. Software Testing Terminologies - Verification, Validation and Testing, Faults, Error and Bug, Test cases, Coverage Criteria. Types of Testing- Unit testing, integration testing, System testing, Acceptance testing, Beta testing, Functional testing, Stress testing, Performance testing, Usability testing and Regression testing. Testing Methods - Black Box testing, White Box testing, Grey Box testing.

Module - 2 (Unit Testing)

Concept of Unit testing. Static Unit testing. Dynamic Unit testing - Control Flow testing, Data Flow testing, Domain testing, Functional Program testing. Mutation testing - Mutation and Mutants, Mutation operators, Mutation score. Junit - Framework for Unit testing. Case Study - Mutation testing using Junit and Muclipse.

Module - 3 (Unit Testing - White Box Approaches)

Overview of Graph Coverage Criteria. Structural Graph Coverage Criteria - Node/vertex coverage, Edge coverage, Edge pair coverage, Path coverage, Complete path coverage, Prime path coverage, Complete round trip coverage, Simple round trip coverage. Data Flow Criteria - du paths, du pairs. Subsumption Relationships among Graph Coverage Criteria. Graph Coverage for Source Code - Control flow graphs for code, CFG: If statement, CFG: If statement with return, CFG: Switch-case, CFG: Loops, CFG: Exceptions (try-catch). Example program – Statistics. Graph Coverage for Design Elements - Call graphs and classes, Class inheritance testing: Coverage criteria, Coverage criteria on inheritance graph, Data flow at the design level, Inter-procedural DU pairs, Coupling du-pairs example. Example - Quadratic Root. Case Study - Graph Based testing using JUnit Framework.

Module - 4 (**Unit Testing - Black Box Approaches**)

Domain Testing / Input Space Partitioning - Partitions of a set. Input domain modelling - Interface-based approach, Functionality-based approach. Identifying values. Multiple partitions of the inputdomain - All Combinations Coverage (ACoC), Each Choice Coverage (ECC), Pair-wise Coverage, T-wise Coverage, Base Choice Coverage, Multiple Base Choices Coverage. TriTyp example. Functional Testing - Functional Testing Concepts of Howden. Functional testing - Important Steps. Types of Functional testing - Equivalence Class Partitioning, Boundary Value Analysis, Decision Tables, Random Testing. Case Study - Black Box testing approaches using JUnit.

Module - 5 (Grey Box Testing Approaches)

Introduction to Grey Box testing - Why Grey Box testing, Gray Box Methodology, Advantages and Disadvantages. Techniques of Grey Box Testing - Matrix Testing, Regression Testing, Orthogonal Array Testing or OAT, Pattern Testing. An Introduction to PEX - Parameterized Unit Testing, The Testing Problem. Symbolic Execution – Example, Symbolic execution tree. PEX application Case Study – PEX.

Text Books

1. Paul Ammann and JeffOffutt ,Introduction to Software Testing.
2. KshirasagarNaik and PriyadarshiTripathy, Software Testing And Quality Assurance: Theory And Practice.

Reference Materials

1. https://www.csc.ncsu.edu/academics/undergrad/honors/thesis/muclipsebinder.pdf - Muclipse tutorial.
2. King, James C, "Symbolic Execution and Program Testing", Association for Computing Machinery, July 1976.
3.

Sample Course Level Assessment Questions

Course Outcome 1 (CO1): Explain the following types of testing methods with examples.

(i) Balck-box testing.

(ii) White-box testing.

(iii) Grey-box testing.

Course Outcome 2 (CO2): Define 12 mutants for the following method *power*() using effective mutation operators. Try to use each mutation operator at least once. Approximately, how many mutants do you think there would be, if all mutants for *power*() were created?

```
public static int power (int left, int right)
{
//*************************************
// Raises Left to the power of Right
// precondition : Right >= 0
// postcondition: Returns Left**Right
//*************************************
        intrslt;
        rslt = Left;
```

```
if (Right == 0)
{
        rslt = 1;
}
else
{
        for (int i = 2; i <= Right; i++)
        rslt = rslt * Left;
}
return (rslt);
}
```

Course Outcome 3 (CO3): Draw the control flow graph and data flow graph of given piece of code.

```
public static double ReturnAverage(int value[],int AS, int MIN, int MAX){
/*
Function: ReturnAverageComputes the averageof all those numbers in the input array in
the positive range [MIN, MAX]. The maximumsize of the array is AS. But, the array size
could be smaller than AS in which case the endof input is represented by -999.
*/
int i, ti, tv, sum;
doubleav;
i = 0; ti = 0; tv = 0; sum = 0;
while (ti< AS && value[i] != -999) {
ti++;
if (value[i] >= MIN && value[i] <= MAX) {
tv++;
sum = sum + value[i];
}
i++;
}
if (tv> 0)
av = (double)sum/tv;
```

```
else
av = (double) -999;
return (av);
}
```

Course Outcome 4 (CO4): Explain the following with examples.

1. Input domain modelling.
2. All Combinations Coverage (ACoC)
3. Each Choice Coverage (ECC)
4. Pair-wise Coverage
5. T-wise Coverage
6. Base Choice Coverage
7. Multiple Base Choices Coverage.

Course Outcome 5 (CO5): Draw the symbolic execution tree for the following program code and explain the symbolic execution of testme (α1, α2).

```
1. int twice (int v) {
2.  return 2 ∗ v;
3.  }
4.  void testme (int x, int y ) {
5.  z = twice ( y);
6.  if ( z == x ){
7.  if ( x > y + 10)
8. ERROR;
9.  }
10.  }
11. int main() {
12.  x = sym input();
13.  y = sym input();
14. testme ( x , y);
15.  return(0);
16. }
```

Model Question Paper

QP CODE: **PAGES: 4**

Reg No:_____ **Name :_____**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION(MINOR), MONTH & YEAR

Course Code: CST 382

Course Name: Introduction to Software Testing

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks

1. Explain the differences between Validation and Verification.
2. Explain the differences between Fault, Error, and Bug?
3. Define Ground string, Mutation score, and Mutants.
4. What are the functions of Test driver and Test stubs in dynamic unit testing?
5. Define Node coverage, Edge coverage and Prime path coverage in a control flow graph.
6. What are du paths and du pairs in a data flow graph?
7. Explain the two approaches in input domain modelling.
8. Explain the difference between Equivalence Class Partitioning and Boundary Value Analysis.
9. Briefly explain three techniques of Grey box testing.
10. Explain the concept of symbolic execution with the help of a toy example.

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Explain the following types of testing
(i) Black Box testing (ii) White Box testing (iii) Grey Box testing **(14)**

(iv) Unit testing (v) Integration testing (vi) System testing (vii) Acceptance testing

OR

12. (a) Explain the following coverage criterias based on the code fragment given below. (i) Functional coverage (ii) Statement coverage (iii)Conditional coverage (iv)Branch coverage **(8)**

```
int foo (int x, int y){
    int z = 0;
    if ((x > 0) && (y > 0)){
    z = x;}
    return z;
    }
```

(b) Write positive and negative test cases for an ATM Machine? **(6)**

13. (a) Explain Dynamic unit test environment with a neat figure. **(8)**

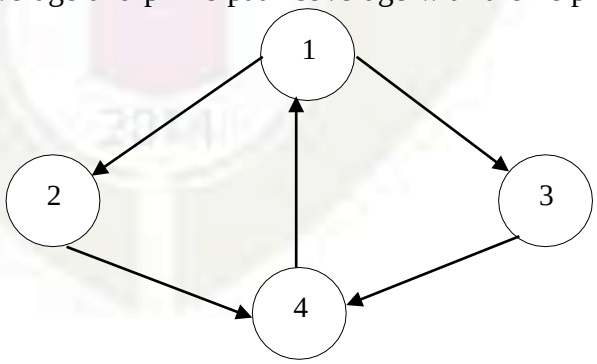
(b) Explain the major difference between control flow testing and data flow testing**.** **(6)**

OR

14. Explain seven types of mutation operators with neat examples. **(14)**

15. (a) Explain touring, side trips and detours with a neat example. **(7)**

(b) Explain simple path coverage and prime path coverage with the help of CFG given below. **(7)**



OR

16. (a) Draw CFG fragment for

(i) Simple *if* (ii) Simple *while* loop (iii) Simple *for* loop **(7)**

(b) Explain the following concepts with examples. **(7)**

(i) Call graph (ii) Inheritance graph (iii) Coupling du-pairs

17. (a) What are the four important steps in functional testing? **(7)**

(b) Briefly explain input domain modelling approaches. **(7)**

OR

18. (a) Consider the triangle classification program with a specification: **(6)**

The program reads floating values from the standard input. The three values *A*, *B*, and *C* are interpreted as representing the lengths of the sides of triangle. The program then prints a message to the standard output that states whether the triangle, if it can be formed, is scalene, isosceles, equilateral, or right angled. Determine the following for the above program:

(i) For the boundary condition $A + B > C$ case (scalene triangle), identify test cases to verify the boundary.

(ii) For the boundary condition $A = C$ case (isosceles triangle), identify test cases to verify the boundary.

(iii) For the boundary condition $A = B = C$ case (equilateral triangle), identify test cases to verify the boundary.

(b) Develop a decision table to generate test cases for this specification. **(8)**

19. (a) Explain the importance of grey box testing, its advantages and disadvantages. **(9)**

(b) Explain the concept of symbolic execution tree. **(5)**

OR

20. (a) Consider the code fragment given below: - **(7)**

```
1. POWER: PROCEDURE(X, Y);
2. Z ← 1;
3. J ← 1;
4. LAB:  IF Y ≥ J THEN
```

```
5. DO; Z← Z * X;
6. J ← J + 1;
7.  GO TO LAB;  END;
8.  RETURN (Z) ;
9.  END;
```

a) Explain Symbolic execution of POWER (αl, α2).

(b) Explain Execution tree for POWER (αl, α2) in the above code fragment. **(7)**

TEACHING PLAN

Index	**Topics**	**No. of Hours (45)**
	Module 1 (Introduction to Software Testing) 9 Hours	
1.1	Some Popular Errors– Ariane 5, Therac 25, Intel Pentium Bug.	1 Hour
1.2	What is Software testing? Why should it be tested? Software Quality, Role of Testing.	1 Hour
1.3	Testing Process - Level 0 thinking, Level 1 thinking, Level 2 thinking, Level 3 thinking, Level 4 thinking.	1 Hour
1.4	Software Testing Terminologies- Verification, Validation and Testing, Faults, Error and Bug, Test cases, Coverage Criteria.	1 Hour
1.5	Types of Testing- Unit testing, integration testing, System testing, Acceptance testing, Beta testing	1 Hour
1.6	Functional testing, Stress testing	1 Hour
1.7	Performance testing, Usability testing and Regression testing.	1 Hour
1.8	Testing Methods - Black Box testing	1 Hour
1.9	Grey Box testing.	1 Hour
Module 2 (Unit testing) 8 Hours		

2.1	Concept of Unit testing.	1 Hour
2.2	Static Unit testing.	1 Hour
2.3	Dynamic Unit testing - Control Flow testing, Data Flow testing	1 Hour
2.4	Domain testing, Functional Program testing.	
2.5	Mutation testing - Mutation and Mutants, Mutation operators, Mutation score.	1 Hour
2.6	Junit - Framework for Unit testing.	1 Hour
2.7	Case Study - Mutation testing using Junit	1 Hour
2.8	Case Study - Mutation testing using Muclipse	1 Hour
Module 3 (Unit Testing:- White Box Approaches) 10 Hours		
3.1	Overview of Graph Coverage Criteria	1 Hour
3.2	Structural Graph Coverage Criteria - Node/vertex coverage, Edge coverage, Edge pair coverage, Path coverage	1 Hour
3.3	Complete path coverage, Prime path coverage, Complete round trip coverage, Simple round trip coverage.	1 Hour
3.4	Data Flow Criteria - du paths, du pairs	1 Hour
3.5	Subsumption Relationships among Graph Coverage Criteria.	1 Hour
3.6	Graph Coverage for Source Code - Control Flow Graphs (CFG) for code, CFG: If statement, CFG: If statement with return, CFG: Switch-case, CFG: Loops, CFG: Exceptions (try-catch). Example program - Statistics	1 Hour
3.7	Graph Coverage for Design Elements - Call graphs and classes, Class inheritance testing: Coverage criteria, Coverage criteria on inheritance graph,	1 Hour

3.8	Data flow at the design level, Inter-procedural DU pairs, Coupling du-pairs example. Example - Quadratic Root	1 Hour
3.9	Case Study - Graph Based testing using JUnit Framework. (Lecture 1)	1 Hour
3.10	Case Study - Graph Based testing using JUnit Framework. (Lecture 2)	1 Hour
Module 4 (Unit Testing:- Black Box Approaches) 9 Hours		
4.1	Domain Testing / Input Space Partitioning - Partitions of a set.	1 Hour
4.2	Input domain modelling - Interface-based approach, Functionality-based approach.	1 Hour
4.3	Identifying values.	1 Hour
4.4	Multiple partitions of the input domain - All Combinations Coverage (ACoC), Each Choice Coverage (ECC), Pair-wise Coverage, T-wise Coverage, Base Choice Coverage, Multiple Base Choices Coverage.	1 Hour
4.5	TriTyp example.	1 Hour
4.6	Functional Testing - Functional Testing Concepts of Howden. Important Steps.	1 Hour
4.7	Types of Functional testing - Equivalence Class Partitioning, Boundary Value Analysis	1 Hour
4.8	Decision Tables, Random Testing.	1 Hour
4.9	Case Study - Black Box testing approaches using JUnit.	1 Hour
Module 5 (Grey Box Testing Approaches) 9 Hours		
5.1	Introduction to Grey Box testing - Why Grey Box testing, Gray Box Methodology, Advantages and Disadvantages.	1 Hour
5.2	Techniques of Grey Box Testing - Matrix Testing, Regression Testing,	1 Hour

5.3	Orthogonal Array Testing or OAT, Pattern Testing.	1 Hour
5.4	An Introduction to Pex - Parameterized Unit Testing, The Testing Problem.	1 Hour
5.5	Symbolic Execution – Example, Symbolic execution tree.	1 Hour
5.6	PEX application.	1 hour
5.7	Case Study – PEX (Lecture 1)	1 Hour
5.8	Case Study – PEX (Lecture 2)	1 Hour
5.9	Case Study – PEX (Lecture 3)	1 Hour

CST 384	CONCEPTS IN DEEP LEARNING	Category	L	T	P	Credits	Year of Introduction
		VAC	3	1	0	4	2019

Preamble:

This course aims to introduce the learner to an overview of the concepts and algorithms involved in deep learning. Deep learning is a subfield of machine learning, a subfield of artificial intelligence. Basic concepts and application areas of machine learning, deep networks, convolutional neural network and recurrent neural network are covered here. This is a foundational program that will help students understand the capabilities, challenges, and consequences of deep learning and prepare them to participate in the development of leading-edge AI technology. They will be able to gain the knowledge needed to take a definitive step in the world of AI.

Prerequisite: Sound knowledge in Basics of linear algebra and probability theory.

CO1	Demonstrate basic concepts in machine learning.(**Cognitive Knowledge Level: Understand)**
CO2	Illustrate the validation process of machine learning models using hyper-parameters and validation sets. **(Cognitive Knowledge Level: Understand)**
CO3	Demonstrate the concept of the feed forward neural network and its training process. **(Cognitive Knowledge Level: Apply)**
CO4	Build CNN and Recurrent Neural Network (RNN) models for different use cases. **(Cognitive Knowledge Level: Apply)**
CO5	Use different neural network/deep learning models for practical applications. **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓								✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓	✓							✓
CO4	✓	✓	✓	✓	✓	✓						✓
CO5	✓	✓	✓	✓	✓	✓						✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks
	Test1 (Percentage)	Test2 (Percentage)	
Remember	**30**	**30**	**30**
Understand	**40**	**40**	**40**
Apply	**30**	**30**	**30**
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance **: 10 marks**

Continuous Assessment Tests **: 25 marks**

Continuous Assessment Assignment : **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carry 14 marks.

Syllabus

INTRODUCTION TO DEEP LEARNING

(General Instructions: Instructors are to introduce students to any one software platform and demonstrate the working of the algorithms in the syllabus using suitable use cases and public datasets to give a better understanding of the concepts discussed. Tutorial hour may be used for this purpose)

Module-1 (Introduction)

Key components - Data, models, objective functions, optimization algorithms, Learning algorithm. Supervised learning- regression, classification, tagging, web search, page ranking, recommender systems, sequence learning, Unsupervised learning, Reinforcement learning, Historical Trends in Deep Learning. Other Concepts - overfitting, underfitting, hyperparameters and validation sets, estimators, bias and variance.

Module- 2 (Optimization and Neural Networks)

Neural Networks –Perceptron, Gradient Descent solution for Perceptron, Multilayer perceptron, activation functions, architecture design, chain rule, back propagation, gradient based learning. Introduction to optimization– Gradient based optimization, linear least squares. Stochastic gradient descent, Building ML algorithms and challenges.

Module -3 (Convolutional Neural Network)

Convolutional Neural Networks – convolution operation, motivation, pooling, Structure of CNN, Convolution and Pooling as an infinitely strong prior, variants of convolution functions, structured outputs, data types, efficient convolution algorithms. Practical challenges of common deep learning architectures- early stopping, parameter sharing, dropout. Case study: AlexNet, VGG, ResNet.

Module- 4 (Recurrent Neural Network)

Recurrent neural networks – Computational graphs, RNN design, encoder – decoder sequence to sequence architectures, deep recurrent networks, recursive neural networks, modern RNNs LSTM and GRU, Practical use cases for RNNs.

Module-5 (Application Areas)

Applications – computer vision, speech recognition, natural language processing, common word embedding: continuous Bag-of-Words, Word2Vec, global vectors for word representation (GloVe). Research Areas – autoencoders, representation learning, boltzmann machines, deep belief networks.

Text Book

1. Ian Goodfellow, YoshuaBengio, Aaron Courville, Deep Learning, MIT Press 2015 ed.
2. Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive into Deep Learning, August 2019.
3. Neural Networks and Deep Learning, Aggarwal, Charu C., c Springer International Publishing AG, part of Springer Nature 2018

Reference Books

1. Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks by Russell Reed, Robert J MarksII, A Bradford Book,2014
2. Practical Convolutional Neural Networks by MohitSewak, Md. Rezaul Karim, PradeepPujari,Packt Publishing 2018
3. Hands-On Deep Learning Algorithms with Python by SudharsanRavichandran,Packt Publishing 2019
4. Deep Learning with Python by Francois Chollet,Manning Publications Co.,2018

Sample Course Level Assessment Questions

Course Outcome 1 (CO1):

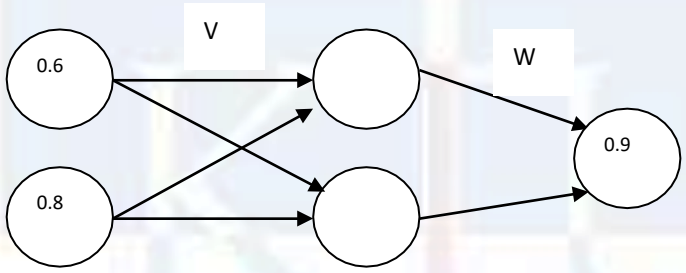
1. Compare regression and classification.
2. Define supervised learning? Distinguish between regression and classification.
3. Discuss the different learning approaches used in machine learning.

Course Outcome 2 (CO2):

1. What are hyperparameters? Why are they needed?
2. What issues are to be considered while selecting a model for applying machine learning in a given problem?

Course Outcome 3 (CO3):

1. Update the parameters V_{11} in the given MLP using back propagation with learning rate as 0.5 and activation function as sigmoid. Initial weights are given as V_{11}= 0.2, V_{12}=0.1, V_{21}=0.1, V_{22}=0.3, V_{11}=0.2, W_{11}=0.5, W_{21}=0.2



2. Draw the architecture of a multi-layer perceptron.
3. Derive update rules for parameters in the multi-layer neural network through the gradient descent.

Course Outcome 4 (CO4):

1. Give two benefits of using convolutional layers instead of fully connected ones for visual tasks.
2. Suppose that a CNN was trained to classify images into different categories. It performed well on a validation set that was taken from the same source as the training set but not on a testing set. What could be the problem with the training of such a CNN? How will you ascertain the problem? How can those problems be solved?
3. Explain how the cell state is updated in the LSTM model from Ct-1 to Ct
4. Show the steps involved in an LSTM to predict stock prices.

Course Outcome 5 (CO5):

1. Explain how the cell state is updated in the LSTM model from Ct-1 to Ct
2. Show the steps involved in an LSTM to predict stock prices.
3. Illustrate the workings of the RNN with an example of a single sequence defined on a vocabulary of four words.

Course Outcome 6 (CO6):

1. Development a deep learning solution for problems in the domain i) natural language processing or ii Computer vision (Assignment
2. Illustrate the workings of the RNN with an example of a single sequence defined on a vocabulary of four words.

Model Question Paper

QP CODE: PAGES:4

Reg No:_____

Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
SIXTH SEMESTER B.TECH DEGREE EXAMINATION(MINOR), MONTH & YEAR

Course Code: CST 384

Course Name: CONCEPTS IN DEEP LEARNING

Max. Marks:100 **Duration: 3 Hours**

PART A
Answer all Questions. Each question carries 3 Marks

1. Distinguish between supervised learning and Reinforcement learning. Illustrate with an example.

2. Differentiate classification and regression.

3. Compare overfitting and underfitting. How it can affect model generalization.

4. Why does a single perceptron cannot simulate simple XOR function? Explain how this limitation is overcome?

5. Illustrate the strengths and weaknesses of convolutional neural networks.

6. Illustrate convolution and pooling operation with an example

7. How many parameters are there in AlexNet? Why the dataset size (1.2 million) is important for the success of AlexNet?

8. Explain your understanding of unfolding a recursive or recurrent computation into a computational graph.

9. Illustrate the use of deep learning concepts in Speech Recognition.

10. What is an autoencoder? Give one application of an autoencoder

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E." What is your understanding of the terms task, performance and experience. Explain with two example **(10)**

(b) "How does bias and variance trade-off affect machine learning algorithms? **(4)**

OR

12. (a) Illustrate the concepts of Web search, Page Ranking, Recommender systems with suitable examples. **(10)**

(b) List and discuss the different hyper parameters used in fine tuning the **(4)**

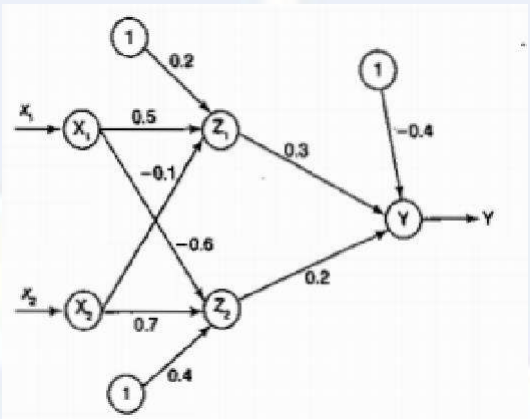
traditional machine learning models

13. (a) How multilayer neural networks learn and encode higher level features from input features. **(7)**

(b) Explain gradient decent and delta rule? Why stochastic approximation to gradient descent is needed? **(7)**

OR

14. (a) Find the new weights for the network using backpropogation algorithm, the network is given with a input pattern[-1,1] and target output as +1, Use learning rate of alpha=0.3 and bipolar sigmoid function. **(7)**



(b) Write an algorithm for backpropgation which uses stochastic gradient descent method. Comment on the effect of adding momentum to the network. **(7)**

15. (a) Input to CNN architecture is a color image of size 112x112x3. The first convolution layer comprises of 64 kernels of size 5x5 applied with a stride of 2 and padding 0. What will be the number of parameters? **(5)**

(b) Let X=[-1, 0, 3, 5] W=[.3 ,.5 .2,.1] be the the input of i^{th} layer of a neural network and to apply softmax function. What should be the output of it? **(4)**

(c) Draw and explain the architecture of convolutional network **(5)**

OR

16. (a) Explain the concept behind i) Early stopping ii) dropout iii) weight decay **(9)**

(b) How backpropagation is used to learn higher-order features in a convolutional Network? **(5)**

17. (a) Explain the working of RNN and discuss how backpropagation through time is used in recurrent networks. **(8)**

(b) Describe the working of a long short term memory in RNNs. **(6)**

OR

18. (a) What is the vanishing gradient problem and exploding gradient problem? **(8)**

(b) Why do RNNs have a tendency to suffer from exploding/vanishing gradient? How to overcome this challenge? **(6)**

19. (a) Explain any two word embedding techniques **(8)**

(b) Explain the merits and demerits of using Auto encoders in Computer Vision. **(6)**

OR

20. (a) Illustrate the use of representation learning in object classification. **(7)**

(b) Compare Boltzmann Machine with Deep Belief Network. **(7)**

Teaching Plan

CONCEPTS IN DEEP LEARNING (45 Hours)		
Module 1 : Introduction (9 hours)		
1.1	Key components - Data, models, objective functions, optimization algorithms. (TB2: Section 1.1-1.2)	1 hour

1.2	Learning algorithm (TB1: Section 5.1), Supervised learning- regression, classification (TB2: Section 1.3.1)	1 hour
1.3	tagging, web search, page ranking (TB2: Section 1.3.1)	1 hour
1.4	Recommender systems, Sequence learning, Unsupervised learning, Reinforcement learning(TB2: Section 1.3.2-1.3.4)	1 hour
1.5	Historical Trends in Deep Learning (TB1: Section 1.2).	1 hour
1.6	Concepts: over-fitting, under-fitting, hyperparameters and validation sets. (TB1: Section 5.2-5.3)	1 hour
1.7	Concepts: Estimators, bias and variance. (TB1: Section 5.4)	1 hour
1.8	Demonstrate the concepts of supervised learning algorithms using a suitable platform.	1 hour
1.9	Demonstrate the concepts of unsupervised using a suitable platform.	1 hour
Module 2 : Optimization and Neural Networks (9 hours)		
2.1	Perceptron, Stochastic Gradient descent, Gradient descent solution for perceptron (TB3: Section 1.1 - 1.2.1)	1 hour
2.2	Multilayer perceptron (TB3: Section 1.2.2), (TB1: Section 6.1,6.3)	1 hour
2.3	Activation functions- Sigmoid, tanh, Softmax, ReLU, leaky ReLU (TB3: Section 1.2.1.3 - 1.2.1.5)	1 hour
2.4	Architecture design (TB1: Section 6.4, TB3: Section 1.6)	1 hour
2.5	Chain rule, back propagation (TB3: Section 1.3)	1 hour

2.6	Gradient based learning (TB1: Section 6.2)	1 hour
2.7	Gradient based optimization (TB1: Section 4.3)	1 hour
2.8	Linear least squares using a suitable platform. (TB1: Section 4.5)	1 hour
2.9	Building ML Algorithms and Challenges (TB3: 1.4, TB1: 5.10-5.11)	1 hour
Module 3 :Convolution Neural Network (10 hours)		
3.1	Convolution operation, Motivation, pooling (TB1:Section 9.1-9.3)	1 hour
3.2	Structure of CNN (TB3: Section 8.2)	1 hour
3.3	Convolution and Pooling as an infinitely strong prior (TB1: Section 9.4)	1 hour
3.4	Variants of convolution functions – multilayer convolutional network, tensors, kernel flipping, downsampling, strides and zero padding. (TB1: Section 9.5)	1 hour
3.5	Variants of convolution functions - unshared convolutions, tiled convolution, training different networks. (TB1: Section 9.5)	1 hour
3.6	Structured outputs, data types (TB1: Section 9.6-9.7)	1 hour
3.7	Efficient convolution algorithms. (TB1: Section 9.8,9.10)	1 hour
3.8	Practical challenges of common deep learning architectures- early Stopping (TB3: 4.6)	1 hour
3.9	Practical challenges of common deep learning architectures- parameter sharing, drop-out (TB3: Section 4.9, 4.5.4)	1 hour
3.10	Case Study: AlexNet,VGG, ResNet. (TB3: Section 8.4.1-8.4.3,8.4.5)	1 hour

Module 4 :Recurrent Neural Network (8 hours)		
4.1	Computational graphs (TB1: Section 10.1)	1 hour
4.2	RNN (TB1: Section 10.2-10.3)	1 hour
4.3	Encoder – decoder sequence to sequence architectures. (TB1: Section 10.4)	1 hour
4.4	Deep recurrent networks (TB1: Section 10.5)	1 hour
4.5	Recursive neural networks , Modern RNNs, (TB1: Section 10.6, 10.10)	1 hour
4.6	LSTM and GRU (TB1: Section 10.10, TB3: Section 7.5-7.6)	1 hour
4.7	Practical use cases for RNNs. (TB1: Section 11.1-11.4)	1 hour
4.8	Demonstrate the concepts of RNN using a suitable platform.	1 hour
Module 5 : Applications and Research (9 hours)		
5.1	Computer vision. (TB1: Section 12.2)	1 hour
5.2	Speech recognition. (TB1: Section 12.3)	1 hour
5.3	Natural language processing. (TB1: Section 12.4)	1 hour
5.4	Common Word Embedding -: Continuous Bag-of-Words, Word2Vec (TB3: Section 2.6)	1 hour
5.5	Common Word Embedding -: Global Vectors for Word Representation(GloVe) (TB3: Section 2.9.1- Pennigton 2014)	1 hour
5.6	Brief introduction on current research areas- Autoencoders, Representation learning. (TB3: Section 4.10)	1 hour

5.7	Brief introduction on current research areas- representation learning. (TB3: Section 9.3)	1 hour
5.8	Brief introduction on current research areas- Boltzmann Machines, Deep belief networks. (TB1: Section 20.1, TB3 Section 6.3)	1 hour
5.9	Brief introduction on current research areas- Deep belief networks. (TB1: Section 20.3)	1 hour

CXT 386	WEB PROGRAMMING FOR GRAPHICS & GAMING	Category	L	T	P	Credit	Year of Introduction
		MINOR	3	1	0	4	2021

Preamble:

This is the course for awarding B.Tech. Minor in Computer Science and Design with specialization in Computer Graphics. The purpose of this course is to make awareness about the basic concepts in web design, standards and Java script. This course helps the learner to understand WebGL for rendering high-performance interactive 3D and 2D graphics in Game design. The study of computer graphics and related web application tools enables for creating various web-based Graphics applications.

Prerequisite: Knowledge of Programming is required.

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

CO#	CO
CO1	Summarize basic principles, rules and design of web site, considering the user requirements. **(CognitiveKnowledge level: Understand)**
CO2	Design websites adhering to current web standards (HTML, XHTML and HTML5). **(CognitiveKnowledge level: Apply)**
CO3	Construct visually formatted web pages using CSS. **(Cognitive Knowledge level: Apply)**
CO4	Use scripting language JavaScript to create interactive components in web pages. **(Cognitive Knowledge level: Apply)**
CO5	Use JavaScript API WEBGL for rendering high-performance interactive 3D and 2D graphics **(Cognitive Knowledge level: Apply)**

Mapping of Course Outcomes with Program Outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓										✓
CO2	✓	✓	✓		✓							✓
CO3	✓	✓	✓		✓							✓
CO4	✓	✓	✓		✓				✓			✓
CO5	✓	✓	✓	✓	✓							✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of Solutions	**PO9**	Individual and Team Work
PO4	Conduct Investigations of ComplexProblems	**PO10**	Communication
PO5	Modern Tool Usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong Learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3 hours

Continuous Internal Evaluation Pattern:

Attendance	**10 marks**
Continuous Assessment Tests (Average of Series Tests - 1& 2)	**25 marks**
Continuous Assessment Assignment	**15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question addingup to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one full question.Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module - 1 (Introduction to World Wide Web)

Introduction to world wide web, basic principles involved in developing a web site, rules of web designing, web standards, Web Browsers, Web Servers, Basic internet protocols, Uniform Resource Locators, Multipurpose Internet Mail Extensions, The Hypertext Transfer Protocol. Common Gateway Interface (CGI), Content Management System – Basics, Design concepts.

Module - 2 (Web Design using HTML5)

Clients, servers, introduction to Markup languages, scripting languages, Introduction to elements of HTML and XHTML, Introduction to Document object model (DOM).

Introduction to HTML5: Structuring & editing an HTML5 document, Fundamentals of HTML - Headings-Hyper Links- Images - Special Characters & Horizontal Rules-Lists- Tables -Forms - Internal Linking- Meta Elements-HTML5 Form input types -Input and Data List Elements and auto complete attribute- Page Structure Elements -Multimedia-HTML5 Audio & video elements.

Module - 3 (Introduction to CSS)

Introduction to Style sheets: Introduction to CSS-Basic syntax and structure-Inline Styles, Embedded Style Sheets, Conflict Resolution, Linking External Style Sheets-Exploring CSS Selectors-Properties, values, Positioning Elements: Absolute Positioning, Relative Positioning - Backgrounds-List Styles-Element Dimensions- Table Layouts-Box Model and Text Flow-div and span -Basics of Responsive CSS, Media port & Media Queries.

Module -4 (JavaScript Fundamentals)

Introduction to Scripting- Programming fundamentals of JavaScript - Data types, Values, Variables, Expressions and Operators. Obtaining User Input with prompt Dialog - Arithmetic-Decision Making -Control Statements - Functions -Arrays - Strings - Objects -Document Object Model (DOM) -Form processing. Difference between server side and client-side JavaScript, embedding JavaScript in HTML and frameworks, DOM and event handling, error handling, mouse, text, drag, drop and keyboard events and node operations, Animation and multimedia Forms of Debugging.

Module - 5 (WEBGL)

WEBGL overview, WEBGL application, Sample Application, Context, Geometry, Shaders, Associating Attributes and Buffer Objects, Drawing a Model.

WEBGL EXAMPLES: Drawing Points, Drawing a Triangle, Modes of Drawing.

Text Book

1. Paul J. Deitel, Harvey M. Deitel, Abbey Deitel, Internet & World Wide Web How to Program 5th Edition, Pearson Education.

2. Thomas A Powell, HTML & CSS: The Complete Reference, Tata McGraw Hill Publications, 5th Edition.
3. Kouichi Matsuda, Rodger Lea, WebGL Programming Guide: Interactive 3D Graphics Programming with WebGL, Pearson Education,

References

1. Scott Guelich, Shishir Gundavaram, Gunther Birzniek; CGI Programming with Perl 2/e, O'Reilly.
2. Doug Tidwell, James Snell, Pavel Kulchenko; Programming Web Services with SOAP, O' Reilly,2001
3. Robert. W. Sebesta, "Programming the World Wide Web", Fourth Edition, Pearson Education, 2007.

Sample Course Level Assessment Questions.

Course Outcome 1 (CO1):

1. What is MIME? Give the expansion of MIME. List four examples for MIME types. State the reason why MIME type specification is necessary in a request-response transaction between a browser and server.
2. Explain Hypertext Transfer Protocol and Common Gateway Interface in detail.

Course Outcome 2 (CO2):

1. Construct a valid HTML document for your personal Profile registration page for a Job Site www.123Jobs.com. Add relevant HTML elements in a table, to accept a minimum of 10 different fields which includes your name, address, phone, email address, your picture, your college; your branch, fields for your personal history (Minimum 3 fields), favorite theory and practical subjects (Checkbox), Username, Password(password).
2. What is codec? Recognize the role of controls attribute in <video> and <audio> tag in HTML. Use the COVID vaccination promotional video 'MySafety.mp4' in a web page with suitable HTML code, 'Autoplay' option enabled and displayed in a standard dimension 750 X500.

Course Outcome 3 (CO3):

1. Organize a sample web page for the event 'Raagam2021' at your campus and use embedded Style sheets to apply a minimum 5 styles. State the Style Specification format of embedded style sheets.
2. Write CSS style rules to implement the following in a web page:
   a. to display the content of hyperlinks with yellow background color and in italics.
   b. to display the contents of unordered lists in bold and in Arial font.
   c. to display a background image titled "birds.jpg" with no tiling.

Course Outcome 4(CO4):

1. Write the code for an HTML document with embedded JavaScript scripts, which initially displays a paragraph with text "Welcome" and a button titled "Click". When the button is clicked, the message "Hello from JavaScript" in bold should replace the paragraph text.
2. Illustrate the usage of JavaScript DOM in event handling and explain any three methods with example.

Course Outcome 5 (CO5):

1. What are the modes of drawing used in WEBGL? Illustrate with suitable examples.
2. Explain the structure of WEBGL application with an example code.

Model Question paper

	QP Code		**Total Pages :3**
Reg No._____		**Name:_____**	
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY **SIXTH SEMESTER B.TECH (MINOR) DEGREE EXAMINATION, MONTH and YEAR**			
Course Code: CXT 386			
Course Name: WEB PROGRAMMING FOR GRAPHICS & GAMING			
Max. Marks: 100		**Duration: 3 Hours**	
PART A			
		Answer all questions, each carries 3 marks.	Marks
1		Define WWW. List any two examples of web server & web browser. Differentiate between URL and a domain.	(3)
2		Explain any three internet protocols.	(3)
3		Explain the creation of hyperlinks in HTML.	(3)
4		Explain different types of list.	(3)
5		Explaintheuseof and.	(3)
6		Discuss the various CSS style sheet levels with suitable examples.	(3)
7		Illustrate the implementation of a JavaScript function greeting () using external .js file, to display a welcome message, when you click on a Button in an HTML page.	(3)
8		Explain arrays in JavaScript with example.	(3)
9		Explain vertex shaders.	(3)
10		Explainanythreedrawingmodessupportedby WebGL.	(3)

PART B

Answer any one Question from each module. Each question carries 14 Marks

11		Explain the design concepts of content management system.	(14)
		OR	
12		What are the basic principles and rules involved in developing a website?	(14)
13	a)	Design a web page that displays the following table.	(6)

Food Item	Recommended Intake			
	age<15		age>15	
	gm	Kcal	gm	Kcal
Cerials	1000	2000	750	1760
NonCerials	450	800	350	600

	b)	What is the difference between radio buttons and check-boxes when implemented using HTML? Write HTML code to implement a form which has the following elements: i. A text-box which can accept a maximum of 25 characters. ii. Three radio buttons with valid Label, Names and values. iii. Three check boxes buttons with valid Label, Names and values. iv. A selection list containing four items, two which are always visible. v. A submit button clicking on which will prompt the browser to send the form data to the server "http://www..mysite.com/reg.php" using "POST" method and reset button to clear its contents. You can use any text of your choice to label the form elements.	(8)
		OR	
14	a)	Write the equivalent HTML code to implement the following in a web page: (i) An image titled "birds.jpg" with a height of 100 pixels and width of 200 pixels. If the image cannot be accessed, a message "No image available" should be displayed (ii) A hyperlink to the URL "www.mysite.com/birds.jpg". The hyperlink should have the label "Click Here".	6
	b)	Create a static HTML document for your portfolio, which includes the following contents: your name, address, Mobile Number and email address. Also add the details about your college, university, your major and the batch of study. Include a picture of yourself and at least one other image (friend/pet/role model) to the document with a short description about that. Add three paragraphs about your personal history, with links to your social media profile. Also create an ordered list for describing your Skill Set & an unordered list showing your Strengths & Weaknesses.	(8)
15	a)	Write CSS and the corresponding HTML code for the following: i. Set the background color for the hover and active link states to "green". ii. Set the list style for unordered lists to "square". iii. Set "Flower.png" as the background image of the page and set 3% margin for the pages. iv. Set dashed border for left and right and double border for top & bottom of a table with 2 rows.	(8)
	b)	Write CSS style rules to implement the following in a web page: a. to display the content of hyperlinks with yellow background color and	(6)

16	a)	Write CSS and the corresponding HTML code for the following: i. Set the background color for the hover and active link states to "green". ii. Set the list style for unordered lists to "square". iii. Set "Flower.png" as the background image of the page and set 3% margin for the pages. iv. Set dashed border for left and right and double border for top & bottom of a table with 2 rows.	(8)
		in italics. b. to display the contents of unordered lists in bold and in Arial font. c. to display a background image titled “birds.jpg” with no tiling.	
		OR	
17	a)	List the order of precedence of style levels. Organize a sample web page for providing ‘KTU B.Tech Honours Regulation 19’ for KTU and use embedded Style sheet to apply minimum 5 styles for list, tables and pages.	(7)
	b)	Organize a sample web page for the event ‘Raagam2021’ at your campus and use embedded Style sheets to apply a minimum 5 styles. State the Style Specification format of embedded style sheets.	(7)
18	a)	Illustrate the different ways of Array declaration in JavaScript. Describe the function of the following JavaScript Array object methods with examples. (i) join (ii) slice	(7)
	b)	Write the code for an HTML document with embedded JavaScript scripts, which initially displays a paragraph with text "Welcome" and a button titled "Click". When the button is clicked, the message "Hello from JavaScript" in bold should replace the paragraph text.	(7)
		OR	
19		Illustrate the usage of JavaScript DOM in event handling and explain any three methods with example.	(14)
		Explain the different modes of drawing in WebGL with examples.	(14)
		OR	
20		With suitable example explain how scaling and rotation of objects can be performed using WebGL.	(14)

Teaching Plan		
No	**Topic**	**No. of Lectures (45)**
	Module-1 (Introduction to World Wide Web)	**7**
1.1	Introduction to world wide web	1
1.2	Basic principles involved in developing a web site, rules of web designing	1
1.3	Web standards, Web Browsers, Web Servers	1
1.4	Basic internet protocols, Uniform Resource Locators	1
1.5	Multipurpose Internet Mail Extensions	1
1.6	The Hypertext Transfer Protocol, Common Gateway Interface (CGI)	1
1.7	Content Management System – Basics, Design concepts	1
	Module-2 (Web Design using HTML5)	**9**
2.1	Clients, servers, Introduction to Markup languages, scripting languages	1
2.2	Introduction to elements of HTML and XHTML, Introduction to Document object model (DOM).	1
2.3	Introduction to HTML5: Structuring & editing an HTML5 document	1
2.4	Fundamentals of HTML - Headings-Hyper Links	1
2.5	Images - Special Characters & Horizontal Rules-Lists	1
2.6	Tables -Forms - Internal Linking	1
2.7	Meta Elements-HTML5 Form input types	1
2.8	Input and Data List Elements and auto complete attribute-Page Structure Elements	1
2.9	Multimedia-HTML5 Audio & video elements	1
	Module-3 (Introduction to CSS)	**9**

3.1	Introduction to Style sheets: Introduction to CSS-Basic syntax and structure	1
3.2	Inline Styles, Embedded Style Sheets	1
3.3	Conflict Resolution, Linking External Style Sheets	1
3.4	Exploring CSS Selectors-Properties, values	1
3.5	Positioning Elements: Absolute Positioning, Relative Positioning	1
3.6	Backgrounds-List Styles-Element Dimensions- Table Layouts-Box Model and Text Flow	1
3.7	div and span -Basics of Responsive CSS	1
3.8	Media port & Media Queries	1
3.9	Media port & Media Queries	1
	Module-4 (JavaScript Fundamentals)	10
4.1	Introduction to Scripting- Programming fundamentals of JavaScript	1
4.2	Data types, Values, Variables, Expressions and Operators	1
4.3	Obtaining User Input with prompt Dialog	1
4.4	Arithmetic-Decision Making -Control Statements	1
4.5	Functions -Arrays - Strings - Objects -Document Object Model (DOM)	1
4.6	Form processing. Difference between server side and client-side JavaScript,	1
4.7	Embedding JavaScript in HTML and frameworks	1
4.8	DOM and event handling, error handling	1
4.9	Mouse, text, drag, drop and keyboard events and node operations	1
4.10	Animation and multimedia Forms of Debugging	1
	Module-5 (WEBGL)	**10**
5.1	WEBGL overview	1

5.2	WEBGL application, Sample Application	1
5.3	Context, Geometry	1
5.4	Shaders, Associating Attributes and Buffer Objects	1
5.5	Shaders, Associating Attributes and Buffer Objects	1
5.6	Drawing a Model	1
5.7	WEBGL EXAMPLES: Drawing Points	1
5.8	Drawing a Triangle	1
5.9	Modes of Drawing	1
5.10	Examples	1

CST 394	NETWORK SECURITY	Category	L	T	P	Credits	Year of Introduction
		VAC	3	1	0	4	2019

Preamble:

The purpose of this course is to create a better understanding of the network security concepts. This course covers network security standards, email security services, web security mechanisms, firewalls and wireless security mechanisms. This course helps the learner to gain insight into the key aspects of secure network communication and enables to apply in real-life scenarios.

Prerequisite: A sound background in Number Theory and Cryptographic Algorithms.

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

CO#	Course Outcomes
CO1	Identify the key aspects of security, intrusion detection systems and digital signature schemes **(Cognitive Knowledge Level: Apply)**
CO2	Explain the security standards used in network communication (**Cognitive Knowledge Level:Understand)**
CO3	Identify the mechanisms in email security services (**Cognitive Knowledge Level: Apply)**
CO4	Summarize the protocols used to provide web security (**Cognitive Knowledge Level: Understand)**
CO5	Explain the fundamental concepts of wireless network security and firewalls **(Cognitive Knowledge Level: Understand)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓								✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓		✓						✓
CO4	✓	✓	✓	✓	✓	✓						✓
CO5	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	40	40	40
Apply	30	30	30
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance **: 10 marks**

Continuous Assessment Tests **: 25 marks**

Continuous Assessment Assignment **: 15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module – 1 (Network Security Basics)

Introduction to network security - Security requirements, Challenges of security, Network security model. Malicious programs – Worms, Viruses, Trojans, Spyware, Adware. Intrusion Detection Systems (IDS) - Uses, Techniques. Digital signatures - ElGamal, Schnorr, Digital Signature Standard (DSS).

Module – 2 (Network Security Standards)

Kerberos v4 – Configuration, Authentication, Encryption, Message formats. Kerberos v5 – Cryptographic algorithms, Message formats. Public Key Infrastructure (PKI) – Trust models, Revocation. Real-time communication security – Perfect Forward Secrecy (PFS), Denial-of-Service protection, Endpoint identifier hiding, Live partner reassurance. Internet Protocol Security (IPSec) - Authentication Header (AH), Encapsulating Security Payload (ESP), Internet Key Exchange (IKE) phases.

Module – 3 (Email Security)

Introduction to email security - Security services for email, Establishing keys, Privacy, Authentication, Message integrity, Non-repudiation. Privacy Enhanced Mail (PEM) – Encryption, Source authentication and integrity protection, Message formats. Secure/Multipurpose Internet Mail Extensions (S/MIME) – Messages, Differences from PEM. Pretty Good Privacy (PGP) - Encoding, Certificate and key revocation, Anomalies, Object formats.

Module – 4 (Web Security)

Introduction to web security - Web security considerations, Threats. Secure Sockets Layer (SSL) – Architecture, Protocols, Transport Layer Security (TLS) – Differences from SSL. Hypertext Transfer Protocol Secure (HTTPS) – Connection initiation, Closure. Secure Shell (SSH) – Transport layer protocol, User authentication protocol, Connection protocol.

Module – 5 (Wireless Network Security and Firewalls)

IEEE 802.11 Wireless LAN - Network components, Architectural model, Services. IEEE 802.11i wireless LAN security - Services, Phases of operation. Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA), WPA2, Wireless Application Protocol (WAP) – Services, Protocol architecture. Firewalls – Need for firewalls, Packet filters, Circuit-level firewalls, Application layer firewalls.

Text Books

1. C. Kaufman, R. Perlman and M. Speciner, “Network Security: Private Communication in a Public World”, 2/e, PHI.
2. William Stallings, “Cryptography and Network Security Principles and Practice”, 5/e, Pearson

Education Asia.

References

1. Behrouz A. Forouzan, Debdeep Mukhopadhyay, "Cryptography and Network Security", 3/e, Tata McGraw Hill.
2. Tyler Wrightson, "Wireless Network Security A Beginner's Guide", 2012, Tata McGraw Hill.
3. William Stallings, "Network Security Essentials: Applications and Standards", 4/e, Prentice Hall.
4. Schiller J., Mobile Communications, 2/e, Pearson Education.
5. Roberta Bragg et. al., "Network Security: The Complete Reference", Tata McGraw Hill.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Using the Schnorr digital signature scheme, let $q = 83$, $p = 997$ and $d = 23$. Find values for e_1 and e_2.
2. The Digital Signature Algorithm (DSA) specifies that if the signature generation process results in a value of zero, a new value of k should be generated and the signature should be recalculated. Give reason.

Course Outcome 2 (CO2):

1. In Kerberos v4, the authenticator field is not of security benefit when asking the Key Distribution Center (KDC) for a ticket for Bob, but useful when logging in as Bob. Give reasons for your answer.
2. How does the stateless cookie protocol provide clogging protection?

Course Outcome 3 (CO3):

1. If Alice is sending an ENCRYPTED message, she first signs the message digest with her private key and then encrypts the message digest with the pre-message secret key. Why this last encryption was considered necessary for encrypted messages and not for MIC-CLEAR or MIC-ONLY?
2. Which security services are considered desirable in the following cases? (i) Sending a purchase order (ii) Sending a ransom note. (iii) Sending a mission description to security officials.
3. Explain the security mechanism used in Gmail communication.

Course Outcome 4 (CO4):

1. Is it possible in SSL for the receiver to reorder SSL record blocks that arrive out of order? If so, how it can be done? If not, why?
2. Describe any five web security threats, their consequences and countermeasures.

Course Outcome 5 (CO5):

1. Explain the security areas addressed by IEEE 802.11i.
2. Describe the advantages and disadvantages of application layer firewalls.

Model Question Paper

QP CODE:
Reg. No:_____
Name:_____ **PAGES : 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH. DEGREE (HONORS) EXAMINATION, MONTH &YEAR
Course Code: CST 394

Course Name: Network Security

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks

1. Distinguish between signature-based and anomaly-based intrusion detection techniques.

2. A trusted third party is considered as a main component in a network security model. Why?

3. How is endpoint identifier hiding achieved in real-time communication?

4. Show how encryption is used to provide privacy and integrity in Kerberos v5.

5. End-to-end privacy is essential for e-mail security. How is this achieved?

6. List the four steps for preparing an EnvelopedData MIME entity.

7. Show the operation of a Secure Sockets Layer (SSL) Record protocol.

8. For Secure Shell (SSH) packets, what is the advantage of not including the MAC in the scope of packet encryption?

9. List the three security services provided by IEEE 802.11i.

10. Define the terms Access Point, Basic Service Set, Extended Service Set.

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Using the ElGamal scheme, let p = 881 and d = 700, find values for e1 and e2. Choose r = 17. Find the value of S1 and S2 if M = 400. **(8)**

(b) Explain the requirements and challenges of network security. **(6)**

OR

12. (a) In ElGamal, Schnorr and DSS, what happens if an attacker can find the value of random secret key used by the signer? Also, what happens if a user uses the same value of random secret key to sign two messages? Explain your answer for each scheme separately. **(8)**

(b) Explain the network security model with the help of a neat diagram. **(6)**

13. (a) Alice wishes to log into Bob's workstation remotely. List the steps involved in this communication if Kerberos v4 is used. **(7)**

(b) How does Diffie-Hellman technique provide perfect forward secrecy using signature keys? **(7)**

OR

14. (a) Explain the algorithm for Message Authentication Code (MAC) calculation and verification in Kerberos v5 rsa-md5-des. **(8)**

(b) Compare the aggressive mode and main mode of Phase 1 Internet Key Exchange (IKE). **(6)**

15. (a) Describe the different methods by which authentication of source is performed in email communication. **(7)**

(b) Explain the Signed data and Clear-signed data functions provided by S/MIME. **(7)**

OR

16. (a) Explain the advantages of Pretty Good Privacy (PGP) over Privacy Enhanced Mail (PEM). **(7)**

(b) Define non-repudiation. Describe the different ways by which it is implemented in email communication. **(7)**

17. (a) Describe the significance of pseudo-random function of Transport Layer Security. **(7)**

(b) Explain the four different phases of Secure Sockets Layer (SSL) HandshakeProtocol. **(7)**

OR

18. (a) Describe how connection initiation and connection closure is done in Hyper Text Transfer Protocol Secure (HTTPS). **(7)**

(b) Illustrate the sequence of events in Secure Shell (SSH) transport layer protocol packet exchanges. **(7)**

19. (a) Explain the Discovery phase and Authentication phase of IEEE 802.11i operation. **(7)**

(b) Why are firewalls needed? Compare the features of packet filters and circuit level firewalls. **(7)**

OR

20. (a) Explain the two authentication methods used in Wired Equivalent Privacy (WEP). **(7)**

(b) Describe the three transaction classes provided by Wireless Transaction Protocol. **(7)**

Teaching Plan

No	Contents	No of Lecture Hrs
Module - 1 (Network Security Basics) (7 hrs)		
1.1	Security requirements, Challenges of security	1
1.2	Network security model	1
1.3	Worms, Viruses, Trojans, Spyware, Adware	1
1.4	Intrusion Detection Systems (IDS) uses, Techniques	1
1.5	ElGamal digital signature	1
1.6	Schnorr digital signature	1
1.7	Digital Signature Standard (DSS)	1
Module - 2 (Network Security Standards) (12 hrs)		
2.1	Kerberos v4 configuration, Authentication	1
2.2	Kerberos v4 encryption	1
2.3	Kerberos v4 message formats	1
2.4	Kerberos v5 cryptographic algorithms – rsa-md5-des, des-mac, des-mac-k	1
2.5	Kerberos v5 cryptographic algorithms - rsa-md4-des, rsa-md4-des-k, Encryption for privacy and integrity	1
2.6	Kerberos v5 message formats	1
2.7	Public Key Infrastructure (PKI) trust models	1
2.8	PKI revocation	1
2.9	Perfect Forward Secrecy (PFS), Denial-of-Service protection	1
2.10	Endpoint identifier hiding, Live partner reassurance	1
2.11	Internet Protocol Security (IPSec) Authentication Header (AH), Encapsulating Security Payload (ESP)	1

2.12	Internet Key Exchange (IKE) phases	1
Module - 3 (Email Security) (9 hrs)		
3.1	Security services for email, Establishing keys, Privacy	1
3.2	Authentication, Message integrity, Non-repudiation	1
3.3	Privacy Enhanced Mail (PEM) encryption, Source authentication	1
3.4	PEM integrity protection, Message formats (Lecture 1)	1
3.5	PEM message formats (Lecture 2)	1
3.6	Secure/Multipurpose Internet Mail Extensions (S/MIME) – Messages, Differences from PEM	1
3.7	Pretty Good Privacy (PGP) encoding, Certificate and key revocation, Anomalies	1
3.8	PGP Object formats (Lecture 1)	1
3.9	PGP Object formats (Lecture 2)	1
Module – 4 (Web Security)(9 hrs)		
4.1	Web security considerations, Threats, Secure Sockets Layer (SSL) architecture	1
4.2	SSL protocols (Lecture 1)	1
4.3	SSL protocols (Lecture 2)	1
4.4	Transport Layer Security (TLS) differences from SSL (Lecture 1)	1
4.5	TLS differences from SSL (Lecture 2)	1
4.6	Hypertext Transfer Protocol Secure (HTTPS) connection initiation, Closure	1
4.7	Secure Shell (SSH) transport layer protocol	1
4.8	SSH user authentication protocol	1
4.9	SSH connection protocol	1

Module - 5 (Wireless Security and Firewalls) (8 hrs)		
5.1	IEEE 802.11 Wireless LAN network components, Architectural model, Services	1
5.2	IEEE 802.11i wireless LAN security services, Phases of operation (Lecture 1)	1
5.3	IEEE 802.11i phases of operation (Lecture 2)	1
5.4	Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA), WPA2	1
5.5	Wireless Application Protocol (WAP) services, Protocol architecture (Lecture 1)	1
5.6	WAP protocol architecture (Lecture 2)	1
5.7	Need for firewalls, Packet filters	1
5.8	Circuit-level firewalls, Application layer firewalls	1

CST 396	ADVANCED TOPICS IN MACHINE LEARNING	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2019

Preamble:

This course enables the learners to understand the advanced concepts and algorithms in machine learning. The course covers the standard and most popular supervised learning algorithms such as linear regression, logistic regression, decision trees, Bayesian learning and the naive Bayes algorithm, basic clustering algorithms, auto encoders, sampling methods and PAC learning. This course helps the students to provide machine learning based solutions to real world problems.

Prerequisite: Basic understanding of probability theory, linear algebra, multivariate calculus and multivariate probability theory.

CO1	Illustrate the concepts of regression and classification techniques **(Cognitive Knowledge Level: Apply)**
CO2	Demonstrate various unsupervised learning techniques **(Cognitive Knowledge Level: Apply)**
CO3	Choose suitable model parameters for different machine learning techniques and to evaluate a model performance **(Cognitive Knowledge Level: Apply)**
CO4	Explain the framework of PAC learning, basic concepts of VC dimension and non-uniform learnability **(Cognitive Knowledge Level: Understand)**
CO5	Construct Bayesian models for data and apply computational techniques to draw inferences **(Cognitive Knowledge Level: Apply)**
CO6	Illustrate the concepts of sampling algorithms, auto encoder, generative adversarial networks **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO 2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO1 2
CO1	✓	✓	✓	✓	✓	✓						✓
CO2	✓	✓	✓	✓	✓	✓						✓
CO3	✓	✓	✓	✓	✓	✓						✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓	✓	✓	✓							✓
CO6	✓	✓	✓	✓	✓	✓						✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks
	Test1 (Percentage)	Test2 (Percentage)	
Remember	**30**	**30**	**30**
Understand	**30**	**30**	**30**
Apply	**40**	**40**	**40**
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Tests : 25 marks

Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks

First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have a maximum 2 sub-divisions and carry 14 marks.

Syllabus

Module -1 (Supervised Learning)

Overview of machine learning - supervised, semi-supervised, unsupervised learning, reinforcement learning Regression algorithms: least squares linear regression, gradient descent, closed form, normal equations, regularization techniques (LASSO, RIDGE), polynomial regression. Discriminative Methods - Logistic Regression, Decision Tree Learning. Generative Methods - Naive Bayes Classifier, Gaussian Discriminant Analysis (GDA).

Module -2 (Unsupervised Learning)

Clustering - Similarity measures, Hierarchical Agglomerative Clustering, K-means partitional clustering, K-medoids clustering, Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model.

Module -3 (Practical aspects in machine learning)

Classification Performance measures - Precision, Recall, Accuracy, F-Measure, ROC, AUC, generalisation and overfitting, cross-validation, bias-variance tradeoff, error estimation, parameter and model selection. Ensemble Methods - Bagging, Boosting, Adaboost, Random Forests.

Module -4 (Statistical Learning Theory)

Models of learnability- learning in the limit, probably approximately correct (PAC) learning. Sample complexity- quantifying the number of examples needed to PAC learn, Computational complexity of training, Sample complexity for finite hypothesis spaces, PAC results for learning conjunctions, Sample complexity for infinite hypothesis spaces, Vapnik-Chervonenkis(VC) dimension.

Module -5 (Advanced Machine Learning Topics)

Graphical models - Bayesian belief networks, Markov random fields(MRFs), Inference on chains and factor graphs, inference on clique trees. Monte Carlo methods – Basic sampling algorithms, rejection sampling, importance sampling, Markov chain Monte Carlo(MCMC), Gibbs sampling. Variational methods. Auto Encoder, Variational AutoEncoder, Generative Adversarial Networks

Textbook

1. Christopher M. Bishop. Pattern recognition and machine learning. Springer 2006.
2. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition, MIT Press 2010.
3. Mohammed J. Zaki and Wagner Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, First South Asia edition, 2016.
4. Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press 2016.
5. Mehryar Mohri, Afshin Rostamizadeh and Ameet Talwalkar. Foundations of Machine Learning. Second edition. MIT Press 2018.
6. Tom Mitchell. Machine Learning. McGraw Hill 1997.
7. Richard O. Duda, Peter E . Hart, David G. Stork. Pattern classification, Second Edition. Wiley.
8. Jiawei Han, Micheline Kamber, Jian Pei. Data Mining Concepts and Techniques, Third Edition. Morgan Kaufmann.
9. David Foster. Generative Deep Learning - Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Inc., June 2019.

Reference Books

1. Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press 2012
2. Carl Edward Rasmussen and Christopher K. I. Williams. Gaussian Processes for Machine Learning. MIT Press 2005.

Sample Course Level Assessment Questions

Course Outcome1 (CO1):

1. Consider a naive Bayes classifier with 3 boolean input variables, $\mathbf{X_1}$, $\mathbf{X_2}$ and $\mathbf{X_3}$, 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?
2. Describe the ID3 algorithm. Is the order of attributes identical in all branches of the decision tree?
3. Explain the difference between (batch) gradient descent and stochastic gradient descent. Give an example of when you might prefer one over the other.
4. Suppose that you are asked to perform linear regression to learn the function that outputs y, given the D-dimensional input x. You are given N independent data points, and that all the D attributes are linearly independent. Assuming that D is around 100, would you prefer the closed form solution or gradient descent to estimate the regressor?
5. Suppose you have a three class problem where class label $y \in 0, 1, 2$ and each training example $\boldsymbol{X}$ has 3 binary attributes $\boldsymbol{X_1, X_2, X_3} \in \mathbf{0, 1}$. How many parameters (probability distribution) do you need to know to classify an example using the Naive Bayes classifier?

Course Outcome 2(CO2):

1. Which similarity measure could be used to compare feature vectors of two images? Justify your answer.
2. Illustrate the strength and weakness of k-means algorithm.
3. 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:

$C_1 = \{x_1, x_2, x_3\}$, $C_2 = \{x_4, x_5, x_6\}$, $C_3 = \{x_7, x_8\}$. Apply the **k**-means algorithm until convergence, using the Manhattan distance.

4. Cluster the following eight points representing locations into three clusters: $A_1(2, 10)$, $A_2(2, 5)$, $A_3(8, 4)$, $A_4(5, 8)$, $A_5(7, 5)$, $A_6(6, 4)$, $A_7(1, 2)$, $A_8(4, 9)$.

   Initial cluster centers are: $A_1(2, 10)$, $A_4(5, 8)$ and $A_7(1, 2)$.

   The distance function between two points $a = (x_1, y_1)$ and $b = (x_2, y_2)$ is defined as $D(a, b) = |x_2 - x_1| + |y_2 - y_1|$

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

Course Outcome 3(CO3):

1. What is ensemble learning? Can ensemble learning using linear classifiers learn classification of linearly non-separable sets?
2. Describe boosting. What is the relation between boosting and ensemble learning?
3. 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.
4. 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?
5. 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.

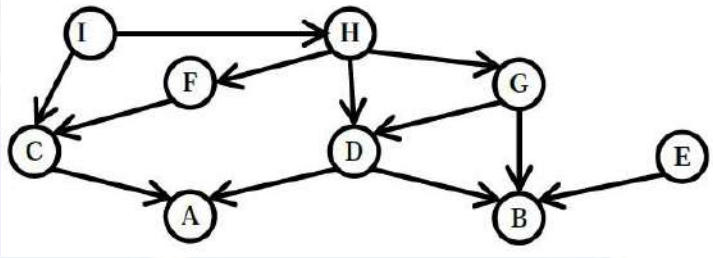
Course Outcome 4(CO4): .

1. A monotone conjunction is a conjunction of the variables such that no variable is negated. Show that the concept class of monotone conjunction is probably approximately correct (PAC)-learnable.
2. Consider a Boolean classification problem with n binary variables and a hypothesis space H, where each hypothesis is a decision tree of depth 2, using only two variables. How many training examples, m suffice to assure that with probability at least 0.99, any consistent learner using H will output a hypothesis with true error at most 0.05
3. Show that the concept class C containing the set of all boolean functions on n variable is not PAC-learnable.

4. What is the Vapnik-Chervonenkis(VC)-dimension of a circle centered at the origin.
5. A hypothesis space that has a high VC dimension is good, bad, or neither? Explain in terms of both (a) richness or expressive power of the hypotheses, and (b) sample complexity.

Course Outcome 5(CO5):

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



2. 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 6 (CO6):

1. Derive an algorithm using the inverse transform method to generate a random sample from the exponential distribution.
2. Explain the pros and cons of importance sampling versus rejection sampling.
3. 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?
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. The word “adversarial” in the acronym for GANs suggests a two-player game. What are the two players, and what are their respective goals?

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 5**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION (HONORS), MONTH & YEAR

Course Code: CST 396

Course Name: Advanced Topics in Machine Learning

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1.	Suppose you have a dataset with m = 1000000 examples and n = 200000 features for each example. You want to use multivariate linear regression to fit the parameters to our data. Should you prefer gradient descent or the normal equation? Justify your answer.	
2.	Define Information gain? How is that different from Gain ratio? Give the advantage of using Gain ratio measure?	
3.	What is cluster analysis? Identify two applications where cluster analysis can be applied to multimedia data?	
4.	Given two objects represented by the tuples (22, 1, 42, 10) and (20, 0, 36, 8): (i) Compute the Euclidean distance between the two objects. (ii) Compute the Manhattan distance between the two objects.	
5.	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.	
6.	How Bias-Variance Tradeoff affects machine learning algorithms?	
7.	For a particular learning task, if the requirement of error parameter ε changes from 0.1 to 0.01. How many more samples will be required for probably approximately correct(PAC) learning?	

8.		Suppose we have a hypothesis set that labels all points inside an interval ***[a, b]*** as class 1. Find its Vapnik-Chervonenkis(VC)- dimension?	
9.		Given a density function ***f(x)***, the rejection sampling is a method that can generate data points from the density function ***f***. List the three steps to generate a random sample from f using rejection sampling.	
10.		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?	**(10x3=30)**
		Part B **(Answer any one question from each module. Each question carries 14 Marks)**	
11.	(a)	. Consider the hypothesis for the linear regression $h_\theta(x) = \theta_0 + \theta_1 x$, and the cost function $J(\theta_0, \theta_1) = 1/2m \Sigma_{i=1 \text{ to } m} (h_\theta(x^{(i)}) - y^{(i)})^2$ where m is the number of training examples. Given the following set of training examples. . Answer the following questions : . 1) Find the value of $h_\theta(2)$ if $\theta_0 = 0$ and $\theta_1 = 1.5$. 2) Find the value of J(0,1) 3) Suppose the value of $J(\theta_0, \theta_1) = 0$. What can be inferred from this.	**(5)**

x	y
3	2
1	2
0	1
4	3

	(b)	. Write a gradient descent algorithm for multivariate regression? Compare the gradient and analytical solution to the multivariate regression?	**(9)**
		OR	
12.	(a)	Consider the collection of training samples (S) in the Figure given below. Drug is the target attribute which describes the Drug suggested for each patient. Find the value of the following . i) Gain(S, Sex) ii) Gain (S, Cholesterol)	**(9)**

		<table><tr><th>Patient ID</th><th>Age</th><th>Sex</th><th>BP</th><th>Cholesterol</th><th>Drug</th></tr><tr><td>p1</td><td>Young</td><td>F</td><td>High</td><td>Normal</td><td>Drug A</td></tr><tr><td>p2</td><td>Young</td><td>F</td><td>High</td><td>High</td><td>Drug A</td></tr><tr><td>p3</td><td>Middle-age</td><td>F</td><td>Hiigh</td><td>Normal</td><td>Drug B</td></tr><tr><td>p4</td><td>Senior</td><td>F</td><td>Normal</td><td>Normal</td><td>Drug B</td></tr><tr><td>p5</td><td>Senior</td><td>M</td><td>Low</td><td>Normal</td><td>Drug B</td></tr><tr><td>p6</td><td>Senior</td><td>M</td><td>Low</td><td>High</td><td>Drug A</td></tr><tr><td>p7</td><td>Middle-age</td><td>M</td><td>Low</td><td>High</td><td>Drug B</td></tr><tr><td>p8</td><td>Young</td><td>F</td><td>Normal</td><td>Normal</td><td>Drug A</td></tr><tr><td>p9</td><td>Young</td><td>M</td><td>Low</td><td>Normal</td><td>Drug B</td></tr><tr><td>p10</td><td>Senior</td><td>M</td><td>Normal</td><td>Normal</td><td>Drug B</td></tr><tr><td>p11</td><td>Young</td><td>M</td><td>Normal</td><td>High</td><td>Drug B</td></tr><tr><td>p12</td><td>Middle-age</td><td>F</td><td>Normal</td><td>High</td><td>Drug B</td></tr><tr><td>p13</td><td>Middle-age</td><td>M</td><td>High</td><td>Normal</td><td>Drug B</td></tr><tr><td>p14</td><td>Senior</td><td>F</td><td>Normal</td><td>High</td><td>Drug A</td></tr></table>	
	(b)	Explain how LASSO regression helps to reduce the overfitting problem?	**(5)**
13.	(a)	Suppose that we have the following data: <table><tr><td>*a*</td><td>*b*</td><td>*c*</td><td>*d*</td><td>*e*</td><td>*f*</td><td>*g*</td><td>*h*</td><td>*i*</td><td>*j*</td></tr><tr><td>(2,0)</td><td>(1,2)</td><td>(2,2)</td><td>(3,2)</td><td>(2,3)</td><td>(3,3)</td><td>(2,4)</td><td>(3,4)</td><td>(4,4)</td><td>(3,5)</td></tr></table> Identify the cluster by applying the k-means algorithm, with k = 2. Try using initial cluster centers as far apart as possible.	**(9)**
	(b)	Describe EM algorithm for Gaussian mixtures.	**(5)**
		OR	
14.	(a)	Illustrate the strength and weakness of k-means in comparison with the k-medoids algorithm.	**(4)**
	(b)	Suppose that we have the following data . Use single linkage Agglomerative clustering to identify the clusters. Draw the Dendogram. Use Euclidean distance measure	**(10)**

		<table><tr><th></th><th>X</th><th>Y</th></tr><tr><td>P1</td><td>0.4</td><td>0.53</td></tr><tr><td>P2</td><td>0.22</td><td>0.38</td></tr><tr><td>P3</td><td>0.35</td><td>0.32</td></tr><tr><td>P4</td><td>0.26</td><td>0.19</td></tr><tr><td>P5</td><td>0.08</td><td>0.41</td></tr><tr><td>P6</td><td>0.45</td><td>0.30</td></tr></table>	
15.	(a)	Define Precision, Recall, Accuracy and F-measure?	**(4)**
	(b)	What does it mean for a classifier to have a high precision but low recall?	**(3)**
	(c)	Fill in the missing values in the accompanying three class confusion matrix. Given that model accuracy is 72% and classification error for class 2 is 20%. Find also the precision and recall for class1 <table><tr><td colspan="2" rowspan="2"></td><td colspan="3">Predicted</td></tr><tr><td>Class 1</td><td>Class 2</td><td>Class 3</td></tr><tr><td rowspan="3">Actual</td><td>Class 1</td><td>14</td><td>2</td><td>5</td></tr><tr><td>Class 2</td><td>?(X)</td><td>40</td><td>2</td></tr><tr><td>Class 3</td><td>1</td><td>?(Y)</td><td>18</td></tr></table>	**(7)**
OR			

16.	(a)	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?	**(4)**
	(b)	Given the following ROC Curve? Find the AUC?	**(3)**
	(c)	How does random forest classifier work? Why is a random forest better than a decision tree?	**(7)**
17.	(a)	Show that the concept class Cn of the conjunction of boolean literals up to n variables is probably approximately correct(PAC)-learnable.	**(8)**
	(b)	Explain the concept of Vapnik-Chervonenkis (VC) dimension using shattering. How the number of training examples required to train the model is related to the VC dimension and what is its relation with training and test errors.	**(6)**
		OR	
18.	(a)	Consider a Boolean classification problem with ***n*** binary variables and a hypothesis space ***H***, where each hypothesis is a decision tree of depth 1. How many training examples, ***m*** suffice to assure that with probability at least 0.99, any consistent learner using ***H*** will output a hypothesis with true error at most 0.05?	**(7)**
	(b)	Prove that VC(H) ≤ log2 |H|, where H is a hypothesis space. (|H| denotes the	**(7)**

		cardinality of the hypothesis space)	
19.	(a)	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). Find P(J = t $\wedge$ M = f $\wedge$ A = f $\wedge$ B = f $\wedge$ E = t).	**(7)**
	(b)	Derive an algorithm using the inverse transform method to generate a random sample from the distribution with density $f_X(x) = 3x^2,\ 0 < x < 1$.	**(7)**
		OR	
20.	(a)	Draw 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)***	**(6)**
	(b)	What is effective sample size (ESS)? Why is a large ESS necessary but not sufficient for good MCMC mixing?	**(3)**
	(c)	Describe the overall Gibbs sampling algorithm briefly	**(5)**

Teaching Plan

Module 1 : (Supervised Learning)(10 hours)		
1.1	Supervised, semi-supervised, unsupervised learning, reinforcement learning (TB 2: Ch 1)	1 hour
1.2	Least squares linear regression (TB 2: Section 2.6)	1 hour
1.3	Gradient descent, closed form, normal equations (TB 2: Section 5.8)	1 hour
1.4	Regularization techniques (LASSO, RIDGE) (TB 4: Section 7.1)	1 hour
1.5	Polynomial regression (TB 2: Section 2.6)	1 hour
1.6	Logistic Regression (TB 6: Section 3.3)	1 hour
1.7	Decision Tree Learning (ID3) (TB 8: Section 8.2)	1 hour
1.8	Decision Tree Learning (C4.5) (TB 8: Section 8.2)	1 hour
1.9	Naive Bayes Classifier (TB 8: Section 8.3)	1 hour
1.10	Gaussian Discriminant Analysis (GDA) (TB 7: Section 5.2,5.3)	1 hour
Module 2 : (Unsupervised Learning)(8 hours)		
2.1	Similarity measures (TB 8: Section 2.4)	1 hour
2.2	Hierarchical Agglomerative Clustering (TB 3: Chapter 14)	1 hour
2.3	Hierarchical Agglomerative Clustering (TB 3: Chapter 14)	
2.4	K-means partitional clustering (TB 3: Chapter 13)	1 hour
2.5	K-medoids partitional clustering	
2.6	Gaussian mixture models (TB 3: Chapter 13)	1 hour
2.7	Expectation Maximization (EM) algorithm for Gaussian mixture model Lecture-1 (TB 3: Chapter 13)	1 hour
2.8	Expectation Maximization (EM) algorithm for Gaussian mixture model Lecture-2 (TB 3: Chapter 13)	1 hour
Module 3 : (Practical aspects in machine learning) (6 hours)		

3.1	Precision, Recall, Accuracy, F-Measure, ROC, AUC (TB8.5/TB 3: Chapter 22.1)	1 hour
3.2	Generalisation and overfitting, cross-validation (TB 2: Section 2.7,4.8)	1 hour
3.3	Bias-variance tradeoff (TB 2: Chapter 22.3)	1 hour
3.4	Error estimation, parameter and model selection (TB 3: Chapter 8.5)	1 hour
3.5	Bagging, Boosting (TB 8: Chapter 8.6)	1 hour
3.6	Adaboost, Random Forests (TB 8: Chapter 8.6)	1 hour
Module 4 : (Statistical Learning Theory) (TB 5 – Chapter 2, 3.3)(7 hours)		
4.1	Learning in the limit, probably approximately correct (PAC) learning	1 hour
4.2	Quantifying the number of examples needed to PAC learn	1 hour
4.3	Computational complexity of training	1 hour
4.4	Sample complexity for finite hypothesis spaces	1 hour
4.5	PAC results for learning conjunctions	1 hour
4.6	Sample complexity for infinite hypothesis spaces	1 hour
4.7	Vapnik-Chervonenkis(VC) dimension	1 hour
Module 5 : (Advanced Machine Learning Topics) (13 hours)		
5.1	Bayesian belief networks (TB 1 – Chapter 8)	1 hour
5.2	Markov random fields (TB 1 – Chapter 8)	1 hour
5.3	Inference on chains and factor graphs (TB 1 – Chapter 8)	1 hour
5.4	Inference on clique trees (TB 1 – Chapter 8)	1 hour
5.5	Basic sampling algorithms (TB 1 – Chapter 11)	1 hour
5.6	Rejection sampling (TB 1 – Chapter 11)	1 hour
5.7	Importance sampling (TB 1 – Chapter 11)	1 hour
5.8	Markov chain Monte Carlo(MCMC) (TB 1 – Chapter 11)	1 hour
5.9	Gibbs sampling (TB 1 – Chapter 11)	1 hour

5.10	Variational method (TB 1 – Chapter 10)	1 hour
5.11	Auto Encoder (TB 4 – Chapter 14)	1 hour
5.12	Variational AutoEncoder (TB 9 – Chapter 3)	1 hour
5.13	Generative Adversarial Networks (TB 9 – Chapter 4)	1 hour

CODE CXT 398	Data Visualization and Open-Source Programming for IOT	CATEGORY	L	T	P	CREDIT	Year of Introduction
		VAC	3	1	0	4	2021

Preamble:

The purpose of this course is to introduce Data Visualization and Open-Source Programming for IOT among learners. Concepts in this course help the learners to understand the visualization of data and programming required for IOT.

Prerequisite: Basic Concepts of IOT.

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

CO 1	Summarize data visualization and analytics concepts for the IoT. (Knowledge Level: **Understand**).
CO 2	Articulate various network analytics and edge streaming analytics in IOT. (Knowledge Level: **Understand).**
CO 3	Describe various IoT development boards. (Knowledge Level: **Understand).**
CO 4	Discuss data analysis and visualization using R programming and Python Programming. (Knowledge Level: **Understand**).
CO 5	Use Visualization techniques of R and Python to plot data related with IOT. (Knowledge Level: **Apply).**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO 1	√	√	√	√		√						√
CO 2	√	√	√	√								√
CO 3	√	√	√	√		√						√

CO 4	√	√	√	√	√	√						√
CO 5	√	√	√	√	√							√

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	**Continuous Assessment Tests**		**End Semester Examination Marks**
	Test 1	**Test 2**	
Remember	20%	20%	20%
Understand	40%	40%	40%
Apply	40%	40%	40%
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Tests : 25 marks

Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contain 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which student should answer anyone. Each question can have a maximum of 2 subdivisions and carries 14 marks.

Syllabus

Module 1 (Data visualization and Analytics for IoT). 9 hours

An Introduction to Data Analytics for IoT, Structured Versus Unstructured Data, Data in Motion Versus Data at Rest, IoT Data Analytics Overview, IoT Data Analytics Challenges, Machine Learning and Getting Intelligence from Big Data, Predictive Analytics. Types of data visualization techniques (e.g., bar charts, line charts, scatter plots, heatmaps), Choosing the right visualization technique for different types of data.

.

Module 2 (Network Analytics). 9 hours

Introduction to Network Analytics, Flexible NetFlow Architecture, FNF Components, Flexible NetFlow in Multiservice IoT Networks. Big Data Analytics Tools and Technology, Massively Parallel Processing Databases. Edge Streaming Analytics, Comparing Big Data and Edge Analytics, Edge Analytics Core Functions, Distributed Analytics Systems.

Module 3 (Development boards for IoT). 9 hours

Development boards for IoT. NodeMCU, Arduino, Raspberry Pi, NVIDIA Jetson. ESP32. Introduction to platform IO IDE. IoT Network architecture and design. Challenges, Different IoT architecture. Example programs for reading sensor values - DHT11 and TLS2561.

Module 4 (Data Visualization using R). 8 hours

Perception and Data Visualization - Introduction to R, Data structures in R, vectors, array, matrix and data frame, Managing graphics in R, Graphical functions, Reading flat files and csv files in R. Plot the data. Scatter plots using gplot. Histogram and density plots.

Module 5 (Data Visualization using Python). 9 hours

Introduction to Numerical python(numpy) and pandas. Visualization with Matplotlib- Line plots, scatter plots and histogram plots. Customizing Matplotlib: Configurations and Stylesheets. Three-Dimensional Plotting in Matplotlib. Plotting DHT11 sensor data using matplotlib.

Text Books

1. IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things" by David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Robert Barton, Cisco Press.2017.
2. *Introduction to IoT*. Misra, S., Mukherjee, A., & Roy, A. (2021)
3. Data Visualization: A Practical Introduction" , Kieran Healy
4. Python Data Science Handbook: Essential Tools for Working with Data" by Jake VanderPlas

References

1. Developing IoT Projects with ESP32 , Vedat Ozan Oner
2. LoRaWAN for IoT: Complete Guide to Developing LoRaWAN Solutions for Smart Cities, Agriculture, and Industry 4.0" , Manish Jha
3. Storytelling with Data: A Data Visualization Guide for Business Professionals , Cole Nussbaumer Knaflic
4. Data Visualization Made Simple, Kristen Sosulski

Sample Course Level Assessment Questions.

Course Outcome 1 (CO1):

1. Explain various data visualization techniques.
2. In the context of data analytics in IOT explain the differences between structured and unstructured data.

Course Outcome 2 (CO2):

1. Explain edge analytics processing unit.
2. What are IoT data analytics challenges? Explain.

Course Outcome 3 (CO3):

1. Discuss the advantages and limitations of using the NVIDIA Jetson development board for IoT applications.
2. How the Raspberry Pi can be used as a development board for IoT projects? Explain.

Course Outcome 4 (CO4):

1. Explain the importance of edges in data visualization and how they contribute to the perception of patterns and shapes.
2. Describe the characteristics of histogram and density plots in visualizing the distribution of data, including their ability to show frequency or probability density.
3. Explain the purpose and applications of three-dimensional (3D) plotting in data visualization.

Course Outcome 5 (CO5):

1. Design an R or Python application that incorporates geolocation data from IoT devices. Use appropriate mapping libraries, such as leaflet (in R) or folium (in Python), to plot the IoT devices on a map and visualize their spatial distribution. Explore additional visualization techniques, such as clustering or heatmaps, to analyze spatial patterns in the IoT data.
2. Develop a Python script that integrates data from IoT devices with external data sources, such as weather data or social media feeds.

Model Question paper

	QP Code :		
Reg No.:_____		**Name:**	
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY **VI SEMESTER B.TECH (HONOURS) DEGREE EXAMINATION, MONTH and YEAR**			
Course Code: CXT 398			
Course Name: Data Visualization and Open-Source Programming for IOT			
Max. Marks: 100			**Duration: 3 Hours**

PART A

		Answer all questions, each carries 3 marks.	Marks
1		In the context of data analytics in IOT explain the differences between structured and unstructured data	3
2		Discuss the advantages and limitations of using the NVIDIA Jetson development board for IoT applications.	3
3		Explain how analytics relates to IoT data.	3
4		Describe flexible-netFlow in multiservice IoT networks.	3
5		Discuss the Scaling problems in IoT data analytics.	3
6		How does the Arduino Uno compare to other Arduino models?	3
7		Compare and contrast the characteristics and use cases of histogram plots and density plots in visualizing data distributions in R.	3
8		Discuss the steps involved in importing a CSV file into R.	3

9		Explain how pandas provides data structures and functions for working with structured data.	3
10		Discuss the importance of choosing appropriate axis labels and titles when creating plots with Matplotlib.	3

		PART B	
		Answer any one Question from each module. Each question carries 14 Marks	
11	a)	Explain edge analytics processing unit.	(4)
	b)	What are IoT data analytics challenges? Explain.	(6)
		OR	
12	a)	Explain various types of data analysis results.	(7)
	b)	Illustrates the four data analysis types and how they rank as complexity and value increase.	(7)
13	a)	Provide an example program using the DHT11 sensor to read temperature and humidity values in an IoT project.	(7)
	b)	Explain in detail the common applications of ML for IoT.	(7)
		OR	
14	a)	In the context of IoT infrastructure deployments and technologies, explain the benefits of flow analytics in addition to other network management services.	(6)
	b)	Illustrate field area network (FAN) traffic analytics performed on the aggregation router in a smart grid.	(8)
15	a)	Explain how NodeMCU is suitable for prototyping IoT projects and its programming language compatibility.	(8)

	b)	Describe the steps involved in setting up a new ESP32 project using Platform IO IDE	(6)
		OR	
16	a)	Explain the steps involved in using the TLS2561 sensor to measure light intensity and retrieve the values in an IoT application.	(8)
	b)	Compare and contrast the features and capabilities of NodeMCU, Arduino, Raspberry Pi, and NVIDIA Jetson as development boards for IoT.	(6)
17	a)	Compare and contrast the advantages and limitations of different data formats (e.g., CSV, JSON, Excel) when importing data into R.	(6)
	b)	Explain the steps involved in creating scatter plots using the "gplot" function in R, including specifying variables and customizing the plot.	(8)
		OR	
18	a)	Write an R program to read a csv file with two numeric columns and plot the data using scatter plot.	(6)
	b)	Design a comprehensive data visualization project for IoT data using R, considering multiple IoT devices or sensors. Incorporate features like interactive elements, data aggregation, or anomaly detection to enhance the usability and insights derived from the visualizations.	(8)
19	a)	Describe the main features of pandas library in Python and how it facilitates data analysis and manipulation through its DataFrame data structure.	(8)
	b)	Describe the steps involved in reading DHT11 sensor data into Python using appropriate libraries or modules.	(6)
		OR	
20	a)	Explain the purpose of line plots in data visualization and discuss how they can be used to represent trends or patterns over time or continuous variables.	(8)
	b)	Discuss the different types of 3D plots available in Matplotlib	(6)
		***	

Teaching Plan		
No	**Topic**	**No. of Lectures (44 Hrs)**
	Module-1 (Data visualization and Analytics for IoT)	**9 hrs**
1.1	An Introduction to Data Analytics for IoT	1
1.2	Structured Versus Unstructured Data	1
1.3	Data in Motion Versus Data at Rest	1
1.4	IoT Data Analytics Overview, IoT Data Analytics Challenges	1
1.5	Machine Learning and Getting Intelligence from Big Data	1
1.6	Predictive Analytics	1
1.7	Types of data visualization techniques (e.g., bar charts, line charts, scatter plots, heatmaps)	1
1.8	Choosing the right visualization technique for different types of data	1
1.9	Choosing the right visualization technique for different types of data	1
	Module-2 (Network Analytics)	**9 hrs**
2.1	Introduction to Network Analytics	1
2.2	Flexible NetFlow Architecture	1
2.3	FNF Components, Flexible NetFlow in Multiservice IoT Networks	1
2.4	Big Data Analytics Tools and Technology	1

2.5	Massively Parallel Processing Databases	1
2.6	Edge Streaming Analytics	1
2.7	Comparing Big Data and Edge Analytics	1
2.8	Edge Analytics Core Functions, Distributed Analytics Systems	1
2.9	Edge Analytics Core Functions, Distributed Analytics Systems	1
	Module-3 (Development boards for IoT)	**9 hrs**
3.1	Development boards for IoT	1
3.2	NodeMCU, Arduino	1
3.3	Raspberry Pi	1
3.4	NVIDIA Jetson. ESP32	1
3.5	Introduction to platform IO IDE. .	1
3.6	IoT Network architecture and design. Challenges, Different IoT architecture.	1
3.7	IoT Network architecture and design. Challenges, Different IoT architecture.	1
3.8	Example programs for reading sensor values - DHT11 and TLS2561	1
3.9	Example programs for reading sensor values - DHT11 and TLS2561	1

	Module-4 (Data Visualization using R)	**8 hrs**
4.1	Perception and Data Visualization	1
4.2	Introduction to R	1
4.3	Data structures in R, vectors, array	1
4.4	matrix and data frame	1
4.5	Managing graphics in R, Graphical functions.	1
4.6	Reading flat files and csv files in R	1
4.7	Plot the data. Scatter plots using gplot. Histogram and density plots.	1
4.8	Plot the data. Scatter plots using gplot. Histogram and density plots.	1
	Module-5 (Data Visualization using Python)	9 hrs
5.1	Introduction to Numerical python(numpy) and pandas.	1
5.2	Visualization with Matplotlib- Line plots ,scatter plots and histogram plots. (Lecture 1)	1
5.3	Visualization with Matplotlib- Line plots ,scatter plots and histogram plots. (Lecture 2)	1
5.4	Visualization with Matplotlib- Line plots ,scatter plots and histogram plots. (Lecture 3)	1
5.5	Customizing Matplotlib: Configurations and Stylesheets	1

5.6	Three-Dimensional Plotting in Matplotlib.	1
5.7	Three-Dimensional Plotting in Matplotlib.	1
5.8	Plotting DHT11 sensor data using matplotlib	1
5.9	Plotting DHT11 sensor data using matplotlib	1