ADT294	COMPUTATIONAL FUNDAMENTALS FOR BIOINFORMATICS	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: Bioinformatics is an interdisciplinary area that combines Computer Science, Molecular Biology, and Mathematics and allied areas of Science. This course covers computational fundamentals of Bioinformatics and Computational Biology such as DNA, genes and proteins, transcription, translation, sequence alignment, representation and basic Python programming required for handling bioinformatics data. The learners will be able to solve basic bioinformatics problems using python programming.

Prerequisite: Basic understanding of programming languages.

Mapping of course outcomes with program outcomes

CO 1	Describe the basic concepts of Bioinformatics with an emphasis on biological macromolecules-DNA, RNA and Protein and synthesis of biomolecules (**Cognitive knowledge level : Understand**)
CO 2	Identify biological data formats and databases, retrieve bio-sequences, and align bio-sequences to identify similarity, dynamic programming (**Cognitive knowledge level : Apply)**
CO 3	Illustrate nucleotide attributes and transcription using programming tools (**Cognitive knowledge level : Apply)**
CO 4	Demonstrate the concepts of Parsing FASTA and Sequences Analysis (**Cognitive knowledge level : Apply)**
CO 5	Compute k-mers, translation of DNA subsequences and Open reading frame. (**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
	Test1 (%)	Test2 (%)	
Remember	10	10	10
Understand	30	30	70
Apply	10	10	20
Analyse			
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 Internal 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. 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 bioinformatics)
Introduction to bioinformatics, Nature & Scope of Bioinformatics, animal vs plants, Eukaryote vs prokaryote,Nucleus. Chromosome, gene DNA, RNA, amino acids, and Protein, The Central Dogma, Messenger RNA, tRNA, rRNA, Genetic code, Gene Structure, Transcription, translation.

Module-2 (Introduction to bio sequences and analysis)
Introduction to Biological Databases and data storage, NCBI, Genbank, Bio sequence formats-Database Similarity Searching, BLAST, Sequence alignment, Scoring Matrices, Multiple-Sequence Alignment, Dynamic programming

Module 3: (Introduction to Processing Nucleotides)
Tetranucleotide Frequency, Counting the Nucleotides, Writing and Verifying a Solution, Transcribing DNA into mRNA: Mutating Strings, Reading and Writing Files, Reverse Complement of DNA, String Manipulation, Iterating Over a Reversed String.

Module 4: (Processing Nucleotides GC Content and Hamming Distance)
Creating the Fibonacci Sequence, Writing, Testing, and Benchmarking Algorithms, retrieving FASTA Using Biopython, Iterating the Sequences Using a for Loop, Parsing FASTA and Analyzing Sequences, Computing GC Content, Finding the Hamming Distance, Counting Point Mutations

Module 5 (Translation of DNA and subsequence)

K-mers and Codons, Translating Codons, Translating mRNA into Protein, Finding Subsequences of DNA, Find a Motif in DNA, Finding Overlapping Patterns Using Regular Expressions, Sequence Similarity, Finding the Shortest Sequence in a FASTA File, Extracting K-mers from a Sequence, Counting Frequencies of K-mers, Finding Open Reading Frames

Text Books

1. Mount, D. W.. Bioinformatics: Sequence and Genome Analysis. India, CBS Publishers & Distributors, 2005.
2. Youens-Clark, Ken. *Mastering Python for Bioinformatics*. United States: O'Reilly Media, 2021.

References

1. Kelley, S.T. and Didulo, D, *Computational Biology: A Hypertextbook*. John Wiley & Sons, 2020

2. Baxevanis, Andreas D., Gary D. Bader, and David S. Wishart, eds. *Bioinformatics*. John Wiley & Sons, 2020.
3. Shaik, Noor Ahmad, et al. *Essentials of Bioinformatics, Volume I.* Springer, 2019
4. Selzer, Paul M., Richard J. Marhöfer, and Andreas Rohwer, *Applied bioinformatics. An introduction–Springer, Verlag,*, 2008.
5. S C Rastogi, N Mendiratta and P Rastogi, *Bioinformatics: Methods and Applications,* PHI Learning Private Limited, New Delhi, 2015.
6. D E Krane and M L Raymer, *Fundamental Concepts of Bioinformatics*, Pearson Education, 2006.
7. Bassi, Sebastian. *Python for Bioinformatics*. United Kingdom: CRC Press, 2017.
8. Model, Mitchell L. Bioinformatics Programming Using Python. United States: O'Reilly Media, 2010.
9. Antao, Tiago. *Bioinformatics with Python Cookbook*. United Kingdom: Packt Publishing, 2015.Antao, Tiago. Bioinformatics with Python Cookbook: Learn how to Use Modern Python Bioinformatics Libraries and Applications to Do Cutting-edge Research in Computational Biology, 2nd Edition. United Kingdom: Packt Publishing, 2018.

Course Level Assessment Questions

Course Outcome 1 (CO1)

1. Compare and contrast the DNA and RNA on the basis of structure and functions.
2. Demonstrate with the help of a flow diagram the generation of protein using the transcription and translation process.

Course Outcome 2 (CO2):

1. Identify the following qualifiers for GenBank and give their definitions: [ACCN], [ALL], [AUTH], [ECNO], [FKEY], [GENE], [JOUR], [KYWD]

2. Find the sequence alignment between the following two sequences, locally and Globally

   Sequence1: GATTCTATCTAACTA, Sequence2: GTTCTATTCTAAC

3. Retrieve sequence of Severe acute respiratory syndrome coronavirus 2 and use BLAST to find the similar sequences

Course Outcome 3 (CO3):

1. Write a Python program pseudocode to read the below given sequence as command line argument and print the counts for each of the bases A, C, G, and T.

   Sequence: ACTGCAACGGGCAATATGTCTC

2. Write a python pseudocode to transcribe the following DNA sequence to its mRNA sequence.

   Sequence: TGCAACGGGCAATATGTCTC

Course Outcome 4 (CO4)

1. Solve the problem of generating the Fibonacci sequence using Python.

2. Use a simple python program using a list to find the DNA string having the highest GC content, provided any 5 random DNA strings.

Course Outcome 5 (CO5)

1. Illustrate with the help of an example how an RNA string is getting converted to a protein string.

2. Write a python code to print the position and the number of times a subsequence is present in a given DNA string.

Model Question Paper

QP CODE:

Reg No: _____

Name: _____   **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE (HONOURS) EXAMINATION, MONTH & YEAR

Course Code: ADT294

Course Name: COMPUTATIONAL FUNDAMENTALS FOR BIOINFORMATICS

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

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Differentiate DNA, Gene, genome and chromosome.

2. What do you mean by Gene expression?

3. Specify the functions of mRNA, tRNA and rRNA?

4. Differentiate between local and global alignment.

5. Find the reverse complement of the following DNA given in 5'-3'order? AAAACCCGGT

6. List any 3 string manipulation construct used in processing nucleotides.

7. Illustrate how recursion is implemented using a Python pseudocode.

8. What is GC content? Give the GC content of the DNA string: "AGCTATAG".

9. Discuss the role of K-mers and codons in protein synthesis.

10. Define motif in DNA. Mention its importance in finding a conserved sequence.

(10x3=30)

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

11. (a) Discuss the central dogma of molecular biology. **(7)**

    (b) How is the primary transcript produced by a prokaryote different from that produced by a eukaryotic cell? **(7)**

OR

12. (a) Differentiate between Prokaryote and Eukaryote Cell **(7)**

    (b) Describe with the help of a neat diagram, the structure of DNA. **(7)**

13. (a) What is sequence alignment? Explain any five applications of sequence alignment in Bioinformatics? **(7)**

    (b) Discuss variants of BLAST with its input and output **(7)**

OR

14. (a) Explain the working principles of the Nucleotide BLAST with an example **(7)**

    (b) Differentiate primary and secondary databases in Bioinformatics. **(7)**

15. (a) How do you find the reverse complement of a DNA sequence? Write at least 2 different Python pseudocodes using different constructs to print the reverse complement of a given the 5'-3' end of a DNA sequence. **(10)**

    (b) Write a Python pseudocode to convert DNA sequence to RNA sequence by using the re.sub() regular expression construct. **(4)**

OR

16. (a) What is the need for 'argparse' module in Python? How can we use this module in different ways to do a tetra nucleotide frequency count? **(7)**

    (b) Write a Python program pseudocode to read the below given sequence as command line argument and print the counts for each of the bases A, C, G, and T. **(7)**

17. (a) Generate a random DNA sequence using python and find the transcribed **(7)**

DNA sequence of its reverse complement

(b)   Write a python code using regular expressions to find the DNA sequence **(7)**
having the highest GC content in a DNA sequence.

OR

18.   (a)   Define Hamming distance. Using hamming distance, find the percentage of **(7)**
similarity   between   the   sequence   AAACCCGGGTTT   and
AACCCGGGTTTA with one sequence in line with other.

(b)   Write a Python code using zip() function to find the hamming distance **(7)**
between 2 sequence. Give comments on each construct used in the code.

19.   (a)   Write a Python program using function and a list comprehension to **(10)**
translate RNA into protein. Illustrate working of the program with an
example RNA string.

(b)   Illustrate with python pseudocode to show how the str.find( ) function can **(4)**
be used to find a substring and its position in an input sequence.

OR

20.   (a)   Illustrate with the help of an example how an RNA string is getting **(6)**
converted to a protein string..

(b)   Write notes on ORF. Write a python code to find the ORF using the **(8)**
str.find( ) and str.partition( ) functions.

TEACHING PLAN

No	Contents	No of Lecture Hrs
	Module-1 (Introduction to bioinformatics)(10 hrs)	
1.1	Introduction to bioinformatics	1
1.2	Nature & Scope of Bioinformatics	1
1.3	Animal vs plants, Eukaryote vs prokaryote	1
1.4	Nucleus. Chromosome, gene	1
1.5	DNA, RNA, and Protein	1
1.6	The Central Dogma introduction	1
1.7	Messenger RNA, tRNA, rRNA,	1
1.8	Genetic code	1
1.9	Gene Structure and Control	1
1.10	Transcription, Translation	1
	Module-2 (Introduction to bio sequences and analysis) (10 hrs)	
2.1	Introduction to Biological Databases and data storage	1
2.2	NCBI, Genbank	1
2.3	NCBI, Genbank Sequence retrieval	1
2.4	Bio sequence formats	1
2.5	Database Similarity Searching, BLAST	1
2.6	BLAST Exercises	1
2.7	Sequence alignment	1
2.8	Scoring Matrices	1
2.9	Multiple-Sequence Alignment	1
2.10	Introduction to Dynamic programming in MSA	1
	Module-3 (Introduction to Processing Nucleotides) (8 hrs)	
3.1	Counting the Nucleotides, Writing and Verifying a Solution	1

3.2	Transcribing DNA into mRNA	1
3.3	Iterating the Input Files	1
3.4	Mutating Strings	1
3.5	Writing and Reading Output Sequences	1
3.6	Reverse Complement of DNA	1
3.7	String Manipulation	1
3.8	Iterating Over a Reversed String	1

Module-4 (Processing Nucleotides GC Content and Hamming Distance) (8 hrs)		
4.1	Creating the Fibonacci Sequence	1
4.2	Writing, Testing, and Benchmarking Algorithms	1
4.3	Retrieving FASTA Using Biopython	1
4.4	Parsing FASTA and Analysing Sequences	1
4.5	Computing GC Content	1
4.6	Finding the Hamming Distance	1
4.7	Iterating the Characters of Two Strings	1
4.8	Counting Point Mutations	1
Module-5 (Translation of DNA and subsequence) (9 hrs)		
5.1	K-mers and Codons	1
5.2	Translating mRNA into Protein	1
5.3	Finding Subsequence of DNA	1
5.4	Find a Motif in DNA	1
5.5	Finding Overlapping Patterns Using Regular Expressions	1
5.6	Sequence Similarity	1
5.7	Finding the Shortest Sequence in a FASTA File , Extracting K-mers from a Sequence	1
5.8	Counting Frequencies of K-mers	1
5.9	Finding Open Reading Frames	1

ADT296	ADVANCED TOPICS IN COMPUTER GRAPHICS	CATEGORY	L	T	P	CREDITS
		VAC	3	1	0	4

Preamble: This course helps the learners to make awareness about strong theoretical concept in computer graphics. It covers the three-dimensional environment representation in a computer, transformation of 2D/3D objects, basic mathematical techniques and algorithms used to build useful applications. This course enables the learners to develop the ability to create image processing frameworks for different domains and develops algorithms for emerging display technologies.

Prerequisite: A sound knowledge of Mathematics and concepts of any 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 and transformations on 2D & 3D objects. **(Cognitive Knowledge level: Apply)**
CO4	Demonstrate the working of various clipping algorithms and projection algorithms. **(Cognitive Knowledge level: Apply)**
CO5	Summarize visible surface detection methods**(Cognitive Knowledge level: Understand)**
CO6	Explain the concept of realism in a scene and its performance preservation**(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	⊘											⊘
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. 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(Line and Circle drawing 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, Color CRT displays, Flat panel display and its categories. Line drawing algorithms - DDA, Bresenham's algorithm. Circle drawing algorithms - Midpoint Circle generation algorithm, Bresenham's algorithm.

Module - 2(Filled Area Primitives and Two dimensional 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.

Module - 3 (Clipping and 3D transformations)
Window to viewport transformation. Cohen Sutherland and Midpoint subdivision line clipping algorithms, Sutherland Hodgeman and Weiler Atherton Polygon clipping algorithms. Three dimensional viewing pipeline. Basic 3D transformations.

Module - 4 (Projections and Visible Surface detection)
Projections- Parallel and Perspective projections. Visible surface detection algorithms- Back face detection, Depth buffer algorithm, Scan line algorithm, A buffer algorithm

Module - 5 (Realism and performance)
Realism - Illumination Shading, Shadows, Texture mapping, Bump mapping, Environment mapping, Transparency, Accumulation Buffer, Back face Culling, Visibility Culling.

Text Books

1. Donald Hearn and M. Pauline Baker, Computer Graphics, PHI, 2e, 1996

2. Aditi Majumder and M.Gopi , Introduction to VISUAL COMPUTING Core Concepts in Computer Vision, Graphics, and Image Processing, 2018

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) Donald Hearn, M. Pauline Baker and Warren Carithers, Computer Graphics with OpenGL, PHI, 4e, 2013

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 accepted from the user(2,3) and (5,8) using Bresenham's line drawing algorithm and implement it using any appropriate programming language. (Assignment)

2. Illustrate how the 4-connected area filling approach differs from 8- connected area filling in boundary filling algorithm and implement it using any appropriate programming language.(Assignment)

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. Implement the above transformation using any appropriate programming language with user inputs. (Assignment)

Course Outcome 4 (CO4):

1. 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).

2. Implement Cohen Sutherland clipping algorithm using any appropriate programming language with user inputs. (Assignment)

Course Outcome 5 (CO5):

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

Course Outcome 6 (CO6):

1. You are rendering a black and white checkered tiled floor using a single texture mapped polygon. The view is simulating a person standing on the floor and looking at a point far away from him on the floor. (1)Artifacts at the distant end of the floor can be seen. How would you remove these artifacts? (2) How can you explain why this method works using the sampling theorem?

2. You are seeing an object which is either texture mapped, bump mapped or displacement mapped but you don't know which one. However, you have the liberty to move the light and the viewpoint of an object and see it from different angles and for different positions of the light. How will you figure out which technique was used?

Model Question Paper

QP CODE:

Reg No: _____

Name: _____   **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE (HONOURS) EXAMINATION, MONTH & YEAR

Course Code:ADT296

Course Name: Advanced Topics in Computer Graphics

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

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Consider a raster system with a resolution of 1024*1024. Compute 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?

2. How 8-way symmetry of circle can be used for writing circle drawing algorithms? Write the symmetric points if (x, y) is a point on the circle with centre at origin.

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. How does Cohen Sutherland algorithm determine whether a line is visible, invisible or a candidate for clipping based on the region codes assigned to the end points of the line?

7. Define the terms (i) Centre of projection (ii) Principal vanishing point

8. Differentiate between the object space and image space method for the hidden surface removal of an image.

9. Describe the steps used to convert the normal map to bump mapping.

10. One artifact of Gouraud shading is that it can miss specular highlights in the interior of the triangles. How can this be explained as an aliasing artifact?

(10x3=30)

Part B

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

11. (a) Derive the initial decision parameter of Bresenham's line drawing algorithm and rasterize a line with endpoints (2,2) and (10,10). **(8)**

(b) Draw the architecture of raster scan display systems and explain its working principle **(6)**

OR

12. (a) Explain the working principle of a Refresh CRT monitor with suitable diagrams. **(7)**

(b) Write Midpoint circle drawing algorithm and plot a circle with radius=20 and center (50,30) using the algorithm. **(7)**

13. (a) Differentiate between boundary fill and flood fill algorithms. **(5)**

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

OR

14. (a) A diamond shaped polygon is located at P(-1,0), Q(0,-2), R(1,0) and S(0,2). Find the transformation matrix which would rotate the triangle by 90 degree counter clockwise about the point Q. Using the transformation matrix, find the coordinates of the rotated polygon. **(7)**

(b) Illustrate the working principle of scan line polygon filling algorithm **(7)**

15. (a) Illustrate Weiler – Atherton polygon clipping algorithm. **(6)**

(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). **(8)**

OR

16. (a) Describe the steps required for a general 3D rotation if the rotation axis is not parallel to any one of the principal axis. The rotation axis is defined by the points P1(x1,y1,z1) and P2(x2,y2,z2). Give its composite matrix representation **(6)**

(b) Describe Sutherland Hodgeman polygon clipping algorithm and list out its limitations **(8)**

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

(b) Define parallel projection. Describe orthographic and oblique parallel projection. **(7)**

OR

18. (a) Illustrate the scan line method used in visible surface detection. **(7)**

(b) Explain the steps involved in performing perspective projections **(7)**

19. (a) Specify any three shading algorithms used in interactive graphics. **(6)**

(b) Explain the procedure of texture to object space mapping. **(8)**

OR

20. (a) Explain the mapping scheme in which the effects of small bumps on the surface of an object can be simulate without changing the number of primitives **(8)**

(b) Describe about object to screen space mapping. **(6)**

TEACHING PLAN

No	Contents	No of Lecture Hrs
Module – 1 (Line and Circle drawing algorithms) (10 hrs)		
1.1	Basics of Computer Graphics and applications	1
1.2	Refresh Cathode Ray Tubes	1
1.3	Random and Raster Scan Displays and systems,	1
1.4	Color CRT displays	1
1.5	Flat panel display and its categories.	1
1.6	DDA Line drawing Algorithm	1
1.7	Bresenham's line drawing algorithm	1
1.8	Midpoint Circle generation algorithm	1
1.9	Bresenham's Circle generation algorithm	1
1.10	Illustration of line and circle drawing algorithms	1
Module - 2 (Filled Area Primitives and Two dimensional transformations) (9 hrs)		
2.1	Scan line polygon filling	1
2.2	Boundary filling and flood filling	1
2.3	Basic 2D transformations-Translation	1
2.4	Basic 2D transformations- Rotation	1
2.5	Basic 2D transformations- Scaling	1
2.6	Reflection and Shearing	1
2.7	Illustration of Basic 2D Transformations	1
2.8	Composite transformations	1
2.9	Matrix representations and homogeneous coordinates	1
Module - 3 (Clipping and 3D transformations) (8 hrs)		
3.1	Window to viewport transformation	1
3.2	Cohen Sutherland Line clipping algorithm	1
3.3	Midpoint subdivision Line clipping algorithm	1
3.4	Sutherland Hodgeman Polygon clipping algorithm	1
3.5	Weiler Atherton Polygon clipping algorithm	1
3.6	Three dimensional viewing pipeline	1

3.7	Basic 3D transformation-Translation and scaling	1
3.8	Basic 3D transformation-Rotation	1

Module - 4 (Projections and Visible Surface detection) (7 hrs)		
4.1	Projections-Parallel projections	1
4.2	Projections- Perspective projections	1
4.3	Illustration of projection methods	1
4.4	Visible surface detection algorithms- Back face detection	1
4.5	Depth buffer algorithm	1
4.6	Scan line visible surface detection algorithm	1
4.7	*A buffer* algorithm	1

Module - 5 (Realism and performance)(10 hrs)		
5.1	Illumination	1
5.2	Shading and Shadows	1
5.3	Texture mapping-Texture to object space mapping	1
5.4	Texture mapping-Object to screen space mapping and Mip Mapping	1
5.5	Bump mapping	1
5.6	Bump mapping-Illustration	1
5.7	Environment mapping and Transparency	1
5.8	Accumulation Buffer and Back face Culling	1
5.9	Visibility Culling	1
5.10	Visibility Culling	1

AIT395	COMPUTATIONAL BIOLOGY	CATEGORY	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: This course helps the learners to understand concepts in Genomics, Proteomics Computational Biology, Next Generation Sequencing, NGS Data Analysis and Systems biology. It enables the learners to understand various Next Generation Sequencing Techniques, analysis and interpretation of the NGS Data. Also, course introduces computational and mathematical analysis and modeling of complex biological systems and Systems Biology

Prerequisite: Basic background in Bioinformatics

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

CO 1	Describe the basic concepts of genomics, microarray, protein structure determination and prediction(**Cognitive knowledge level: Understand**)
CO 2	Explain the fundamental aspects drug discovery and molecular modelling (**Cognitive knowledge level: Apply**)
CO 3	Demonstrate Networks in Biology, types of networks and its representation (**Cognitive knowledge level : Apply**)
CO 4	Explain Next Generation sequencing Technologies and DNA Protein interaction analysis(**Cognitive knowledge level: Understand**)
CO 5	Illustrate Next Generation sequence analysis, Mapping approaches and algorithms (**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	⊘	⊘			⊘							⊘

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
	Test1 (%)	Test2 (%)	
Remember	30	30	30
Understand	50	50	50
Apply	20	20	20
Analyse			
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 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. 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 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. Each question can have maximum 2 sub-divisions and carries 14 marks.

SYLLABUS`

Module -01 (Genomics and Proteomics)

Genes, Genes in genomes, Genomes of prokaryotes and Eukaryotes, Protein-coding genes, RNA, Single-nucleotide polymorphisms, Microarray, Analysis of microarray data, Proteins and peptides, Experimental Protein structure identification, computational methods for protein structure prediction, Homology modelling, Protein folding and fold recognition.

Module-02 (Computer Aided Drug Discovery)

Drug discovery pipeline, Drug target identification & validation, Active site identification, pharmacophore, Lead/Ligand identification, lead compound optimization, Binding energy calculation, Energy Minimization. Molecular modelling in drug discovery, concept of Molecular Dynamics, concept of Absorption, Distribution, Metabolism and Excretion (ADME), Quantitative Structure-Activity Relationships.

Module-03 (Network Biology)

Transcriptional Regulatory Networks, Genes and DNA Regulatory Regions, Genetic Interaction Map, Protein Interaction Networks, Experimental methodologies to obtain Protein Interaction Data, Computational methods to Predict Protein-Protein Interactions, Visualization of Protein Interaction Networks,Metabolic Networks, Interacting Partners, Mathematical Representation

Module-04 (Next Generation Sequencing and analysis)

A Typical NGS Experimental Workflow, Next-Generation Sequencing (NGS) Technologies, Illumina Reversible Dye-Terminator Sequencing, Ion Torrent Semiconductor Sequencing,

Pacific Biosciences Single Molecule Real-Time (SMRT) Sequencing, RNA-sequencing (RNA Seq), Protein-DNA Interaction Analysis (ChIP-Seq)

Module-05 (NGS Data Analysis)

Base Calling, FASTQ File Format, and Base Quality Score, NGS Data Quality Control and Preprocessing, Reads Mapping, Mapping Approaches and Algorithms, Selection of Mapping Algorithms and Reference Genome Sequences, SAM/BAM as the Standard Mapping File Format, Mapping File Examination and Operation, Tertiary Analysis

Books
1. Lesk, Arthur M. Introduction to Bioinformatics. United Kingdom, Oxford University Press, 2019.
2. Biological Networks. Singapore, World Scientific Publishing Company, 2007.
3. Wang, Xinkun. Next-Generation Sequencing Data Analysis. United States, CRC Press, 2016.

References
1. Tiwary, Basant K.. Bioinformatics and Computational Biology: A Primer for Biologists. Singapore, Springer Singapore, 2021.
2. Benfey, Philip N.. Quickstart Molecular Biology: An Introductory Course for Mathematicians, Physicists, and Computational Scientists. United States, Cold Spring Harbor Laboratory Press, 2014.
3. Baxevanis, Andreas D., Gary D. Bader, and David S. Wishart, eds. *Bioinformatics*. John Wiley & Sons, 2020.
4. Shaik, Noor Ahmad, et al. *Essentials of Bioinformatics, Volume I.* Springer, 2019
5. Selzer, Paul M., Richard J. Marhöfer, and Andreas Rohwer, *Applied bioinformatics. An introduction–Springer, Verlag,*, 2008.
6. S C Rastogi, N Mendiratta and P Rastogi, *Bioinformatics: Methods and Applications* , PHI Learning Private Limited, New Delhi, 2015.
7. D E Krane and M L Raymer, *Fundamental Concepts of Bioinformatics*, Pearson Education, 2006.

Course Level Assessment Questions

Course Outcome 1 (CO1):
1. Compare and contrast the genomes of Prokaryotes and Eukaryotes
2. Summarize the method of DNA microarray and its analysis.
3. Using the online tool SWISS-MODEL, develop model of Homo sapiens (Human) Leptin protein and interpret your result

Course Outcome 2 (CO2):
1. Explain the process of computer aided drug discovery and various step involved in it
2. Explain the process of molecular modelling in drug discovery

Course Outcome 3 (CO3):

1. Differentiate between Transcriptional and protein interaction networks
2. From the STRING database identify the interactions of Homo sapiens TP53 protein and interpret your result

Course Outcome 4 (CO4):

1. Summarize Next Generation Sequencing methods.
2. Explain The Protein- DNA interaction analysis with the help of ChIP-Seq
3. What can RNA-seq reveal?

Course Outcome 5 (CO5):

1. Illustrate the process involved in Data Quality control and preprocessing in Next Generation Sequencing
2. Explain the mapping algorithms and reference genome sequences

TEACHING PLAN

No	Contents	No of Lecture (45Hrs)
Module -01 (Genomics and Phylogenetics) (9hrs)		
1.1	Genes, Genes in genomes.	1
1.2	Genomes of prokaryotes and Eukaryotes	1
1.3	Protein-coding genes, RNA, Single-nucleotide polymorphisms	1
1.4	Microarrays	1
1.5	Analysis of microarray data	1
1.6	Proteins and peptides	1
1.7	Experimental Protein structure identification	1
1.8	Computational methods for protein structure prediction	1
1.9	Homology modelling, Protein folding and fold recognition	1
Module-02 (Computer Aided Drug Discovery)(9hrs)		
2.1	Drug discovery pipeline	1
2.2	Drug target identification & validation	1
2.3	Active site identification, pharmacophore	1
2.4	Lead/Ligand identification	1
2.5	lead compound optimization, Binding energy calculation, Energy Minimization	1
2.6	Molecular modelling in drug discovery	1

2.7	Concept of Molecular Dynamics	1
2.8	Concept of Absorption, Distribution, Metabolism and Excretion (ADME)	1
2.9	Quantitative Structure-Activity Relationship	1
Module-03 (Network Biology)(9hrs)		
3.1	Transcriptional Regulatory Networks	1
3.2	Genes and DNA Regulatory Regions,	1
3.3	Genetic Interaction Map,	1
3.4	Protein Interaction Networks	1
3.5	Experimental methodologies to obtain Protein Interaction Data	1
3.6	Computational methods to Predict Protein-Protein Interactions	1
3.7	Visualization of Protein Interaction Networks	1
3.8	Metabolic Networks- Interacting Partners	
3.9	Metabolic Networks- Mathematical Representation	

Module-04 (Next Generation Sequencing and analysis) (8hrs)		
4.1	A Typical NGS Experimental Workflow	1
4.2	Next-Generation Sequencing (NGS) Technologies	1
4.3	Next-Generation Sequencing (NGS) Technologies	1
4.4	Illumina Reversible Dye-Terminator Sequencing	1
4.5	Ion Torrent Semiconductor Sequencing	1
4.6	Pacific Biosciences Single Molecule Real-Time (SMRT) Sequencing	1
4.7	RNA-sequencing (RNA Seq)	1
4.8	Protein-DNA Interaction Analysis (ChIP-Seq)	1
Module-05 (NGS Data Analysis)(10hrs)		
5.1	NGS data,FASTQ File Format	1
5.2	Base Calling, Base Quality Score	1
5.3	NGS Data Quality Control	1
5.4	NGS data Preprocessing	1
5.5	Reads Mapping, Mapping Approaches and Algorithms,	1

5.6	Selection of Mapping Algorithms and Reference Genome Sequences	1
5.7	SAM/BAM as the Standard Mapping	1
5.8	Mapping File Examination and Operation	1
5.9	Tertiary Analysis	1
5.10	Demonstration of NGS Data Analysis	1
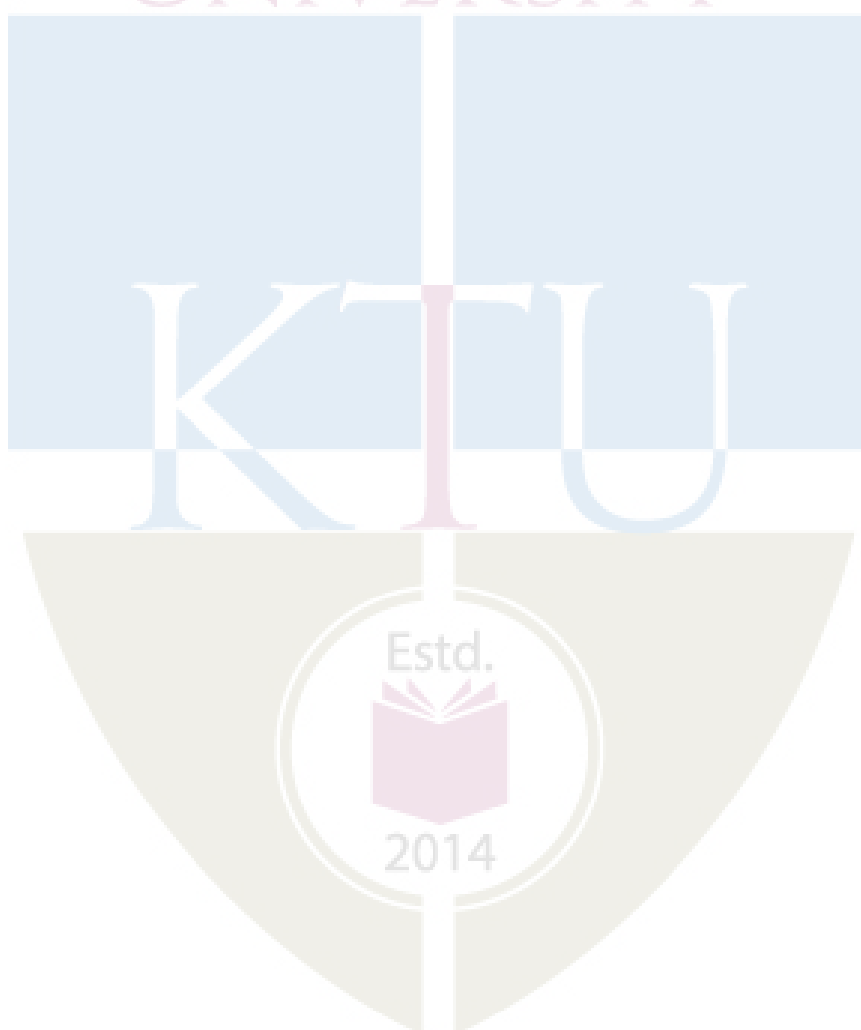
Model Question Paper

QP CODE:

Reg No: _____

Name: _____   **PAGES: 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

EIGHTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: AIT395

Course Name: COMPUTATIONAL BIOLOGY

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

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Distinguish between Genes, Genes in genomes.

2. What are the structural features of Eukaryotic cells?

3. What are SNPs and why are they important?

4. How do you identify the active site of a protein?

5. What is protein energy minimization?

6. List any three types of biochemical networks with one line description

7. What are reversible Dye-Terminators in NGS sequencing?

8. What is the difference between the DNA sent for Whole Exome sequencing vs ChIP sequencing?

9. List any three features of FastQ file format.

10. What is SAM format? How is BAM different from SAM?   **(10x3=30)**

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

11. (a) With the help of a neat diagram, explain a prokaryotic gene structure. Is a promoter at the upstream or downstream of a transcription unit?   **(7)**

    (b) What is homology modeling? Discuss the steps involved in the same   **(7)**

                                OR

12. (a) Explain the design of a microarray experiment, detailing the various phases. **(7)**

(b) What experimental method is used to determine the tertiary protein structure? **(7)**
What are the computational methods?

13. (a) Illustrate the computational drug discovery pipeline with a suitable flowchart **(7)**

(b) What is Molecular modeling in drug discovery? Explain the process of **(7)**
molecular modelling.

OR

14. (a) Explain the scoring functions in molecular docking. **(7)**

(b) Explain lead compound optimization, Binding energy calculation, Energy **(7)**
Minimization in the process of Computer aided drug discovery

15. (a) What is transcriptional control and why is it important? Explain how **(7)**
transcriptional regulatory networks plays an important role in gene
expression and control?

(b) Explain how the computational methods helps in identifying the Protein– **(7)**
Protein Interactions

OR

16. (a) How the Protein–Protein Interactions are identified by using experimental **(7)**
methods.

(b) What is metabolic network? What are type of data are needed for metabolic **(7)**
network reconstruction?

17. (a) Explain any two next-generation sequencing techniques with their steps. **(7)**

(b) How do you interpret a FastQC report? **(7)**

OR

18. (a) What are the steps in RNA sequencing?Why is RNA-seq better than **(7)**
microarrays?

(b) illustrate the steps involved in mapping protein-DNA interactions using ChIP- **(7)**
sequencing

19. (a) How do you interpret per base sequence quality?What is the purpose of **(7)**
mapping reads to a reference genome?

(b) Explain any three mapping algorithms for the NGS. **(7)**

OR

20. (a) Illustrate steps involved in the NGS data Preprocessing and Quality Control **(7)**

(b) Discuss the significance of NGS in clinical diagnosis. **(7)**

AIT396	MACHINE LEARNING IN COMPUTATIONAL BIOLOGY	CATEGORY	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: This course is intended to provide the learners a outlook towards application of Machine learning algorithms in the field of computational biology. This course helps the learners to apply the Machine learning methods - clustering algorithms, dimensionality reduction, decision drees, Artificial Neural Network, Support Vector Machine to the computational biology problems. Also the course discuss Challenges of Machine Learning in Computational Biology and Future directions of Machine Learning in Computational Biology.

Prerequisite: Basic background in Bioinformatics and Machine Leaning

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

CO 1	Describe the basic concepts of Machine Leaning, Classification, regression and clustering problems, parameters and measures (**Cognitive knowledge level: Understand**)
CO 2	Demonstrate the clustering algorithm on computational biology problems (**Cognitive knowledge level: Apply**)
CO 3	Explain Dimensionality reduction techniques and Decision Trees in computational biology (**Cognitive knowledge level : Apply**)
CO 4	Illustrate Feature Extraction and Pattern recognition and Classification in the domain of Computational Biology analysis (**Cognitive knowledge level: Apply**)
CO 5	Explain the role and challenges of Machine Learning in Computational (**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	⊘	⊘			⊘							⊘

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
	Test1 (%)	Test2 (%)	
Remember	30	30	30
Understand	50	50	50
Apply	20	20	20
Analyse			
Evaluate			
Create			

Mark Distribution

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

Continuous Internal Evaluation Pattern:

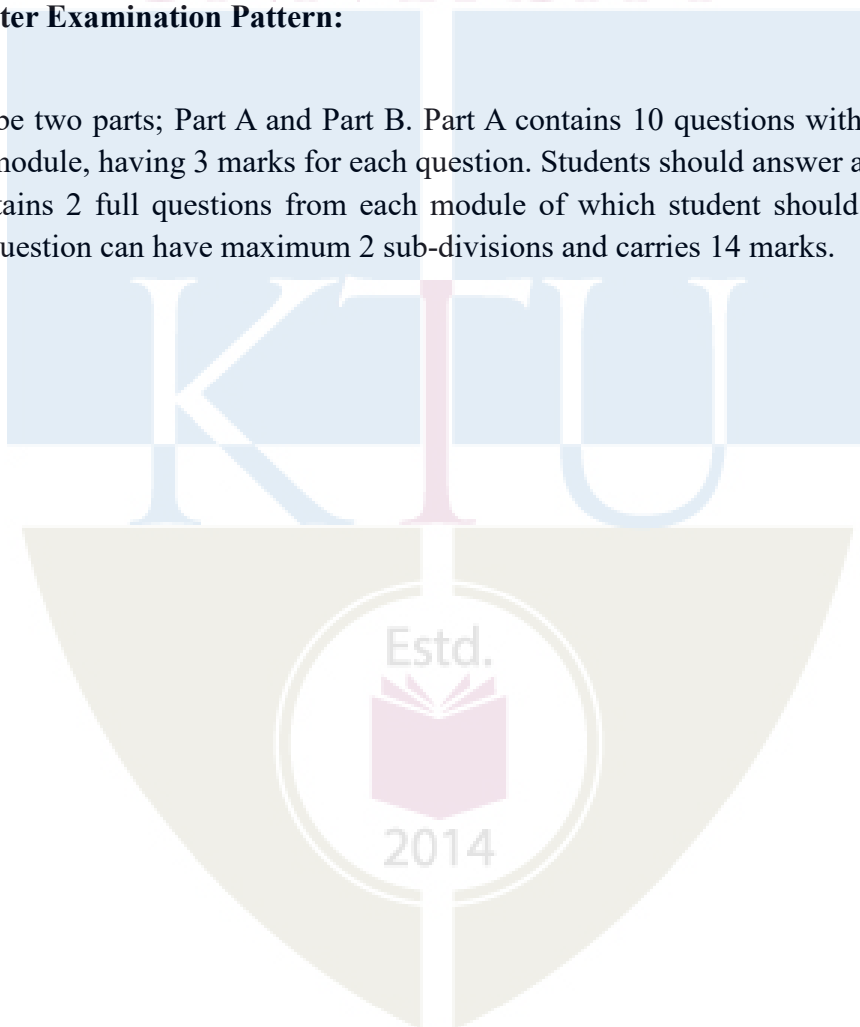|	
Attendance	**10 marks**
Continuous Assessment Tests (Average of Series Tests 1& 2)	**25 marks**
Continuous Assessment Assignment	**15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. 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 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. Each question can have maximum 2 sub-divisions and carries 14 marks.

Machine Learning in Computational Biology

Module 1 (Overview of Machine Learning)

Overview of Machine Learning, fitting predictive models to data, Supervised and unsupervised learning, Classification, regression and clustering problems, Loss or cost functions. Parameters and hyperparameters, Training. validation and testing, Inductive bias and the bias variance trade-off, Use of clustering models.

Module 2 (Clustering problems Computational Biology)

Hierarchical Clustering, Partition Clustering, Overview Model-Based Clustering, k-Means clustering, k-Means clustering algorithm, Advantages, Disadvantages, illustrative example of k-Means clustering, Clustering for creating phylogenetic trees, Using Clustering Approach to Identify Patients' Subtypes, Application of clustering algorithms on gene expression data.

Module 3 (Supervised techniques for Computational Biology)

Proteomics Dataset, Data Pre-processing Algorithms, Dimension and Feature Subset Selection, Dimensionality reduction - Principal Component Analysis (PCA), Partial Least Square (PLS), Linear Discriminant Analysis (LDA), Protein Classification, Decision Trees in Bioinformatics, Proteomic Mass Spectra Classification Using Decision Tree Technique.

Module 4 (Machine-Learning Algorithms for Computational Biology)

Machine-Learning Algorithms for Feature Selection from Gene Expression Data, Feature Extraction and Pattern recognition from sequence data, measures of a Feature. Artificial Neural Network (ANN) in Bioinformatics, Genetic Algorithms (GA) in Bioinformatics, Designing ANN for Bioinformatics, ANN in Protein Bioinformatics, Support Vector Machine with Feature Elimination.

Module 5 (Scope of Machine Learning in Computational Biology)

Role of Machine Learning in Computational Biology, Creation and analysis of sequence data, Challenges of Machine Learning in Computational Biology, Data Errors, Mean Square Error Generative versus Discriminative, Approximation Versus Explanation, Single Versus Multiple Methods, Future directions of Machine Learning in Computational Biology.

Text Books

1. Statistical Modelling and Machine Learning Principles for Bioinformatics Techniques, Tools, and Applications. Germany, Springer Singapore, 2020.
2. Yang, ZhengRong. Machine Learning Approaches to Bioinformatics. Singapore, World Scientific Publishing Company, 2010.

References

1. Izadkhah, Habib. Deep Learning in Bioinformatics: Techniques and Applications in Practice. Netherlands, Elsevier Science, 2022.
2. Agapito, Giuseppe, et al. Artificial Intelligence in Bioinformatics: From Omics Analysis to Deep Learning and Network Mining. Netherlands, Elsevier Science, 2022.
3. Data Analytics in Bioinformatics: A Machine Learning Perspective. United States, Wiley, 2021.
4. Michailidis, George, et al. Introduction to Machine Learning and Bioinformatics. United Kingdom, CRC Press, 2008.
5. Zhang, Yanqing, and Rajapakse, Jagath C, Machine Learning in Bioinformatics, Germany, Wiley, 2009.
6. Baldi, Professor Pierre, et al. Bioinformatics, Second Edition: The Machine Learning Approach. India, Bradford, 2001.

Course Level Assessment Questions

Course Outcome 1 (CO1):
1. Compare and contrast Supervised and unsupervised learning
2. Differentiate Classification with regression with an example
3. Explain the parameters and hyperparameters of a model?
4. Summarize validation and testing in machine learning?

Course Outcome 2 (CO2):
1. Write K-means algorithm and separate {5, 11, 19, 27, 23, 25, 6, 18, 2, 8, 10, 12, 31, 29, 4} into 3 clusters
2. Illustrate application of clustering algorithms on gene expression data
3. Differentiate K-means clustering and hierarchical clustering

Course Outcome 3 (CO3):
1. Illustrate dimensionality reduction methods - Principal Component Analysis (PCA), Partial Least Square (PLS), Linear Discriminant Analysis (LDA)
2. Explain Decision trees in Bioinformatics with a toy example.

Course Outcome 4 (CO4):
1. Explain the process involved in feature extraction and pattern recognition from sequence data
2. Design and implement an ANN model for the prediction of relative solvent accessibility

Course Outcome 5 (CO5):
1. Summarize role of Machine Learning in Computational Biology
2. Explain Challenges of Machine Learning approaches in Computational Biology

Model Question Paper	
QP CODE:	
Reg No: _____	
Name: _____	**PAGES: 3**
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY	
SIXTH SEMESTER B.TECH (Honors) DEGREE EXAMINATION, MONTH & YEAR	
Course Code: AIT 396	
Course Name: MACHINE LEARNING IN COMPUTATIONAL BIOLOGY	
Max. Marks: 100	**Duration: 3 Hours**
PART A	
Answer All Questions. Each Question Carries 3 Marks	

1.	What does the regression line equation tell you?	**(3)**
2.	How do you create a predictive data model using machine learning?	**(3)**
3.	Write the major differences between K-means clustering and hierarchical clustering	**(3)**
4.	List any three resources of Proteomics Datasets	**(3)**
5.	What is the importance of using PCA before applying Machine learning method?	**(3)**
6.	Draw example of an ANN architecture including 4 independent variables, one hidden layer with 3 hidden neurons and 2 dependent variables	**(3)**
7.	What is the role of the Activation functions in Neural Networks?	**(3)**
8.	What is Hinge Loss in SVM?	**(3)**
9.	What is mean square error? how will you evaluate it?	**(3)**
10.	What are discriminative machine learning models?	**(10x3=30)**

Part B		
(Answer any one question from each module. Each question carries 14 Marks)		
11.	(a)	With example, differentiate Supervised and unsupervised learning	**(7)**

	(b)	What is loss function and cost function in machine Learning. write the difference and example of loss function and cost function	(7)
		OR	
12.	(a)	Define Train, Validation, and Test Datasets. how do you divide the data into Train, Validation, and Test Datasets.	(7)
	(b)	Explain Classification, regression and clustering methods with examples of each	(7)
13.	(a)	Use K Means clustering to cluster the following data into two groups. Assume cluster centroid are m1=2 and m2=4. The distance function used is Euclidean distance. { 2, 4, 10, 12, 3, 20, 30, 11, 25 }	(7)
	(b)	Illustrate with a toy example the application of clustering algorithms on gene expression data	(7)
		OR	
14.	(a)	Explain the advantages, disadvantages of k-Means clustering	(7)
	(b)	What is the advantage of using hierarchical clustering over K means clustering? When to use the hierarchical clustering?	(7)
15.	(a)	Explain Dimension and Feature Subset Selection	(7)
	(b)	20 physicochemical properties of 100 set of proteins were given with the help of PCA, explain how will you reduce 20x100 in to Five properties (5x100) for the next level analysis	(7)
		OR	
16.	(a)	Explain how Linear Discriminant Analysis can be used for the dimensionality reduction with the help of a scenario in computational biology	(7)
	(b)	How do decision tree classifiers work? what types of problems can they solve in Computational Biology	(7)
17.	(a)	Explain the process of Feature Extraction and Pattern recognition from sequence data	(7)
	(b)	Illustrate the design of Artificial Neural Network for solving Computational Biology question	(7)
		OR	
18.	(a)	Explain crossover and mutation in genetic algorithm with an example	(7)
	(b)	Explain how to construct a support vector machine (SVM) to classify ovarian	(7)

|			
		cancer from 30 individuals from the 15 features obtained from each patient.	
19.	(a)	What role does machine learning and have to play in Computational Biology?	**(7)**
	(b)	Explain different kinds of Data Errors in Machine Learning that would happen in case of applying it in to the Computational Biology domain?	**(7)**
OR			
20.	(a)	What are the advantages and disadvantages of application of machine learning in Computational Biology?	**(7)**
	(b)	"The transformation of huge volume of data into knowledge is the biggest challenge faced in computational biology" How can machine learning techniques help in this?	**(7)**

TEACHING PLAN

No	Contents	No of Lecture (45 Hrs)
Module 1 (Overview of Machine Learning) (9 hrs)		
1.1	Overview of Machine Learning	1
1.2	Fitting predictive models to data	1
1.3	Supervised and unsupervised learning	1
1.4	Classification, regression and clustering problems	1
1.5	Loss or cost functions	1
1.6	Proteins and peptides	1
1.7	Parameters and hyperparameters	1
1.8	Training. validation and testing	1
1.9	Inductive bias and the bias variance trade-off, Use of clustering models	1
Module 2 (Clustering problems Computational Biology) (9 hrs)		
2.1	Hierarchical Clustering	1
2.2	Partition Clustering, Overview Model-Based Clustering	1
2.3	k-Means clustering, k-Means clustering algorithm	1
2.4	k-Means clustering advantages, disadvantages	1
2.5	illustrative example of k-Means clustering	1

2.6	Clustering for creating phylogenetic trees	1
2.7	Using Clustering Approach to Identify Patients' Subtypes	1
2.8	Application of clustering algorithms on gene expression data	1
2.9	Application of clustering algorithms on gene expression data	1
Module 3 (Supervised techniques for Computational Biology) (9 hrs)		
3.1	Proteomics Datasets	1
3.2	Data Pre-processing Algorithms	1
3.3	Dimension and Feature Subset Selection	1
3.4	Dimensionality reduction	1
3.5	Principal Component Analysis (PCA)	1
3.6	Partial Least Square (PLS), Linear Discriminant Analysis (LDA)	1
3.7	Protein Classification case study	1
3.8	Decision Trees in Bioinformatics	1
3.9	Proteomic Mass Spectra Classification Using Decision Tree Technique	1

Module 4 (Machine-Learning Algorithms for Computational Biology) (8 hrs)		
4.1	Machine-Learning Algorithms for Feature Selection from Gene Expression Data	1
4.2	Feature Extraction and Pattern recognition from sequence data	1
4.3	Measures of a Feature	1
4.4	Artificial Neural Network (ANN) in Bioinformatics	1
4.5	Genetic Algorithms (GA) in Bioinformatics	1
4.6	Designing ANN for Bioinformatics	1
4.7	Designing ANN for Bioinformatics	1
4.8	ANN in Protein Bioinformatics	1
4.9	Support Vector Machine with Feature Elimination.	1
Module 5 (Scope of Machine Learning in Computational Biology) (10 hrs)		
5.1	Role of Machine Learning in Computational Biology	1
5.2	Creation and analysis of sequence data	1

5.3	Challenges of Machine Learning in Computational Biology	1
5.4	Data Errors in Machine Learning, Mean Square Error	1
5.5	Generative versus Discriminative	1
5.6	Approximation Versus Explanation	1
5.7	Single Versus Multiple Methods	1
5.8	Future directions of Machine Learning in Computational Biology	1
5.9	Future directions of Machine Learning in Computational Biology	1
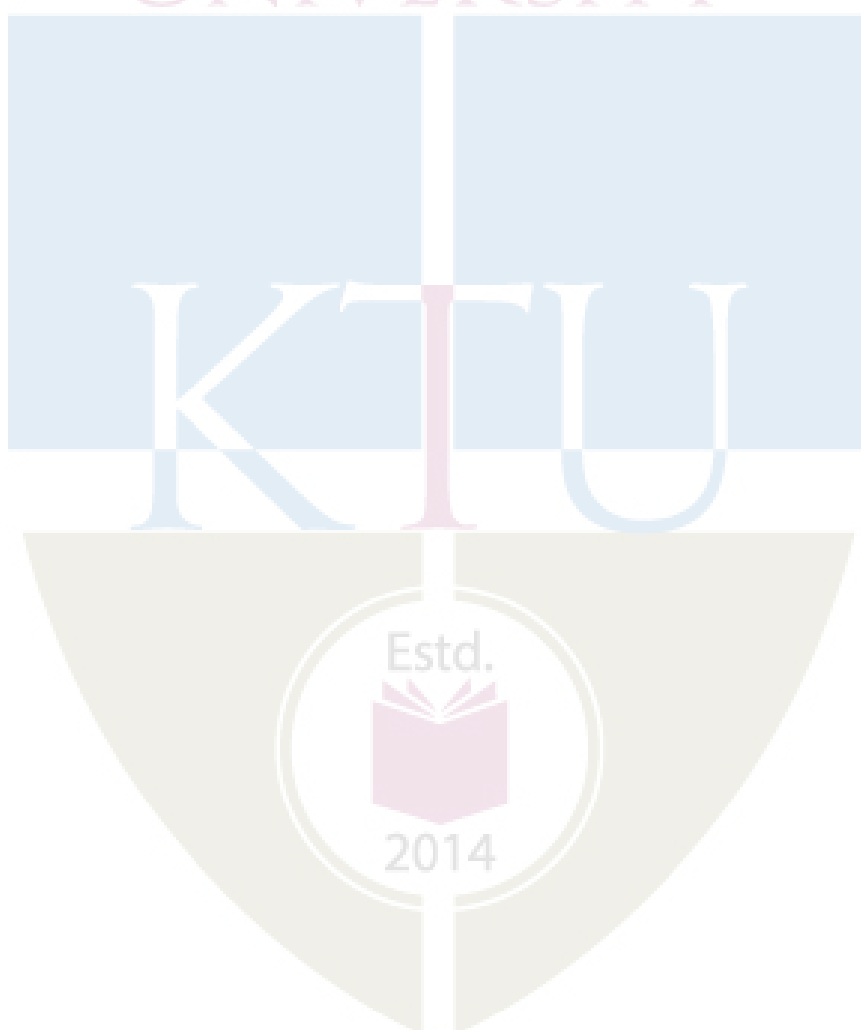
AIT397	ADVANCED CONCEPTS IN COMPUTER VISION	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: This course enables the learners to understand the advanced concepts in computer vision. The course covers the basics of image processing, imaging geometry, image segmentation, feature extraction, object recognition and classification and common applications of computer vision. This course helps the students to design solutions for complex real-life problems.

Prerequisite: A sound knowledge of Mathematics and concepts of any programming language.

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

CO1	Illustrate the concepts of image formation and image model. **(Cognitive Knowledge Level: Understand)**
CO2	Demonstrate various feature extraction and edge detection techniques. **(Cognitive Knowledge Level: Apply**
CO3	Apply edge-based and region-based image segmentation techniques. **(Cognitive Knowledge Level: Apply)**
CO4	Understand and implement image recognition and classification methods. **(Cognitive Knowledge Level: Apply)**
CO5	Explain the various applications of computer vision. **(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**	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 Internal 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. 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 a maximum of 2 subdivisions and carries 14 marks.

Syllabus

Module – 1 (Image Formation and Processing)

Image formation and Image model- Components of a vision system- Cameras- camera model and camera calibration-Radiometry- Light in space- Light in surface - Sources, shadows and shading.

Fundamentals of Image processing: Basic steps of Image processing system sampling and quantization of an Image – Basic relationship between pixels.

Module - 2(Feature Extraction)

Points and Patches – Feature detectors, feature descriptors, feature matching, feature tracking. **Edges –** edge detection, edge linking. **Lines** - Successive approximation, Hough transforms, Vanishing points.

Module - 3 (Image Segmentation)

Classification of segmentation techniques, Edge detection, Edge linking, Thresholding, Region

growing, Region splitting and merging, Watershed based segmentation. Shadow detection and removal. Image processing using OpenCV - blending, smoothing, and reshaping.

Module - 4 (Image Recognition and Classification)

Shape based object classification, Motion based object classification, Viola Jones Object Detection Framework, Object classification using CNNs, use of RCNN for object classification.

Module - 5 (Applications)

Speech and Handwriting Recognition, Automatic Face Recognition, Video Segmentation and Keyframe Extraction, Real-Time Hand Pose Recognition.

Text Books

1. David A. Forsyth & Jean Ponce, Computer vision – A Modern Approach, Prentice Hall, 2002.

2. Richard Szeliski, Computer Vision: Algorithms and Applications, Springer.

3. Maheshkumar H Kolekar, "Intelligent Video Surveillance Systems: An Algorithmic Approach", CRC Press.

4. Francesco Camastra, Alessandro Vinciarelli, "Machine Learning for Audio, Image and Video Analysis: Theory and Applications", Springer 2015.

Reference Books

1. Reinhard Klette, "Concise Computer Vision: An Introduction into Theory and Algorithms", Springer London, 2014.

2. Olivier Faugeras, "Three-Dimensional Computer Vision", The MIT Press, 1993.

Course Level Assessment Questions

Course Outcome1 (CO1):

1.Explain the components of a visual system.

2.Elaborate on the image formation model.

Course Outcome 2(CO2):

1. Explain edge linking through Hough Transform.

2. Discuss how feature extraction is done in image processing.

Course Outcome 3(CO3):

1. Compare the following methods for image segmentation: a) multiple thresholding, b) global thresholding c) local thresholding.

2. Justify the role of region growing, region splitting and region merging operations in any of the computer vision applications.

Course Outcome 4(CO4): .

1. Explain convolution stage and pooling stage of a typical CNN layer.

2. Illustrate Viola Jones object detection framework.

Course Outcome 5(CO5):

1. Elaborate on how computer vision helps in automatic face recognition applications.

2. Discuss how computer vision helps in tackling complex real world problems.

Model Question Paper

QP CODE:

Reg No: _____

Name: _____   **PAGES : 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIFTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: AIT397

Course Name: Advanced Concepts in Computer Vision

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

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Explain the working of a pinhole camera, Derive the expression for pinhole perspective projection.

2. Illustrate "foreshortening" with a neat diagram.

3. Explain edge linking through Hough Transform.

4. Illustrate any two techniques for vanishing point detection in an image.

5. Compare following methods for image segmentation
   a, multiple thresholding, b, global thresholding c, local thresholding.

6. Draw the flowchart of foreground-pixel extraction by edge-based shadow removal

7. Why is a convolutional neural network preferred over a dense neural network for an image classification task?

8. Assess the relevance of selective search algorithm in RCNN for object classification

9. Draw the diagram which shows the general scheme of a recognition system.

10. Illustrate steps in feature extraction from handwritten images.

(10x3=30)

Part B

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

11. (a) State different limitations of pinhole cameras and how to overcome these limitations. **(9)**

    (b) What are shadows? Differentiate umbra from penumbra. How is a self shadow different from a cast shadow? **(5)**

OR

12. (a) Explain the local shading model.How are area sources different from line sources? **(7)**

    (b) Define Camera Calibration. Explain intrinsic and extrinsic parameters of a camera. **(7)**

13. (a) Assess the role of adaptive non-maximal suppression (ANMS) in feature detection. **(4)**

    (b) Illustrate following techniques: **(10)**
    i) Bias and gain normalization (MOPS).
    ii) Gradient location-orientation histogram (GLOH)

OR

14. (a) Illustrate any 2 techniques in Successive approximation. **(4)**

    (b) Compare Scale invariant feature transform (SIFT) and PCA-SIFT. **(5)**

15. (a) Illustrate Gradient operator and Laplacian operator with one example for each. **(10)**

    (b) Illustrate Watershed Algorithms. **(4)**

OR

16. (a) With the help of a diagram illustrate region splitting and merging. **(7)**

(b) Compare blending, smoothing, and reshaping functions using OpenCV. **(7)**

17. (a) Differentiate between convolution stage and pooling stage of a typical CNN layer. **(8)**

(b) Assess the role of dispersedness in shape based object classification. **(6)**

OR

18. (a) Illustrate Viola Jones object detection framework. **(8)**

(b) Explain the steps in motion based object classification. **(6)**

19. (a) Illustrate shot boundary detection through pixel-based approaches and block-based approaches . **(7)**

(b) Explain different approaches in keyframe extraction problems. **(7)**

OR

20. (a) Illustrate shot boundary detection through histogram-based approaches and clustering-based approaches. **(6)**

(b) Illustrate HMM training in speech and handwriting recognition. **(8)**

TEACHING PLAN

No	Contents	No. of Lecture Hours (42 hrs)
Module – 1 (Image Formation and Processing) (8 hours)		
1.1	Image formation and Image model-Introduction	1 hour
1.2	Components of a vision system- Cameras-Camera model	1 hour
1.3	Camera calibration	1 hour
1.4	Radiometry- Light in space-Light in surface	1 hour
1.5	Sources-Shadows and shading	1 hour
1.6	Fundamentals of Image processing: Basic steps of Image processing system	1 hour
1.7	Sampling and quantization of an Image	1 hour
1.8	Basic relationship between pixels.	1 hour
Module-2(Feature Extraction) (8 hours)		
2.1	**Points and Patches** – Feature detectors	1 hour
2.2	Feature descriptors	1 hour
2.3	Feature matching	1 hour
2.4	Feature tracking.	1 hour
2.5	**Edges** – edge detection, edge linking.	1 hour
2.6	**Lines** - Successive approximation	1 hour
2.7	Hough transforms	1 hour
2.8	Vanishing points	1 hour

	Module-3(Image Segmentation)(9 hours)	
3.1	Classification of segmentation techniques, Edge detection	1 hour
3.2	Edge linking	1 hour
3.3	Thresholding, Region growing	2 hours
3.4	Region splitting and merging	1 hour
3.5	Watershed based segmentation.	1 hour
3.6	Shadow detection and removal	1 hour
3.7	Image processing using OpenCV - blending	1 hour
3.8	Smoothing, and reshaping	1 hour
	Module-4(Image Recognition and Classification) (9 hours)	
4.1	Shape based object classification	1 hour
4.2	Motion based object classification	2 hours
4.3	Viola Jones Object Detection Framework	2 hours
4.4	Object classification using CNNs	2 hours
4.6	Use of RCNN for object classification.	2 hours
	Module-5(Applications)(8 hours)	
5.1	Speech and Handwriting Recognition	1 hour
5.2	Handwriting Recognition	1 hour
5.3	Automatic Face Recognition	2 hours
5.4	Video Segmentation	2 hours
5.5	Keyframe Extraction	1 hour
5.6	Real-Time Hand Pose Recognition.	1 hour

AIT398	IMAGE AND VIDEO PROCESSING	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: This course enables the learners to understand how digital images are stored and processed. The learners are exposed to different spatial and frequency domain methods for image enhancement, image restoration techniques, morphological operations that could be performed on digital images and also various image and video compression techniques. The course also gives an introduction to the basics of video processing and video segmentation.

Prerequisite: Advanced Computer Graphics, Advanced Concepts in Computer Vision

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

CO1	Summarize the steps of digital image processing and pixel relationships. **(Cognitive Knowledge Level: Understand)**
CO2	Apply spatial and frequency domain methods for image enhancement. **(Cognitive Knowledge Level: Apply**
CO3	Apply restoration techniques and morphological operations on digital images. **(Cognitive Knowledge Level: Apply)**
CO4	Compare different methods for digital image and video compression. **(Cognitive Knowledge Level: Apply)**
CO5	Understand the basics of video processing and video segmentation. **(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	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 hrs

Continuous Internal Evaluation Pattern:

Attendance	**10 marks**
Continuous Assessment Tests(Average of Internal 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. 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 a maximum of 2 subdivisions and carries 14 marks.

SYLLABUS

Module – 1

Fundamentals of Image processing: Basic steps of Image processing system, sampling and quantization of an Image, basic relationship between pixels and connectivity.

Image Enhancement: Spatial Domain methods - Gray level transformations, Histogram Processing, Fundamentals of Spatial Filtering, Smoothing Spatial filters, Sharpening Spatial filters.

Module -2

Image Transforms: Unitary transforms, 2D Discrete Fourier Transform, Discrete Cosine Transform (DCT), Discrete Wavelet transforms.

Frequency Domain methods: Basics of filtering in frequency domain, image smoothing, image sharpening, homomorphic filtering.

Module - 3

Image Restoration: Image degradation/Restoration model, Noise models, Restoration in presence of noise only - spatial filtering, Periodic Noise reduction by frequency domain filtering.

Morphological Operations: Erosion, Dilation, Opening, Closing, Hit-or-miss transformation, Boundary extraction.

Module - 4

Image compression fundamentals – Coding Redundancy, spatial and temporal redundancy.

Compression models : Lossy and Lossless, Huffman coding, Arithmetic coding, LZW coding, run length coding, Bit Plane coding, JPEG standards.

Module - 5

Video processing: Basics of Video Processing: Analog video, Digital Video.

Video segmentation: Introduction to video segmentation, Change detection.

Video Compression: Introduction to video compression, video compression based on motion compensation, Search for motion vectors, H.261 standard, Transform coding, predictive coding-MPEG.

Text Books

1. Gonzalez and Woods , "Digital Image Processing", 3rd edition , Pearson, 2009.

2. Li, Ze-Nian, Mark S. Drew, and Jiangchuan Liu. "Fundamentals of multimedia", Pearson Prentice Hall, 2004.

3. Bovik, Alan C. "Handbook of image and video processing", Academic press, 2010.

Reference Books

1. David A. Forsyth & Jean Ponce, Computer vision – A Modern Approach, Prentice Hall, 2002.

2. Richard Szeliski, Computer Vision: Algorithms and Applications, Springer.

3. Maheshkumar H Kolekar, "Intelligent Video Surveillance Systems: An Algorithmic Approach", CRC Press.

4. Francesco Camastra, Alessandro Vinciarelli, "Machine Learning for Audio, Image and Video Analysis: Theory and Applications", Springer 2015.

5. M. Tekalp ,"Digital video Processing", Prentice Hall International

6. Relf, Christopher G.,"Image acquisition and processing with LabVIEW", CRC press

7 Chris Solomon, Toby Breckon ,"Fundamentals of Digital Image Processing A Practical Approach with Examples in Matlab", John Wiley & Sons,

8. Yao wang, Joem Ostarmann and Ya – quin Zhang, "Video processing and communication ",1st edition , PHI

Course Level Assessment Questions

Course Outcome1 (CO1):

1. Illustrate how the image is digitized by sampling and quantization.
2. 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.

3	1	2	1q
2	2	0	2
1	2	1	1
p 1	0	1	2

Course Outcome 2(CO2):

1. Determine whether the given matrix is unitary or not:

$$A = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

2. Explain any five properties of 2D Fourier Transform.

Course Outcome 3(CO3):
1. Discuss how restoration is done in digital images.
2. Explain with examples the different morphological operations applied to images.

Course Outcome 4(CO4): .

1. With suitable examples, clearly bring out the need for compression in images and videos.
2. Discuss any one method for finding motion vectors.

Course Outcome 5(CO5):

1. Explain any one technique used for segmenting a video.
2. Compare and contrast analog video and digital video in multimedia.

Model Question Paper

QP CODE:

Reg No: _____

Name: _____   **PAGES : 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: AIT 398

Course Name: Image and Video Processing

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

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Explain bit plane slicing and contrast stretching.

2. Discuss about pixel relationships.

3. Find the 4 order forward and inverse DFT for the following image segment:

1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1

4. Define DCT. Write the properties of DCT.

5. Discuss hit or miss transformation with appropriate examples.

6. Explain about the morphological operation dilation.

7. Explain the significance of image compression.

8. Distinguish between lossy and lossless compression.

9. Discuss the significance of change detection.

10. Explain how transform coding is used in compression algorithms.

(10x3=30)

Part B

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

11. (a) Perform histogram specification of the following 3 bit gray scale image whose gray level distribution is given as follows. **(9)**

Input image

Gray level	0	1	2	3	4	5	6	7
No. of Pixels	8	10	10	2	12	16	4	2

Target image

Gray Level	0	1	2	3	4	5	6	7
No. of Pixels	0	0	0	0	20	20	16	8

(b) Design Laplacian filter for image enhancement in spatial domain. **(5)**

OR

12. (a) What is histogram equalization? Explain the procedure for histogram equalization. **(7)**

(b) Explain the gray level transformation functions: a) image negatives and b) log transformation c) power law transformation. **(7)**

13. (a) Compute the 2D DFT of the 4 X 4 grayscale image given below. **(4)**

(b) Explain about smoothing and sharpening frequency domain filters. **(10)**

OR

14. (a) Explain Butterworth filters for image smoothening and image sharpening. **(4)**

(b) Explain the steps followed in frequency domain filtering? **(5)**

15. (a) Apply opening and closing operation on the image sample A given below with structuring element B **(10)**

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

(b) Explain Morphological operations a) opening b) closing with suitable examples. **(4)**

OR

16. (a) Discuss about different noise models. **(7)**

(b) Explain how periodic noise reduction can be done using frequency domain filtering. **(7)**

17. (a) Comment on JPEG compression standard. **(8)**

(b) Discuss on run-length encoding with the help of an example. **(6)**

OR

18. (a) Explain LZW coding with the help of a suitable example. **(8)**

(b) Illustrate the concept of arithmetic coding. **(6)**

19. (a) Compare and contrast MPEG video coding and H.261 standard. **(7)**

(b) Explain video segmentation with an example. **(7)**

OR

20. (a) Illustrate how motion compensation is used in video compression. **(7)**

(b) With the help of a neat block diagram explain predictive coding methods. **(7)**

Teaching Plan

No	Contents	No. of Lecture Hours (44 hrs)
Module – 1 (7 hours)		
1.1	Fundamentals of Image processing: Basic steps of Image processing system, Sampling and quantization of an Image.	1 hour
1.2	Basic relationship between pixels and connectivity.	1 hour
1.3	Image Enhancement: Gray level transformations	1 hour
1.4	Histogram, Histogram Equalization	1 hour
1.5	Histogram specification	1 hour
1.6	Fundamentals of Spatial Filtering	1 hour
1.7	Smoothing Spatial filters	1 hour
1.8	Sharpening Spatial filters	1 hour
Module-2 (8 hours)		
2.1	Image Transforms: Unitary transforms.	1 hour
2.2	2D Discrete Fourier Transform	1 hour

2.3	Discrete Cosine Transform (DCT)	1 hour
2.4	Discrete Wavelet transforms	1 hour
2.5	Basics of filtering in frequency domain	1 hour
2.6	Image smoothing	1 hour
2.7	Image sharpening	1 hour
2.8	Homomorphic filtering.	1 hour
Module-3 (9 hours)		
3.1	**Image Restoration:** Image degradation/Restoration model	1 hour
3.2	Noise models	1 hour
3.3	Restoration basics	1 hour
3.4	Restoration in presence of noise only - spatial filtering	1 hour
3.5	Periodic Noise reduction by frequency domain filtering.	1 hour
3.6	Morphological Operations: basics	1 hour
3.7	Erosion, Dilation, Opening, Closing	1 hour
3.8	Hit-or-miss transformation	1 hour
3.9	Boundary extraction.	1 hour
Module-4 (10 hours)		
4.1	Image compression fundamentals - Coding Redundancy	1 hour
4.2	Spatial and temporal redundancy.	1 hour
4.3	Compression models : Lossy and Lossless	1 hour
4.4	Huffman coding	1 hour

4.6	Arithmetic coding	1 hour
4.7	LZW coding	1 hour
4.8	Run length coding	1 hour
4.9	Bit Plane coding,	1 hour
4.10	JPEG standards	1 hour
Module-5 (10 hours)		
5.1	Basics of Video Processing: Analog video, Digital Video.	1 hour
5.2	Video segmentation: Introduction to video segmentation	1 hour
5.3	Change detection.	1 hour
5.4	Introduction to video compression	1 hour
5.5	Video compression based on motion compensation	1 hour
5.6	Search for motion vectors	1 hour
5.7	Transform coding	1 hour
5.8	Predictive coding	1 hour
5.9	MPEG standards	1 hour
5.10	H.261 standard	1 hour

CST 393	CRYPTOGRAPHIC ALGORITHMS	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2019

Preamble:

The course on Cryptographic Algorithms aims at exploring various algorithms deployed in offering confidentiality, integrity, authentication and non-repudiation services. This course covers classical encryption techniques, symmetric and public key crypto-system, key exchange and management, and authentication functions. The concepts covered in this course enable the learners in effective use of cryptographic algorithms for real life applications.

Prerequisite: A sound background in Number Theory**.**

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

CO1	Identify the security services provided for different types of security attacks. (**Cognitive Knowledge Level : Understand**)
CO2	Summarize the classical encryption techniques for information hiding. (**Cognitive Knowledge Level: Apply**)
CO3	Illustrate symmetric / asymmetric key cryptographic algorithms for secure communication.(**Cognitive Knowledge Level: Apply**)
CO4	Interpret key management techniques for secure communication.(**Cognitive Knowledge Level: Understand)**
CO5	Summarize message authentication functions in a secure communication scenario.(**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
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks
	Test1 (Percentage)	Test2 (Percent	

		age)	
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   : **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 carries 14 marks.

Syllabus

Module-1 (Introduction to the Concepts of Security)

Need for security, Security approaches, Principles of security, Types of attacks, OSI Security Architecture, Classical encryption techniques - Substitution techniques, Transposition techniques. Stream cipher, Block cipher, Public key cryptosystems vs. Symmetric key cryptosystems, Encrypting communication channels.

Module-2 (Symmetric Key Cryptosystems)

Overview of symmetric key cryptography, Block cipher principles, Data Encryption Standard (DES), Differential and Linear cryptanalysis, Double DES, Triple DES, International Data Encryption Algorithm (IDEA), Advanced Encryption Algorithm (AES),Block cipher modes of operation, Stream cipher, RC4.

Module-3 (Public Key Cryptosystems)

Principles of public key cryptosystems, RSA algorithm, RSA illustration, Attacks, ElGamal cryptographic system, Knapsack algorithm, Diffie-Hellman key exchange algorithm, Elliptical curve cryptosystems.

Module-4 (Key Management)

Symmetric key distribution using symmetric encryption, Symmetric key distribution using asymmetric encryption, Distribution of public keys, Generating keys, transferring keys, Verifying keys, Updating keys, Storing keys, Backup keys, Compromised keys, Public key infrastructure.

Module – 5 (Authentication)

Authentication requirements, Authentication functions, Message authentication codes (MAC), Hash functions, Security of Hash functions and MAC, Message Digest 5 (MD5), Secure Hash Algorithm (SHA)-512, Hash-based Message Authentication Code (HMAC), Cipher-based Message Authentication Code (CMAC), X.509 Authentication services.

Text Books

1. William Stallings, Cryptography and Network Security Principles and Practice, Pearson Edu, 6e.
2. Bruice Schneier, Applied Cryptography Protocols, Algorithms and source code in C, Wiley,2e.

References

1. Behrouz A. Forouzan, Cryptography and Network Security, McGraw Hill, 2e.

2. Johannes A. Buchmann, Introduction to Cryptography, Springer, 2e.

3. DouglasR. Stinson, Cryptography Theory and Practice, 3e,Chapman & Hall/CRC, 2006.

4. Bernard Menezes, Network Security and Cryptography, Cengage Learning, 2011.

Sample Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Consider an automated teller machine (ATM) in which users provide a personal identification number (PIN) and a card for account access. Give examples of confidentiality, integrity, and availability requirements associated with the system and, in each case, indicate the degree of importance of the requirement.

2. Discuss the different security services provided for preventing security attacks.

Course Outcome 2 (CO2):

1. The encryption key in a transposition cipher is (3,2,6,1,5,4). Find the decryption key

2.Discuss the process of encryption in Vernam cipher

Course Outcome 3 (CO3):

1. Devise a meet-in-the-middle attack for a triple DES.

2. Write an algorithm for the InvSubBytes transformation and implement using python (**Assignment**)

3. Consider the following elliptic curve signature scheme. We have a global elliptic curve, prime p, and "generator" G. Alice picks a private signing key X_A and forms the public verifying $Y_A = X_A G$. To sign a message M:

   - Alice picks a value k

   - Alice sends Bob M, k and the signature $S = M - kX_A G$.

   - Bob verifies that $M=S+kY_A$.

   Show that the verification process produces an equality if the signature is valid.

4. Write an algorithm to add two points on an elliptic curve over GF(p) and implement using Python. (**Assignment**)

5. Write an algorithm for encryption using knapsack cryptosystem and implement using Java. (**Assignment**)

Course Outcome4 (CO4):

1. List four general categories of schemes for the distribution of public keys.

2. What are the essential ingredients of a public-key directory?

Course Outcome 5 (CO5):

1. State the value of the length field in SHA-512 if the length of the message is 1919 bits and 1920 bits.

2. Write an algorithm in pseudo code for HMAC and implement using Python (**Assignment**)

Model Question Paper

QP CODE:

Reg No:_____

Name :_____   **PAGES : 3**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIFTH SEMESTER B.TECH DEGREE EXAMINATION(HONORS), MONTH & YEAR

Course Code: CST 393

Course Name: Cryptographic Algorithms

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

PART A

Answer all Questions. Each question carries 3 Marks

1. State the two approaches in attacking a cipher.

2. Define Substitution Cipher. Encrypt using one time pad M = HONORS and K = CIPHER.

3. Specify the purpose of S-Boxes in Data Encryption Standard (DES).

4. Differentiate between diffusion and confusion.

5. Perform encryption using RSA Algorithm for the following p=7; q=11; e=13; M=5.

6. Is Diffie-Hellman key exchange protocol vulnerable? Justify.

7. List the techniques for distribution of public keys.

8. Define a certificate authority and its relation to public key cryptography.

9. Distinguish between integrity and message authentication.

10. What types of attacks are addressed by message authentication?

(10x3=30)

Part B

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

11. (a) With a neat sketch, Explain OSI Security architecture model.

(8)

(b) How does link encryption differ from end-to-end encryption? Explain.

(6)

OR

12. (a) Encrypt the text "cryptography" using the Hill Cipher with the key $\begin{pmatrix} 9 & 4 \\ 5 & 7 \end{pmatrix}$. Show the calculations.

(8)

(b) Illustrate the steps involved in encrypting a plain text using playfair cipher with an example.

(6)

13. (a) With a neat sketch, explain a single round in DES.

10

(b) Explain encryption and decryption using 2 keys and 3 keys of triple DES.

(4)

OR

14. (a) Explain the block cipher modes i) Cipher feedback mode ii) Output feedback mode.

(8)

(b) Describe the four types of transformations in AES.

(6)

15. (a) Write an algorithm for generating public and private key using Elliptical curve cryptography.

(10)

(b) The equation $y^2=x^3 +x+1$, the calculation is done modulo 13. Add two points R= P+Q, where P= (4,2) and Q= (10,6). **(4)**

OR

16. User A and B use the Diffie-Hellman key exchange technique with a common prime q=71 and primitive root alpha=7.

(a) If user A has private key X_A =3, What is A's public key Y_A? **(7)**

(b) If user B has private key X_B =6, What is A's public key Y_B? **(7)**

17. (a) Define a session key and show how a KDC can create can create a session key between Alice and Bob. **(7)**

(b) What are the requirements for the use of a public-key certificate scheme? **(7)**

OR

18. (a) What are the core components of a PKI? Briefly describe each component. **(8)**

(b) Describe the following (i) Updating keys (ii) Compromised Keys. **(6)**

19. (a) Describe how SHA-512 logic produce message digest **(10)**

(b) Distinguish between HMAC and CMAC **(4)**

OR

20. (a) Specify the format for X.509 certificate. Explain the steps required to obtain user's certificate. **(7)**

(b) With suitable block diagrams, explain the types of functions that may be used to produce an authenticator. **(8)**

Teaching Plan

No	Contents	No of Lecture Hrs
	Module - 1 (Introduction to the Concepts of Security) (9 hrs)	
1.1	Need for security, Security approaches	1 hour
1.2	Principles of security, Types of attacks	1 hour
1.3	OSI Security Architecture	1 hour
1.4	Classical encryption techniques: Substitution techniques(Caesar cipher, Monoalphabetic cipher, Playfair cipher)	1 hour
1.5	Classical encryption techniques: Substitution techniques (Hill cipher, Polyalphabetic cipher, One-time pad)	1 hour
1.6	Classical encryption techniques: Transposition techniques	1 hour
1.7	Stream cipher, Block cipher	1 hour
1.8	Public- key cryptosystems vs. Symmetric key cryptosystems	1 hour
1.9	Encrypting communication channels	1 hour
	Module - 2 (Symmetric key cryptosystems) (11 hrs)	
2.1	Overview of symmetric key cryptography	1 hour
2.2	Block cipher principles	1 hour
2.3	Data Encryption Standard (DES)	1 hour
2.4	DES design criteria	1 hour
2.5	Differential and Linear cryptanalysis	1 hour
2.6	Double DES, Triple DES	1 hour

2.7	IDEA	1 hour
2.8	Advanced Encryption Algorithm (AES structure)	1 hour
2.9	Advanced Encryption Algorithm (Transformations)	1 hour
2.10	Block cipher modes of operation	1 hour
2.11	Stream cipher, RC4	1 hour
Module - 3 (Public key cryptosystems) (8 hrs)		
3.1	Principles of public key cryptosystems	1 hour
3.2	RSA algorithm	1 hour
3.3	RSA illustration, Attacks	1 hour
3.4	ElGamal cryptographic system	1 hour
3.5	Knapsack algorithm	1 hour
3.6	Diffie-Hellman key exchange algorithm	1 hour
3.7	Elliptical curve cryptosystems(Elliptical curve arithmetic)	1 hour
3.8	Elliptical curve cryptosystems (Elliptical curve algorithm)	1 hour
Module - 4 (Key Management) (8 hrs) [Text book-2]		
4.1	Symmetric key distribution using symmetric encryption	1 hour
4.2	Symmetric key distribution using asymmetric encryption	1 hour
4.3	Distribution of public keys	1 hour
4.4	Generating keys, Transferring keys	1 hour

4.5	Verifying keys, Updating keys	1 hour
4.6	Storing keys, Backup keys	1 hour
4.7	Compromised keys	1 hour
4.8	Public key infrastructure	1 hour
Module - 5 (Authentication) (9 hrs)		
5.1	Authentication requirements	1 hour
5.2	Authentication functions	1 hour
5.3	Message Authentication Codes (MAC)	1 hour
5.4	Hash functions	1 hour
5.5	Security of Hash functions and MAC	1 hour
5.6	MD5	1 hour
5.7	SHA-512	1 hour
5.8	HMAC, CMAC	1 hour
5.9	X.509 Authentication services	1 hour

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
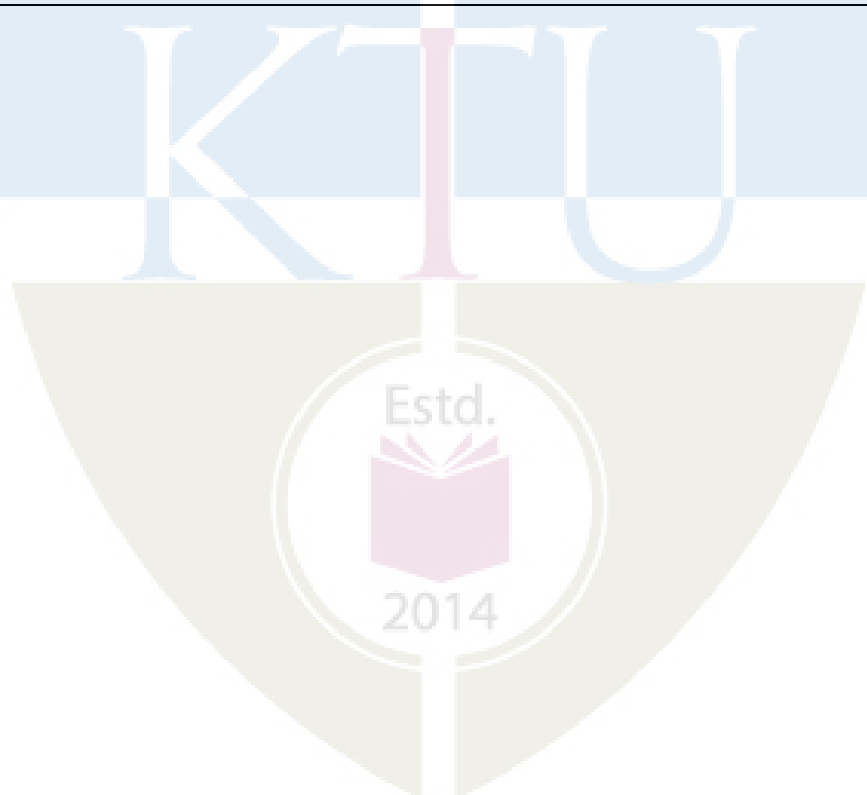
SEMESTER IV

HONOURS

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT	Year of Introduction
CST292	NUMBER THEORY	VAC	4	0	0	4	2019

Preamble: This is the foundational course for awarding B. Tech. Honours in Computer Science and Engineering with specialization in *Security in Computing*. The purpose of this course is to create awareness among learners about the important areas of number theory used in computer science. This course covers Divisibility & Modular Arithmetic, Primes & Congruences, Euler's Function, Quadratic Residues and Arithmetic Functions, Sum of Squares and Continued fractions. Concepts in Number Theory help the learner to apply them eventually in practical applications in Computer organization & Security, Coding & Cryptography, Random number generation, Hash functions and Graphics.

Prerequisite: A sound background in Higher Secondary School Mathematics

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

CO1	Illustrate modular arithmetic operations, methods and techniques **(Cognitive Knowledge Level:Understand)**
CO2	Use the methods - Induction, Contraposition or Contradiction to verify the correctness of mathematical assertions **(Cognitive Knowledge Level: Apply)**
CO3	Utilize theorems and results about prime numbers, congruences, quadratic residues and integer factorization for ensuring security in computing systems **(Cognitive Knowledge Level: Analyse)**
CO4	Illustrate uses of Chinese Remainder Theorem & Euclidean algorithm in Cryptography and Security **(Cognitive Knowledge Level: Apply)**
CO5	Explain applications of arithmetic functions in Computer Science **(Cognitive Knowledge Level:Understand)**
CO6	Implement Number Theoretic Algorithms using a programming language **(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 (Percentage)
	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 maximum 2 sub-divisions and carries 14 marks.

SYLLABUS
Module 1

Divisibility and Modular Arithmetic:

Finite Fields – Groups, Rings and Fields.

Divisibility - Divisibility and Division Algorithms, Well ordering Principle,Bezout's Identity.

Modular Arithmetic- Properties, Euclid's algorithm for the greatest common divisor, Extended Euclid's Algorithm, Least Common multiple, Solving Linear Diophantine Equations, Modular Division.

Module 2

Primes and Congruences:

Prime Numbers-Prime Numbers andprime-powerfactorization, Fermat and Mersenne primes., Primality testing and factorization.

Congruences-Linear congruences, Simultaneous linear congruences, Chinese Remainder Theorem, Fermat's little theorem, Wilson's theorem.

Module 3

Congruences with a Prime-Power Modulus&Euler's Function:

Congruences with a Prime-Power Modulus-Arithmetic modulo p, Pseudoprimes and Carmichael numbers, Solving congruences modulo prime powers.

Euler's Function-Euler's Totient function, Applications of Euler's Totient function, Traditional Cryptosystem, Limitations.

The Group of units- The group U_n,Primitive roots, Existence of primitive roots, Applications of primitive roots.

Module 4

Quadratic Residues & Arithmetic Functions :

Quadratic Residues- Quadratic Congruences, The group of Quadratic residues, Legendre symbol, Jacobi Symbol, Quadratic reciprocity.

Arithmetic Functions- Definition and examples, Perfect numbers, Mobius function and its properties, Mobius inversion formula, The Dirichlet Products.

Module 5

Sum of Squares and Continued Fractions:

Sum of Squares- Sum of two squares, The Gaussian Integers, Sum of three squares, Sum of four squares.

Continued Fractions -Finite continued fractions, Infinite continued fractions, Pell's Equation, Solution of Pell's equation by continued fractions.

Text Books

1. G.A. Jones & J.M. Jones, Elementary Number Theory, Springer UTM, 2007.

2. Joseph Silverman, A Friendly introduction to Number Theory, Pearson Ed. 2009.

Reference Books

1. William Stallings, Cryptography and Network Security Principles and Practice, Pearson Ed.

2. Tom M.Apostol, 'Introduction to Analytic Number Theory', Narosa Publishing House Pvt. Ltd, New Delhi, (1996).

3. Neal Koblitz, A course in Number Theory and Cryptography, 2nd Edition, Springer ,2004.

Sample Course Level Assessment Questions

Course Outcome 1 (CO1): Describe the properties of modular arithmetic and modulo operator.

Course Outcome 2 (CO2): Prove that the equation $y^2 = x^3 - 2$ has only the integer solution $(3, \pm5)$.

Course Outcome 3 (CO3): State the law of reciprocity for Jacobi symbols and use it to determine whether 888 is a quadratic residue or non residue of the prime 1999.

Course Outcome 4 (CO4): Using Chinese remainder theorem, solve the system of congruence x $\equiv$2(mod 3), x $\equiv$3(mod 5), x $\equiv$2(mod 7)

Course Outcome 5(CO5): State and prove Dirichlet product.

Course Outcome 6 (CO6): Use extended Euclid's algorithm to solve Diophantine equations efficiently. Given three numbers $a>0$, $b>0$, and c, the algorithm should return some x and y such that a x + b y = c.

Model Question Paper

QP CODE: PAGES: 03

RegNo : …................................   Name   :………..……………

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER BTECH (HONOURS) DEGREE EXAMINATION, MONTH &YEAR

Course Code:CST 292 Course
Name: Number Theory

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

PART A

Answer all Questions. Each question carries 3 Marks   (10x3=30)

1. State and prove well ordering principle.

2. Find gcd d of x=525 and y=231 and express d as ax + by where a and b are integers.

3. Solve the congruence equation 103 x≡ 57 (mod 211).

4. Use Fermat's Little theorem to show that 91 is not a prime.

5. If m is relatively prime to n , show that $\Phi(mn) = \Phi(m)\, \Phi(n)$.

6. Explain how public key cryptography can be used for digital signatures.

7. Define Mobius function and prove Mobius function is a multiplicative.

8. State and prove Dirichlet product.

9. Show that every prime of the form 4k+1 canbe represented uniquely as the sum of two squares.

10. Find the continued fraction representation of the rational number 55/89.

Part B

Answer any one Question from each module.

Each question carries 14 Marks

11.   (a) State the Euclidean algorithm and its extension with an example.   (7)

(b) Find all the solutions of 24x + 34 y = 6.   (7)

OR

12.   (a) Describe the properties of modular arithmetic and modulo operator.   (7)

(b) Explain Extended Euclidean algorithm. Using the algorithm find the

multiplicative inverse of 135 mod 61 (7)

13. (a) State and prove Wilson's theorem . (7)

(b) Explain Fermat's factorization method and use it to factor 809009 (7)

OR

14. (a) Using Chinese remainder theorem, solve the system of congruences,
x ≡2(mod 3), x ≡3(mod 5), x ≡2(mod 7) (7)
(b) Define Fermat primes. Show that any two distinct Fermat numbers are
Relatively prime. (7)

15. (a) Distinguish between public key and private key encryption techniques.
Also point out the merits and demerits of both. (7)

(b) Define Carmichael number and show that a Carmichael number must

be the product of at least three distinct primes. (7)

OR

16. (a)Define a pseudo prime to a base and find all non trivial bases for which

15 is a pseudo prime. (6)

(b) Find an element of

i) order 5 modulo 11   ii) order 4 modulo 13

iii) order 8 modulo 17   iv) order 6 modulo 19 (8)

17. (a) Determine the quadratic residues and non residues modulo 17. Also

determine whether 219 is a quadratic residue or non residue of the prime 383.
(8)

(b) State the law of quadratic reciprocity. Determine those odd primes p for

which 3 is a quadratic residue and those for which it is a non residue. (6)

OR

18. (a) State and prove properties of Legendre's symbol. (7)
(b) State the law of reciprocity for Jacobi symbols and using it determine

whether 888 is a quadratic residue or non residue of the prime 1999. (7)

19. (a) Prove that the equation $y^2 = x^3 - 2$ has only the integer solution $(3, \pm 5)$. (7)

(b) Define a Gaussian integer. Factorize the Gaussian integer $440 - 55i$. (7)

OR

20. (a) If m, and n can be expressed as sum of four squares, then show that mn can also be expressed the sum of four squares. (7)

(b) Find all the solutions of the Diophantine equation $x^2 - 6y^2 = 1$. (7)

Teaching Plan

Module 1: Divisibility and Euclidean Algorithm		9 hours
1.1	Finite Fields – Groups and Rings.	1 hour
1.2	Finite Fields – Fields.	1 hour
1.3	Divisibility and Division Algorithms, Well ordering Principle.	1 hour
1.4	Decimal Expansion of a positive Integer, Greatest Common Divisor, Bezout's Theorem.	1 hour
1.5	Modular Arithmetic- Properties of congruences, Modular Arithmetic Operations, Properties of Modular Arithmetic.	1 hour
1.6	Euclid's algorithm for the greatest common divisor, Extended Euclid's Algorithm.	1 hour
1.7	Solving Linear Diophantine Equations.	1 hour
1.8	Least Common multiple and Modular Division.	1 hour
1.9	Implementation of Euclid's algorithm, Extended Euclid's Algorithm and solution of Linear Diophantine Equations.	1 hour
Module 2: Primes and Congruences		9 hours
2.1	Prime Numbersand prime-powerFactorization.	1 hour
2.2	Fermat and Mersenne primes.	1 hour
2.3	Primality testing and factorization, Miller -Rabin Test for Primality.	1 hour
2.4	Pollard's Rho Method for Factorization, Fermat's Factorization.	1 hour

2.5	Linear congruences, Simultaneous linear congruences.	1 hour
2.6	Chinese Remainder Theorem.	1 hour
2.7	Implementation of Chinese Remainder Theorem.	1 hour
2.8	Fermat's little theorem.	1 hour
2.9	Wilson's theorem.	1 hour
Module 3: Congruences with a Prime-Power Modulus &Euler's Function		**9 hours**
3.1	Congruences with a Prime-Power Modulus, Arithmetic modulo p.	1 hour
3.2	Pseudo-primes and Carmichael numbers.	1 hour
3.3	Solving congruences modulo prime powers.	1 hour
3.4	Definition of Euler Totient function, Examples and properties.	1 hour
3.5	Multiplicativity of Euler's Totient function.	1 hour
3.6	Applications of Euler's function, Euler's Theorem.	1 hour
3.7	Traditional Cryptosystem, Limitations, Public Key Cryptography.	1 hour
3.8	The Group of Units, Primitive Roots.	1 hour
3.9	Existence of primitive roots for Primes, Applications of primitive roots.	1 hour
Module 4: Quadratic Residues and Arithmetic Functions		**9 hours**
4.1	Quadratic congruences, The group of Quadratic Residues.	1 hour
4.2	Legendre symbol, Jacobi Symbol.	1 hour
4.3	Quadratic reciprocity.	1 hour
4.4	Quadratic residues for prime-power moduli.	1 hour
4.5	Arithmetic Functions: Definition and examples.	1 hour

4.6	Perfect numbers, Definition and proposition.	1 hour
4.7	Mobius inversion formula., application of the Mobius inversion formula.	1 hour
4.8	Mobius function and its properties.	1 hour
4.9	The Dirichlet Product, Definition and proof.	1 hour
Module 5: Sum of Squares and Continued Fractions		**9 hours**
5.1	Sum of Squares, Sum of two squares.	1 hour
5.2	The Gaussian Integers.	1 hour
5.3	Sum of three squares.	1 hour
5.4	Sum of four squares.	1 hour
5.5	Continued Fractions, Finite continued fractions.	1 hour
5.6	Continued Fractions, Finite continued fractions.	1 hour
5.7	Infinite continued fractions.	1 hour
5.8	Pell's Equation, Definition.	1 hour
5.9	Solution of Pell's equation by continued fractions.	1 hour