



VIT
Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: A2

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025**

	i) A negative sample with <Drama, Long, High, Actor A, Director Z, Old> ii) A positive sample with <Action, Long, High, Actor C, Director X, Young>																													
3.	a) When can an algorithm be considered PAC learnable for the concept C? Give the mathematical expression for the same. b) Prove that Gradient descent algorithm helps to attain the global minimum for convex optimization problems regardless of the starting point of search.	2																												
	c)		1	4																										
	<table border="1" data-bbox="339 819 878 1241"> <thead> <tr> <th>Actual Output</th><th>Algorithm 1 Predictions</th><th>Algorithm 2 Predictions</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>2</td></tr> <tr><td>1</td><td>2</td><td>1</td></tr> <tr><td>1</td><td>2</td><td>1</td></tr> <tr><td>2</td><td>2</td><td>2</td></tr> <tr><td>2</td><td>2</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>2</td><td>2</td><td>1</td></tr> <tr><td>2</td><td>1</td><td>1</td></tr> </tbody> </table> <p>The table above shows the actual output for a sample dataset and the predictions of two ML algorithms. Evaluate the appropriate classification metric to select the best algorithm (1 or 2) for an application that requires minimizing the fake intruders.</p>	Actual Output	Algorithm 1 Predictions	Algorithm 2 Predictions	1	1	2	1	2	1	1	2	1	2	2	2	2	2	1	1	1	1	2	2	1	2	1	1	4	
Actual Output	Algorithm 1 Predictions	Algorithm 2 Predictions																												
1	1	2																												
1	2	1																												
1	2	1																												
2	2	2																												
2	2	1																												
1	1	1																												
2	2	1																												
2	1	1																												
4	a) Assume that you are assigned the task of designing a machine translation model to translate Hindi to English. How will you perform the first step of feature representation for the text, efficiently? b) A company wants to know if customer gender (male/female) is related to their product preference (Product A/Product B). They collect survey data as follows:	5	1	2																										
	<table border="1" data-bbox="372 1695 931 1907"> <thead> <tr> <th></th><th>Product A</th><th>Product B</th><th>Total</th></tr> </thead> <tbody> <tr><td>Male</td><td>30</td><td>20</td><td>50</td></tr> <tr><td>Female</td><td>20</td><td>30</td><td>50</td></tr> <tr><td>Total</td><td>50</td><td>50</td><td>100</td></tr> </tbody> </table>		Product A	Product B	Total	Male	30	20	50	Female	20	30	50	Total	50	50	100	5												
	Product A	Product B	Total																											
Male	30	20	50																											
Female	20	30	50																											
Total	50	50	100																											



Programme Name & Branch	: M. Tech Computer Science and Engineering (MCI & MCS)
Course Code and Course Name	: MCSE612L – Cyber Security
Faculty Name(s)	: Rhymend Uthariaraj V
Class Number(s)	: VL2024250503999
Date of Examination	: 28-01-2025
Exam Duration	: 90 minutes
General instruction(s):	Maximum Marks: 50

Answer All Questions

Q. No	Questions	Marks
1.	<p>a) Illustrate the role of human layer within the cybersecurity framework and explain how it addresses the threat of “an employee falling for a phishing attack”.</p> <p>b) What are the major types of attacks on data confidentiality, integrity, and availability (CIA)? Highlight a defensive mechanism to counter a usurpation attack.</p>	5 5
2.	<p>a) Describe the key phases of the cybersecurity risk management cycle and explain the logical flow of activities within each phase.</p> <p>b) List the major constraints in security governance. Additionally, considering advancements in technology (e.g., AI, IoT), predict one new constraint that may arise in security governance over the next decade.</p>	5 5
3.	<p>a) How is “SIM cloning and hijacking” protected against authentication-based threats in mobile networks? Illustrate with a simple example.</p> <p>b) Highlight the key cybersecurity challenges in Internet of Things (IoT) environments and propose a solution to mitigate DDoS attacks originating from IoT devices.</p>	5 5
	<p>a) User Behavior Analytics (UBA) leverages machine learning to identify deviations from normal behavior patterns and enhances cybersecurity measures. Provide a simple case study to illustrate this.</p> <p>b) Explain how an Endpoint Detection and Response (EDR) system can identify and mitigate potential ransomware attack. Use an example to illustrate your explanation.</p>	5 5
	Demonstrate a real-world cyberattack scenario by identifying the vulnerabilities exploited , the threats involved , the type of attack , the incident detection and response process , and the security controls to mitigate the risk.	

Vijay



VIT[®]

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: A2

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

Apply Chi-square test to determine if the gender influences the product preference. (assume the critical value to be 3.841 for a significance level of 5% and 1 degree of freedom).

5. Given below are the dataset with 3 features, their mean values and the covariance matrix. Compute the first principal component and transform the data onto it.

Sam ple	Feature 1 (x1)	Feature 2(x2)	Feature 3 (x3)
1	2	4	1
2	0	0	0
3	3	5	2
4	5	7	3

$$\mu_1 = 2.5, \mu_2 = 4, \mu_3 = 1.5$$

$$\text{Cov} = \begin{bmatrix} 2.5 & 4.33 & 1.25 \\ 4.33 & 7.33 & 3.5 \\ 1.25 & 3.5 & 0.83 \end{bmatrix}$$

10



VIT

Vellore Institute of Technology
Instituted by the University under section 3 of UGC Act 1956

REG.NO.:

24MGS0006

SLOT: A2

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025**

Programme Name & Branch	: MTech CSE
Course Code and Course Name	: MCSE602L Machine Learning
Faculty Name(s)	: Dr. Varalakshmi, Dr. Mythili T
Class Number(s)	: VL2024250502116, VL2024250502112
Date of Examination	: 27/1/2025
Exam Duration	: 90 minutes
	Maximum Marks: 50

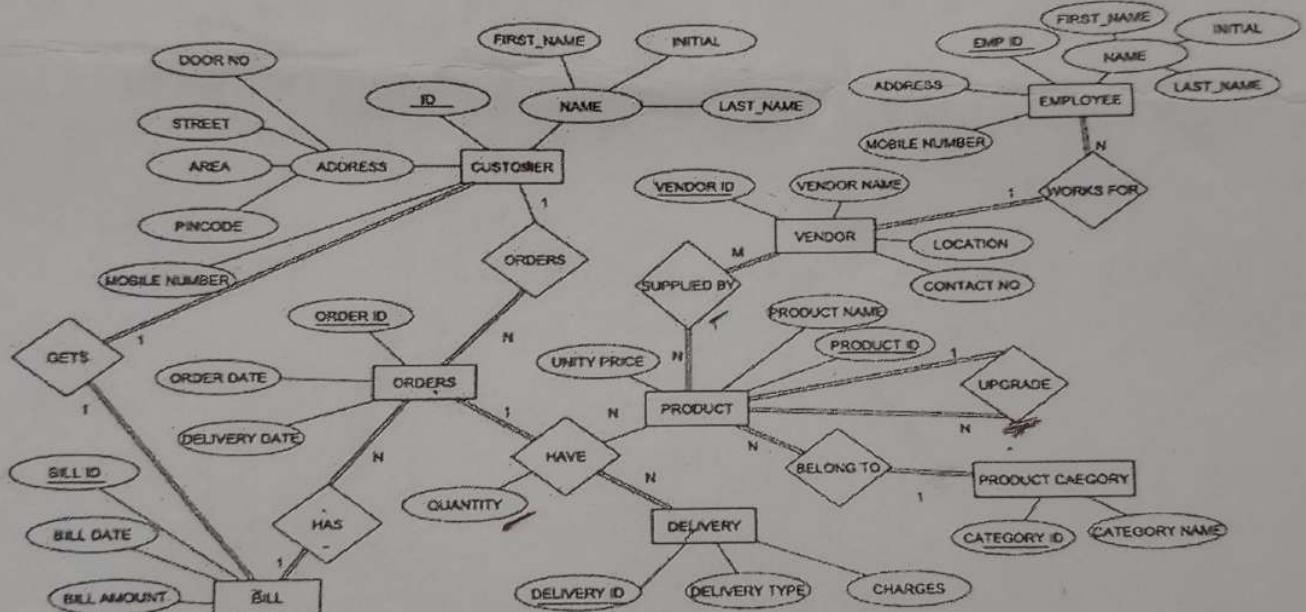
Answer All Questions

Course Outcomes

1. Recognize the characteristics of machine learning strategies
2. Analyze and apply the suitable supervised learning methods for real-world problems

Q. No	Question					M	CO	BL																																															
1.	<table border="1"> <thead> <tr> <th>Sample</th> <th>Temperature</th> <th>Humidity</th> <th>Location</th> <th>Class</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Hot</td> <td>High</td> <td>Indoor</td> <td>Positive</td> </tr> <tr> <td>2</td> <td>Cold</td> <td>Low</td> <td>Outdoor</td> <td>Negative</td> </tr> <tr> <td>3</td> <td>Warm</td> <td>Medium</td> <td>Indoor</td> <td>Positive</td> </tr> <tr> <td>4</td> <td>Hot</td> <td>High</td> <td>Outdoor</td> <td>Negative</td> </tr> <tr> <td>5</td> <td>Cold</td> <td>Medium</td> <td>Indoor</td> <td>Positive</td> </tr> <tr> <td>6</td> <td>Warm</td> <td>Low</td> <td>Outdoor</td> <td>Negative</td> </tr> <tr> <td>7</td> <td>Hot</td> <td>Medium</td> <td>Indoor</td> <td>Positive</td> </tr> <tr> <td>8</td> <td>Cold</td> <td>High</td> <td>Outdoor</td> <td>Negative</td> </tr> <tr> <td>9</td> <td>Warm</td> <td>High</td> <td>Indoor</td> <td>Positive</td> </tr> </tbody> </table> <p>Apply Naïve Bayes classifier for the above dataset to predict the class for the test input <Warm, High, Outdoor></p>	Sample	Temperature	Humidity	Location	Class	1	Hot	High	Indoor	Positive	2	Cold	Low	Outdoor	Negative	3	Warm	Medium	Indoor	Positive	4	Hot	High	Outdoor	Negative	5	Cold	Medium	Indoor	Positive	6	Warm	Low	Outdoor	Negative	7	Hot	Medium	Indoor	Positive	8	Cold	High	Outdoor	Negative	9	Warm	High	Indoor	Positive		10	2	3
Sample	Temperature	Humidity	Location	Class																																																			
1	Hot	High	Indoor	Positive																																																			
2	Cold	Low	Outdoor	Negative																																																			
3	Warm	Medium	Indoor	Positive																																																			
4	Hot	High	Outdoor	Negative																																																			
5	Cold	Medium	Indoor	Positive																																																			
6	Warm	Low	Outdoor	Negative																																																			
7	Hot	Medium	Indoor	Positive																																																			
8	Cold	High	Outdoor	Negative																																																			
9	Warm	High	Indoor	Positive																																																			
2.	<p>Assume a dataset contains the following features to predict if a person likes / dislikes a movie: - Genre, duration, rating, lead actor, director, age group</p> <p>After processing a few samples of the dataset, the candidate elimination algorithm results in the following specific and general hypothesis.</p> <p>S = <Action, Medium, High, Actor A, Director X, Young></p> <p>G = <Action,?????, <Drama,?????, <Horror,?????, <?,Medium,?????, <?,Long,?????, <??,High,???, <???,Actor A,??>, <???,Actor C,??>, <???,Director X,??>, <????, Director Z, ?>, <?????, Young>, <?????, Old></p> <p>Continue to apply the same algorithm to process the two samples given below and enumerate the resulting specific and general hypothesis.</p>		10	1	4																																																		

management, concurrency control, and recovery mechanisms.

Q. No	Question
1.	<p>Depict the given ER model to relational schema by following appropriate mapping steps.</p>  <pre> erDiagram CUSTOMER --o{ ADDRESS : "has" CUSTOMER --o{ ORDERS : "gets" CUSTOMER --o{ BILL : "has" VENDOR } --o{ PRODUCT : "supplies" EMPLOYEE } --o{ VENDOR : "worksFor" PRODUCT }--o{ QUANTITY : "has" PRODUCT }--o{ PRODUCTCATEGORY : "belongsTo" BILL }--o{ DELIVERY : "has" } </pre>



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: C2 + TC2

Programme Name & Branch

: M. Tech (All Specialization)

Course Code and Course Name

: MCSE506L & Database Systems

Faculty Name(s)

: Dr. Manjula & Dr. Ramanathan L

Class Number(s)

: VL2024250502063

Date of Examination

: 29/01/2025 2.00 PM – 3.30 PM

Exam Duration

: 90 minutes

Maximum Marks: 50**General instruction(s):**

- Answer All Questions

Course Outcomes (Type the CO statements covered in this question paper. Use the CO number as per the syllabus copy)

CO1 - Design and implement a database depending on the business requirements, considering various design issues.

CO2 - Understand the concepts of Indexing, Query optimization, transaction management, concurrency control, and recovery mechanisms.

No	Question	M	CO	B
1.	<p>Depict the given ER model to relational schema by following appropriate mapping steps.</p>	10	CO1	BI
	<p>A file consists of 30,000 un-spanned records each of size 100 bytes. Block size is 1024 bytes. If clustering indexing is used on a field such that every value is repeated 6 times. Size of block pointer is 6 bytes and clustering field value is of 9 bytes.</p> <ol style="list-style-type: none"> Find the blocking factor for data file and index file (4 marks) How many blocks are needed to store index file and data file (4 marks) Find the number of block access needed to access data and assuming index file is in hard disk and all blocks addresses are available. (2 marks) 	10	CO1	BI
	Describe the distinctions between external, internal, and conceptual schemas in	10	CO1	B

4.	<p>b) A sender using the Go-Back-N protocol transmits 10 packets to a receiver. The receiver acknowledges the first 3 packets but drops the 4th acknowledgment due to network congestion. The sender has a timer for retransmissions.</p> <p>Question:</p> <p>Analyze the behaviour of the Go-Back-N protocol in this scenario. How does the sender handle the missing acknowledgment for the 4th packet? Explain the role of the timer in this process and evaluate the trade-offs of retransmitting all unacknowledged packets versus just the lost one.</p>	5
5.	<p>A sender wants to transmit a 7-bit data sequence 1101011 using a generator polynomial $G(x)=x^3+x+1$</p> <ul style="list-style-type: none"> • Compute the CRC codeword that the sender will transmit. • If the receiver receives the sequence 1101011010, verify whether the received data contains any errors. 	CO 3 10

SCHOOL OF
ENGINEERING**VIT**

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

COMPUTER SCIENCE AND

CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: F2 +TF2

Programme Name & Branch	:	M.Tech	
Course Code and Course Name	:	MCSE505L	
Faculty Name(s)	:	Prof. Manikandan K, and Prof. Vijayakumar K	
Class Number(s)	:	VL2024250501775, VL2024250501786	
Date of Examination	:	01-02-2025	
Exam Duration	:	90 minutes	Maximum Marks: 50

General instruction(s):

- Answer All Questions
- M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)
- Course Outcomes

CO1: Explore the basics of Computer Networks and various performance metrics.

CO2: Interpret the application layer services and their protocols

CO3: Evaluate the requirements for reliable services and implications of congestion at the transport layer services.

Q. No	Question	M	CO	BL
1.	<p>A company has two offices connected through a packet-switching network. The sending host in Office A sends a packet of size $L = 1,000$ bytes to a receiving host in Office B. The packet switch between the two offices uses store-and-forward switching.</p> <p>The transmission rate between the sending host and the switch is $R_1 = 10$ Mbps. The transmission rate between the switch and the receiving host is $R_2 = 5$ Mbps. The propagation delay on the first link (Office A to switch) is $d_1 = 2$ km, and on the second link (switch to Office B) is $d_2 = 5$ km. The signal propagation speed is $s = 2 \times 10^8$ m/s.</p> <p>Assume the queuing delay at the switch is $T_{psp} = 2$ ms, and the processing delay is $T_{pp} = 1$ ms.</p> <p>Question:</p> <p>Calculate the total end-to-end delay for the packet to travel from the sending host in Office A to the receiving host in Office B, considering all components (transmission, propagation, queuing, and processing delays).</p>	10	CO 1	BL 3
2.	<p>Explain how the concepts of service layering and encapsulation are applied during the following</p> <p>A client on a laptop initiates a connection to a web server using TCP over the internet. The client sends an HTTP request for a webpage, and the server responds with the requested content. The network between the client and the server follows a layered architecture (e.g., TCP/IP stack) and utilizes service layering and encapsulation.</p>	10	CO 1	BL 2
3.	a) Analyze the possible reasons for the failure of an email sent from a company to its client by considering SMTP's working principles and interaction with other protocols (like DNS and TCP). Propose a systematic troubleshooting process to identify and resolve the issue, including how each layer in the network stack (application, transport, network) contributes to successful email delivery.	10	CO 2	BL 4
4.	a) Design chat application using TCP socket API's. The application must support any one of the features like private messaging, group chats, and message delivery confirmation.	5	CO 2	BL 3



VIT

Vellore Institute of Technology
Approved to be University under sec 3 of UGC Act, 1956

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: D2 + TD2

Original Image

0	255	127
10	250	50
200	100	50

4. Segment the given image using region growing algorithm. Use the points marked in the image as seed points and assume the threshold value as 1.

5	6	1	1	1	0	(0)
6	7	7	1	1	1	6
3	(6)	7	7	7	6	5
4	3	7	7	7	5	3
3	3	2	1	(3)	3	4
2	2	1	2	1	2	3
1	1	(1)	1	1	2	2

10 1 6

5. How the dilation and erosion morphological operations affect the shape of objects in a binary image? Explain with example

10 1 1



VIT[®]

Vellore Institute of Technology
(Desired to be University under section 3 of UGC Act, 1956)

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: D2 + TD2

Programme Name & Branch	: M.Tech & Computer Science and Engineering
Course Code and Course Name	: MCSE605L and Machine Vision
Faculty Name(s)	: Dr. G. N. Balaji
Class Number(s)	: VL2024250502175
Date of Examination	: 30.01.2025
Exam Duration	: 90 minutes
	Maximum Marks: 50

Answer All Questions

- M - Max mark; CO – Course Outcome; BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Create)
- CO1: Discover and understand enhancement, segmentation and morphological operation on images for further analysis
- CO2: Acquire the knowledge of various image transforms, wavelets and multiresolution analysis for better interpretation.

Q. No	Question	M	CO	BL																		
1.	Discuss the relationship between the physical properties of a camera, image formation, and the resulting image quality. What are the critical factors to consider when capturing an image for further processing?	10	1	2																		
2.	a) Explain the concepts of sampling and quantization in digital image processing. How do they affect image quality and resolution? (6 Marks)			2																		
	b) Apply down sampling in the following image (4 Marks)	10	2	3																		
	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>(3)</td><td>3</td><td>(3)</td><td>3</td></tr> <tr> <td>9</td><td>9</td><td>9</td><td>9</td></tr> <tr> <td>(3)</td><td>3</td><td>3</td><td>3</td></tr> <tr> <td>9</td><td>9</td><td>9</td><td>9</td></tr> </table>	(3)	3	(3)	3	9	9	9	9	(3)	3	3	3	9	9	9	9					
(3)	3	(3)	3																			
9	9	9	9																			
(3)	3	3	3																			
9	9	9	9																			
3.	a) Perform Histogram equalization for the given image, plot the original image histogram, and equalized image histogram (7 Marks)																					
	Original Image																					
	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>Gray Level</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr> <tr> <td>No. of pixels</td><td>550</td><td>900</td><td>650</td><td>150</td><td>300</td><td>250</td><td>110</td><td>90</td></tr> </table>	Gray Level	0	1	2	3	4	5	6	7	No. of pixels	550	900	650	150	300	250	110	90	10	2	4
Gray Level	0	1	2	3	4	5	6	7														
No. of pixels	550	900	650	150	300	250	110	90														
	b) Apply Logarithmic Transformation in the following 8 bit image, write necessary equations (3 Marks)																					



VIT
Vellore Institute of Technology

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: C2 + TC2

	database design. Furthermore, how do these three schema layers enable logical and physical data independence, and what implications does this independence have for database management and application development?	
4.	Given a relation R(P, Q, R, S, T, U, V, W, X, Y) and Functional Dependency set FD = {PQ → R, P → ST, Q → U, U → VW, and S → XY}, determine whether the given R is in 3NF? If not convert it into 3 NF.	CO1 10
5.	Consider the following Schema Student(StudId, , Fname, Lname, Major) Grades(StudId, CrsId, Grade) Course(CrsId, Cname, Credit_Points, Dept) Construct an Initial Query Tree and Optimized Query Tree using Heuristic Query Optimization Techniques for the SQL query given below: SELECT StudId, Fname FROM Student s, Grades g, Course c WHERE s.StudId=g.StudId AND g.CrsId= c.CrsId AND Grade = 'S' AND Dept = 'DataScience' AND Major = 'AI' pStudId, FName (σ s.StudId=g.StudId ∧ g.CrsId=c.CrsId ∧ Grade='S' ∧ Major = 'AI' ∧ Dept = 'DataScience' (Course × (Student × Grades)))	CO2 10



Shot on moto g31

1 Feb 2025, 7:22 pm



40 - 013 - 0065



VIT*

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

0.75 - 0.25 + 0.5
0.77 + 0.32
~~0.02~~
= 0.27

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: D2 + TD2

Programme Name & Branch

: M.Tech & Computer Science and Engineering

Course Code and Course Name

: MCSE605L and Machine Vision

Faculty Name(s)

: Dr. G. N. Balaji

Class Number(s)

: VL2024250502175

Date of Examination

: 19.03.2025

Exam Duration

: 90 minutes

Maximum Marks: 50

Answer All Questions

- M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)
- CO1: Discover and understand enhancement, segmentation and morphological operation on images for further analysis
- CO2: Acquire the knowledge of various image transforms, wavelets and multiresolution analysis for better interpretation.

Q. No	Question	M	CO	BL
1.	<p>Apply DCT in the following 2D image and discuss the steps involved in converting the image from spatial domain to frequency domain with necessary equations</p> $\mathbf{f} = \begin{bmatrix} 52 & 55 & 61 & 66 \\ 62 & 59 & 55 & 59 \\ 63 & 58 & 60 & 64 \\ 64 & 59 & 55 & 59 \end{bmatrix}$	10	2	3
2.	<p>Choose a suitable kernel to remove the noise in the following signal and explain its impact, further discuss how convolution in the spatial domain corresponds to multiplication in the frequency domain.</p>	10	2	4
3.	<p>Identify suitable technique to recognize the following template in the given image and explain the steps involved.</p>	(5 Marks)		



VIT[®]

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

9 - 0.63 + 17.01 - 2.09
5 + 17 + 31.45 + 1.60
+ 1.53 - 41.08
- 0.86

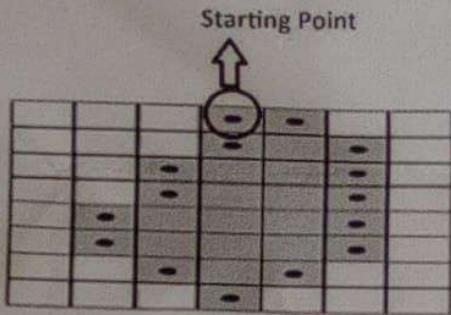
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: D2 + TD2

Template	Reference						10	3	5
	1	1	1	-1	1	1			
7 7 5	1	5	7	7	5	1			
4 3 2	1	2	4	3	2	1			
3 8 2	1	2	3	8	2	1			
	1	2	2	2	1	1			
	1	1	1	1	1	1			

b) Discuss the steps to identify the object in an image when template matching fails
(5 Marks)

4. Extract the Chaincode using 4-direction and 8- direction, also find curvature and bending energy using the 8-direction chaincode in the given image $f(x,y)$. Use the starting point mentioned



31 $2^2 + 2^2$ $2^2 + 2^2$
 $4 + 4$

5.

Point	X1	X2	Class
Point 1	1	2	A
Point 2	2	3	A
Point 3	3	3	B
Point 4	6	5	B
Point 5	7	7	B
Point 6	8	B	B
Point 7	2	2	A
Point 8	5	5	B
Point 9	4	4	A
Point 10	5	3	B

$$\begin{aligned}
 & (1-3)^2 + (2-4)^2 = 10 \\
 & (2-3)^2 + (3-4)^2 = 2 \\
 & (3-3)^2 + (3-4)^2 = 1 \\
 & (6-3)^2 + (5-4)^2 = 10 \\
 & (6-3)^2 + (7-4)^2 = 25 \\
 & (7-3)^2 + (7-4)^2 = 20
 \end{aligned}$$

Apply k-NN classifier and classify Point P ($X_1=3, X_2=4$) assume K=5. To calculate the distance between the new point and each training point use Euclidean distance. Also, mention the options to choose an optimal k value.



VIT[®]

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

**SCHOOL COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST – II (Re-Test)
WINTER SEMESTER 2024-2025**

REG.NO.:

17 AP

SLOT: B1+TB1

Programme Name & Branch : M. Tech Computer Science and Engineering (MCI & MCS)

Course Code and Course Name : MCSE612L – Cyber Security

Faculty Name(s) : Rhymend Uthariaraj V

Class Number(s) : VL2024250503999

Date of Examination : 09-04-2025

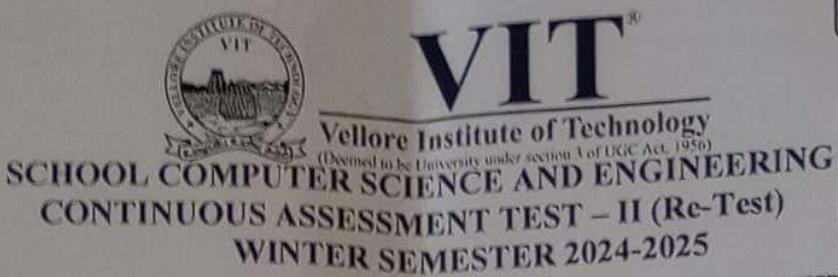
Exam Duration : 90 minutes

Maximum Marks: 50

General instruction(s):

- Answer All Questions

Q. NO	Questions	Marks
1.	<p>Weak authentication remains a primary vulnerability in cyber security.</p> <ol style="list-style-type: none"> Explain how weak authentication mechanisms contribute to security breaches. Analyze a real-world cyber-attack where weak authentication played a critical role. Suggest an improved authentication framework to mitigate such attacks. 	10
2.	<p>A company uses passwords of length L composed of uppercase, lowercase, numbers, and special characters. The total possible character set is C. Derive a mathematical model to compute the entropy (H) of a password and analyze its strength for five different values of L and C. Suggest a guideline to create a strong password for a University Data Centre with hybrid network Architecture.</p>	10
3.	<p>a) A monitoring system in a large enterprise logs 1,000,000 security events per day. If 0.1% of these events are flagged as suspicious and 5% of the suspicious events are false positives, calculate:</p> <ol style="list-style-type: none"> The total number of suspicious events flagged per day. The number of events that are likely true positives (i.e., real suspicious events). 	5



REG.NO.: 24 MCS 6006

SLOT: B1+TB1

<p>b) Discuss the role of ethical hacking in modern threat management. Compare any two ethical hacking techniques and explain how they contribute to the identification and mitigation of threats.</p>	5
<p>a) Discuss the role of access control and monitoring systems in ensuring organizational cyber security. Compare and contrast at least two access control models and explain how continuous monitoring systems contribute to the detection of unauthorized access.</p>	5
<p>b) An organization faces a vulnerability in its operating system with an initial exploitation probability of 20% and an estimated breach cost of \$100,000. After applying the patch, the exploitation probability drops to 5%.</p> <ul style="list-style-type: none"> i. Calculate the expected risk (in dollars) before and after patching. ii. Determine the risk reduction and discuss the cost-benefit of timely patch management. 	5
<p>Discuss the role of cyber security tool in mitigating Distributed Denial of Service (DDoS) attacks. illustrate with any suitable tool and explain how they help detect, filter, and mitigate DDoS traffic. Provide a real-world example of an organization that effectively used such tools.</p>	10



VIT

Vellore Institute of Technology
Approved by the University Grants Commission of India

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: C2 + TC2

Programme Name & Branch

: M.Tech(All Specialization)

Course Code and Course Name

: MCSE506L & Database Systems

Faculty Name(s)

: Dr. Manjula R & Dr. Ramanathan L

Class Number(s)

: VL2024250502031 and VL2024250502063

Date of Examination

: 18/03/2025

Exam Duration

: 90 minutes

Maximum Marks: 50

Q. No	Question	M	CO
1.	<p>Draw the serializable graphs for the schedule S1 and state whether the schedule is (i)conflict serializable or not (ii)view serializable or not. If the schedule is conflict/view serializable, write down its equivalent serial schedule</p> <p>S1: :r1(X);r2(Z);r3(X);r1(Z);r2(Y);r3(Y);w1(X);w2(Z);w3(Y);w2(Y);</p>	10	C02
2	<p>Consider two transaction T0 and T1, with the values A=1000, B=2000 and C=700</p> <p>T0: read(A) A:=A-50 Write(A) Read(B) B=B+50 Write(b) commit</p> <p>T1: read(C) C=C-100 Write(C) Commit</p> <p>If the system failure occurs after all the transactions is committed,</p> <ol style="list-style-type: none"> What will be the sequence of content in each of the log file and database if deferred update is applied. Also, what will be the sequence of content in each of the log file and database if immediate update is applied Finally give your inference about the two techniques in recovery the database when a failure occurs 	10	C01
3	<p>a. What kind of parallelism can be applied in a data warehouse to improve query performance by breaking down complex queries into smaller sub-tasks, such as filtering, sorting, and aggregating, and executing these sub-tasks concurrently using multiple processors or nodes, and what benefits does it provide?" Also explain the parallelism with neat diagram. [4]</p>	10	C02
	Marks]		

→ Intra query



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: C2 + TC2

Interoperation

- b. How can parallelism be applied to scientific simulations, such as weather forecasting, to improve performance, and what are the benefits of doing so? [3]

Marks]*Intra operation*

- c. What parallelization technique can be applied to complex operations like matrix multiplications to accelerate their execution, and what are the resulting benefits? [3 Marks]

Apply data localization step of distributed query processing for the given query

Select emp_name, city from employee, address where employee.empid = address.id and state = TamilNadu

10

CO3

BL4

Illustrate query tree for each of this case: Global Relation, Fragments and Fragments with optimized results

Consider a distributed database management system (DBMS) with the following schema:

- Employees (emp_id, name, department) stored at New York
- Projects (proj_id, emp_id, project_name) stored at Los Angeles

Each Projects tuple is 50 bytes long, and a page can hold 200 Projects tuples. The Projects relation occupies 1000 pages. Each Employees tuple is 40 bytes long, and a page can hold 300 Employees tuples. The Employees relation occupies 500 pages.

10

CO3

BL3

- a. What is the cost of answering the query:

"Select * from Employees E, Projects P where E.emp_id = P.emp_id" using a page-oriented nested loop join at New York and shipping the result to Chicago? [5 Marks]

- b. Using the same distributed DBMS and schema as before, what is the cost of answering the query:

"Select * from Employees E, Projects P where E.emp_id = P.emp_id" by moving both relations to Los Angeles, performing a sort-merge join, and then moving the result to New York? [5 Marks]



VIT

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.: _____

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: A2

Programme Name & Branch	: M. Tech & CSE, AI/ML
Course Code and Course Name	: MCSE602L - Machine Learning
Faculty Name(s)	: Prof. Mythili T & Prof. Varalakshmi M
Class Number(s)	: VL2024250502112 & VL2024250502116
Date of Examination	: 16-03-2025
Exam Duration	: 90 minutes
	Maximum Marks: 50

Course Outcomes

- CO2 - Analyze and apply the suitable supervised learning methods for real-world Problems
- CO3- Identify and integrate more than one technique to enhance the performance of Learning

Answer All Questions

Q. No	Question		M	CC	BL																		
1.	A linear regression model is used to predict student attendance based on several features (e.g., facial landmarks, lighting conditions, room temperature). Below are the training and validation errors at different stages of model training:		10	3	4																		
	<table border="1"> <thead> <tr> <th>Iteration n</th> <th>Training Error</th> <th>Validation Error</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>0.45</td> <td>0.50</td> </tr> <tr> <td>2</td> <td>0.40</td> <td>0.48</td> </tr> <tr> <td>3</td> <td>0.38</td> <td>0.50</td> </tr> <tr> <td>4</td> <td>0.37</td> <td>0.60</td> </tr> <tr> <td>5</td> <td>0.36</td> <td>0.65</td> </tr> </tbody> </table>	Iteration n	Training Error	Validation Error	1	0.45	0.50	2	0.40	0.48	3	0.38	0.50	4	0.37	0.60	5	0.36	0.65				
Iteration n	Training Error	Validation Error																					
1	0.45	0.50																					
2	0.40	0.48																					
3	0.38	0.50																					
4	0.37	0.60																					
5	0.36	0.65																					

a) Analyse the Iteration 4 and Iteration 5 of the model and is the model experiencing over fitting or under fitting? Justify your answer by analysing the trends in the training and validation error.

b) At which iteration(s) would you consider the model to be best fit (not over fitting or under fitting)? Justify your answer using the error values.

c) If you were to apply regularization in Iteration 5, what impact do you expect on the training and validation error, and why?

**VIT**Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: A2

2.	Assume you are developing an attendance recognition system using machine learning techniques to identify students based on facial recognition or other biometric features. The system collects a variety of features such as facial landmarks, timestamps, room temperature, lighting conditions, and background noise. Specifically, what learning method would you recommend for training the system? Additionally, which perceptron model would be most suitable for training the system to achieve higher accuracy? Explain the identified suitable model with neat diagram.	10	2	5
3.	Assume you are provided with the following sample dataset:	10	2	3

Feature 1 (x_1)	Feature 2 (x_2)	Label (y)
2	3	1
4	5	1
1	2	0
3	3	0
5	6	1
6	7	1
1	1	0
4	4	0

Apply the Support Vector Machine (SVM) algorithm to analyse whether the given dataset is linearly separable. Additionally, generate a visualization to demonstrate the separation, including the decision boundary and support vectors.

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: A2

4. Given dataset with features like Facial Landmarks (X_1, X_2), Timestamp, Room Temperature, Lighting (lux), Background Noise, and the target variable Attendance, how can the AdaBoost algorithm be used to compare the predictive accuracy of Room Temperature versus Facial Landmarks as weak learners (decision stumps)?

Additionally, consider six randomly misclassified samples in the dataset, which are used to create a new dataset. How would these misclassifications impact the model's performance, and which feature—Room Temperature or Facial Landmarks—proves more effective in improving the accuracy of attendance classification?

Student ID	Facial Landmarks (X_1)	Facial Landmarks (X_2)	Timestamp (seconds)	Room Temp ($^{\circ}\text{C}$)	Lighting (lux)	Background noise	Attendance (Target)
1	0.23	0.45	120	22.5	200	50	1
2	0.24	0.46	121	22.6	195	52	0
3	0.22	0.43	122	22.4	210	48	1
4	0.25	0.47	123	22.7	205	49	0
5	0.20	0.40	124	22.8	215	51	1
6	0.21	0.42	125	22.3	202	50	0

Misclassified Samples

7	0.26	0.48	130	22.9	210	53	0
8	0.18	0.38	127	23.0	190	47	1
9	0.24	0.45	126	22.7	205	50	0
10	0.22	0.44	128	22.5	215	52	1
11	0.23	0.46	129	22.8	200	49	1
12	0.20	0.42	132	22.6	190	48	0

5. a) In the development of a facial recognition-based attendance system, the research team is evaluating different machine learning models with varying complexities for predicting student attendance. The team has trained three models, each with different numbers of features, and calculated their corresponding training errors (empirical risks) and complexity penalties mentioned in the below table:

Model	Training Error(R_{emp})	Complexity
-------	------------------------------------	------------

5 2



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: A2

Model 1	0.12	Penalty(C)	
Model 2	0.08	1	
Model 3	0.05	2	

- i. Calculate the structural risk for each model (Note: regularization parameter $\lambda=0.1$)
- ii. Based on Structural Risk Minimization (SRM), which model should the research team choose? Justify your answer.
- iii. How does the trade-off between training error and complexity penalty affect the choice of the best model for the attendance prediction task?

b) Consider a classification problem where you are given a dataset of students and the target is to predict whether each student will attend class. Compute the Average Voting for each student. Determine the final prediction (attend or absent) based on the average voting method: (Note: If the average prediction is greater than or equal to 0.5, the final prediction is attend (1), If the average prediction is less than 0.5, the final prediction is absent (0)).

5 2

Student ID	Model Prediction(P1)	Model Prediction(P2)	Model Prediction (P3)	Final Prediction (Using Average Voting)
1	0.8	0.6	0.7	?
2	0.3	0.4	0.2	?
3	0.9	0.95	0.85	?
4	0.45	0.5	0.55	?
5	0.1	0.2	0.15	?



School of Computer Science and Engineering

Winter Semester 2024-25

RECAT II

SLOT:F1+F2+TF1+TF2

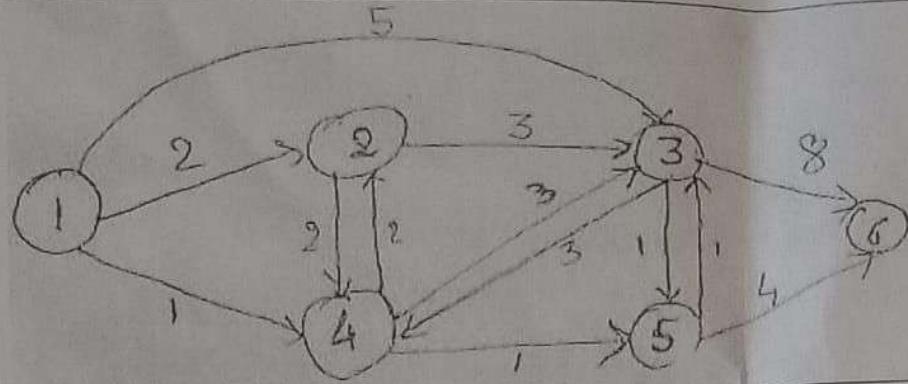
Programme Name & Branch: M.Tech(All Branches)

Course Name & Code: Computer Networks (MCSE505L)

Exam Duration: 90 Min.

Maximum Marks: 50

Q. No .	Question	Max Marks	CO
1.	<p>Make the routing table for R1, R2, R3, R4 and R5</p>	10	CO4
2.	<p>Assume that data of 5000 bytes has to be sent from a sender A (192.168.64.3) is attached to Ethernet Network to the receiver B (200.1.2.5) is a part of X.25 network. The intermediate network between sender's network and receiver's network is FDDI. Router A and Router B connects sender and receiver network to the intermediate network. Trace and Sketch each IP4 fragment (in Hexadecimal) and assembly at Router A and Router B respectively. Assume the Identification of IPv4 datagram in decimal is 68 and the upper layer protocol is OSPF(Assume default values wherever necessary)</p>	10	CO4
3.	<p>List the various actions to be performed in Link State Routing. Apply Dijkstra's algorithm for the given sample graph to form the shortest path tree and construct the routing table for root node A</p>	10	CO5



- | | | | |
|----|--|----|-----|
| | 4. With a neat diagram discuss in detail how SDN is defined by three fundamental principles which differentiate it from traditional networking. Examine the architectural diagram that depicts various protocols that helps to achieve the above | 10 | CO5 |
| 5. | Describe briefly the packet format in TCP and how flow control is achieved in the above. Imagine a TCP connection is transferring a file of 6000 bytes. The first byte is numbered 10010. What are the sequence numbers for each segment if data is sent in five segments with the first four segments having rwnd = 1,000 bytes and for the last segment having rwnd = 2,000 bytes? | 10 | CO3 |

VIT[®]Vellore Institute of Technology
The Government Engineering College of Tamil Nadu

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: F2+TF2

Programme Name & Branch : M.Tech (Common to all Branches)
 Course Code and Course Name : MCSE505L-Computer Networks
 Faculty Name(s) : Dr.Manikandan K, Dr Vijaya kumar K
 Class Number(s) : VL2024250501775, VI.2024250501786
 Date of Examination : 21-Mar- 2025
 Exam Duration : 90 minutes

Maximum Marks: 50

General instruction(s):**Answer All Questions**

M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)

Course Outcomes:

CO3. Evaluate the requirements for reliable services and implications of congestion at the transport layer services. **CO4.** Analyse various functionalities required in the control and data plane at network layer services.

Q. No	Questions	M	CO	BL
		5	3	5
1.	a) Consider a scenario where multiple users are sharing the same bandwidth in a local area network (LAN). How would you design a preventive congestion control protocol to ensure fairness, where each user gets a fair share of the bandwidth?	5	3	5
	b) Consider an instance of TCP's Additive Increase Multiplicative Decrease (AIMD) algorithm where the window size at the start of the slow start phase is 2 MSS and the threshold at the start of the first transmission is 8 MSS. Assume that a timeout occurs during the fifth transmission. Find the congestion window size at the end of the tenth transmission.			
2.	The Internet Service Provider (ISP) is planning to expand its network by adding more routers in the core and extending the backbone to new customer locations. The expansion includes adding new switches and ensuring efficient routing. The ISP's core network is currently using Open Shortest Path First (OSPF) Areas to segment the network for scalability, and each building or data centre is within its own area. Questions: i) How should the ISP design the OSPF areas to ensure optimal performance and avoid issues like LSA (Link-State Advertisement) flooding and routing loops? What factors should be considered when deciding the number of areas, such as router capabilities, network size, and topology? ii) Explain how the Shortest Path First (SPF) algorithm works in OSPF. Given a network topology, how does OSPF calculate the routing table based on the SPF tree? Provide an example with three routers connected in a triangular topology.	4	3	10

Page 1 of 2



3.	An ISP is granted a block of addresses starting with 140.80.0.0/16. The ISP wants to distribute these blocks to 2600 customers as follows. i) The first group has 200 medium-size businesses; each needs 128 addresses. ii) The second group has 400 small businesses; each needs 16 addresses. iii) The third group has 2000 households; each needs 4 addresses. Design the sub-blocks and give the slash notation for each sub-block. Find out how many addresses are still available after these allocations.	10	4	6
4.	Consider the network shown below. Show the operation of Distance vector routing algorithm for computing the least cost path from 'A' to all destinations. Also explicitly list all the shortest path routes from 'A' to all destinations that are the result of the algorithm's computation. Depict all the steps through diagram.	10	4	4
5.	<p>a) You are given the following IPv4 header fields for a fragmented packet: Version: 4, IHL: 5 (20 bytes header), Total Length: 0x0800 (2048 bytes), Identification: 0x1234, Flags: 0x01 (DF flag), Fragment Offset: 0x0000 (first fragment), TTL: 128, Protocol: 6 (TCP)</p> <p>Questions:</p> <ul style="list-style-type: none"> i) What is the size of the data part of the packet (excluding the IPv4 header)? ii) Based on the Flags field (0x01), what does this indicate regarding the packet's fragmentation? iii) If the Fragment Offset is 0x0000, what does that indicate about the position of this fragment? <p>b) Assume you are comparing two packets: one is IPv4, and the other is IPv6.</p> <p>IPv4 Packet: Version: 4, IHL: 5, Total Length: 128 bytes, Protocol: 17 (UDP), Source Address: 192.168.1.1, Destination Address: 192.168.1.2,</p> <p>IPv6 Packet: Version: 6, Payload Length: 100 bytes, Next Header: 17 (UDP), Source Address: 2001:0db8:85a3:0000:0000:8a2e:0370:7334, Destination Address: 2001:0db8:85a3:0000:0000:8a2e:0370:1234</p> <p>Questions:</p> <ul style="list-style-type: none"> i) What is the total length of the IPv4 packet, including the header and data? ii) What is the IPv6 packet's total size, considering both the header and data? iii) What is the difference in size between the IPv4 and IPv6 packets, and why might this difference exist? 	5	4	3



VIT[®]

REG.NO.: 24MCS0017

Vellore Institute of Technology
 (Deemed to be University under section 3 of UGC Act, 1956)
SCHOOL COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

SLOT: B1+TB1

Programme Name & Branch	: M. Tech Computer Science and Engineering (MCI & MCS)
Course Code and Course Name	: MCSE612L - Cyber Security
Faculty Name(s)	: Rhymend Uthariaraj V
Class Number(s)	: VL2024250503999
Date of Examination	: 17-03-2025
Exam Duration	: 90 minutes
	Maximum Marks: 50

General instruction(s):

- Answer All Questions

Q. No	Questions	Marks
1.	Analyse the impact of common software vulnerabilities - such as buffer overflows, SQL injection, cross-site scripting (XSS) and unpatched zero-day vulnerabilities - on modern cyber security infrastructures . Examine how the transition to proactive cyber security strategies (e.g., Patch Management and vulnerability scanning) strengthens organizational security.	10
2.	Design an Access Control List (ACL) for a shared folder in a small organization with four categories of IT professionals (Design team, Development team, Deployment team and Cyber security team). The folder contains both confidential and non-confidential files. How would you configure ACLs to ensure that only authorized employees can access the files? Suggest a simple method for tracking and logging access to these files to detect any unauthorized access.	10
3.	a) Bring out an innovative threat management framework that integrates predictive threat intelligence, and proactive risk mitigation, aimed at ensuring the organization can handle both existing and future cyber threats.	5
	b) A network has been subjected to a Distributed Denial-of-Service (DDoS) attack. The attacker is using 2000 botnets , each sending 50 requests per second to the target server. The server can handle 5000 requests per second before it becomes unresponsive.	5



VIT®

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

SCHOOL COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2024-2025

REG.NO.:

SLOT: B1+TB1

	<ul style="list-style-type: none">i. Calculate the total number of requests per second (RPS) generated by the botnets.ii. Determine how long the server will continue to function properly before reaching its maximum capacity if the attack continues at the current rate.iii. If the attacker increases the number of botnets by 25%, how many more requests per second will be sent to the server?	
4.	<p>a) Evaluate the effectiveness of traditional perimeter defences, such as firewalls and IDPS, in safeguarding the inner perimeter of a network within today's rapidly evolving cyber landscape. How do these conventional methods stack up against more modern approaches, including zero-trust architecture, and micro-segmentation?</p> <p>b) A company uses a VPN with two-factor authentication (2FA) for remote access, requiring RSA-based token generators (first factor) and SMS-based one-time passcodes (second factor). The company has 5000 users, each logging in 5 times per day. The RSA system processes 2000 tokens per second, taking 0.5 seconds per token, and the SMS gateway sends 5000 SMS per second, taking 1 second to deliver.</p> <ul style="list-style-type: none">i. How long will it take to authenticate all remote users for one round of login attempts?ii. If the system experiences a load of 20% more users logging in simultaneously during peak hours, how much additional load (in terms of processing time) will the system experience	5
5.	<p>Using a simple scenario, demonstrate how a Cyber security tool can be used to defend against a specific cyber-attack. In your explanation, include the following:</p> <ul style="list-style-type: none">i. The tool's objective in the given scenarioii. A step-by-step flow of how the tool worksiii. Example code or configuration for implementationiv. The tool's limitations in defending against the attackv. Suggestions to enhance the tool's effectiveness	10



Final Assessment Test - April 2025

Course: MCSE505L - Computer Networks

Class NBR(s): 1775/1786

Slot: F2+TF2

Max. Marks: 100

Time: Three Hours

- > KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- > DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer ALL Questions

(10 X 10 = 100 Marks)

Q1. What are the needs for layered architecture in network communication? How the less number of layers in TCP/IP compensates the operations carried out by the layers in ISO/OSI model in network communication? Justify your answer.

Q2. a) For each of the following four networks, discuss the consequences if a connection fails. [5]

- Five devices arranged in a mesh topology
- Five devices arranged in a star topology (not counting the hub)
- Five devices arranged in a bus topology
- Five devices arranged in a ring topology

b) We transmit data directly between two servers 6,000 km apart through a geostationary satellite situated 10,000 km from Earth exactly between the two servers. The data enters this network at 100Mb/s. [5]

- Find the propagation delay if data travels at the speed of light (2.3×10^8 m/s).
- Find the number of bits in transit during the propagation delay.
- Determine how long it takes to send 10 bytes of data and to receive 2.5 bytes of acknowledgment back.

$10^8 \times 10^6$
10000

Q3. Why do you think DNS uses UDP, instead of TCP, for its query and response messages? Describe the functions of SNMP, SMI, MIB with an example.

a) How to improve the quality of service in networking? Describe the traffic shaping techniques used to improve the quality of service. [5]

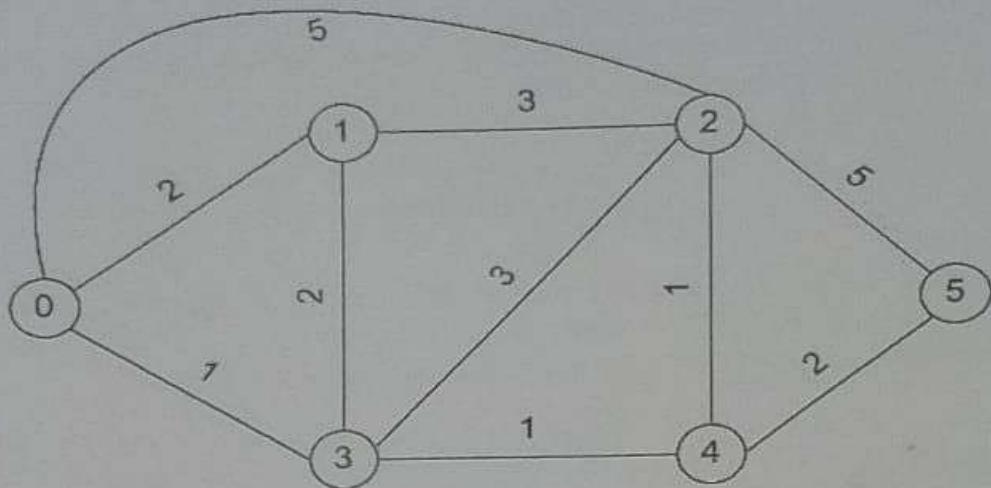
b) Suppose you are designing the communication part of the following three Internet applications. The table below shows the QoS (quality of service) the general users demand for the three types of network applications. [5]

SNTP

	Loss	Time Sensitive
Web	no loss	no
Skype	loss tolerant	yes
Online Game	no loss	yes

Which transport layer services, will you choose to transfer web data, game data and skype calls? Justify your answer?

5. Explain the phases of link state routing protocol. Find the link state information of each node and shortest path from node '0' to all the members in the below network.



6. Explain the routing protocol taxonomy with neat diagram. Compare performance evaluation between link state routing protocol and distance vector routing protocol.

7. Two neighbouring nodes A and B use Selective Repeat ARQ with a 3 bit sequence number. Assuming that Communication between A and B is bidirectional, show the window position and frame flow for the following sequence of events :-

- Initial Position. Before A and B sends any frames.
- After A sends 0, 1, 2 and B sends 2 Data frames with ACK 0 and ACK1
- A sends frames 3, 4 and then receives NAK3 from B. B receives the corrupted frame 3 and sends 1 Data frame with ACK4.
- A sends frames 3, 5, 6. The acknowledgement send by B is lost and A does not receive it.

Indicate the Sender's Window Size and Receiver's Window (S_r , S_n and S_w) at both A and B.

0, 3, 1, 4, 2, 5
S →

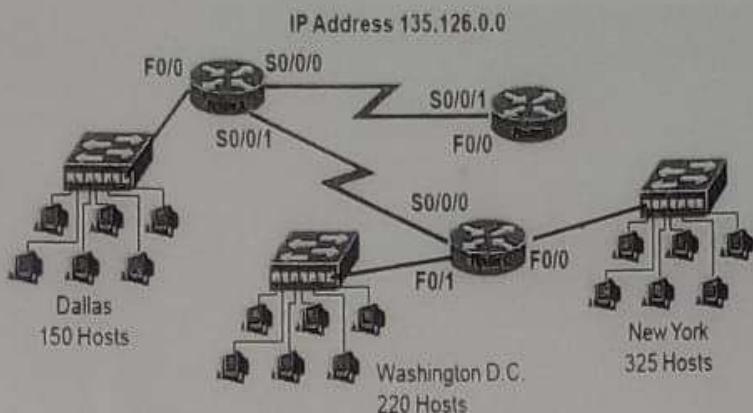
8

Given the dataword 1010011010 and the divisor 10111,

- Show the generation of the codeword at the sender site (using binary division).
- Suppose the leftmost bit of the message is inverted due to the noise on the transmission link. What is the result of the receiver's CRC calculation? How does the receiver know that an error has occurred?

9.a)

Based on the information in the graphic shown, design a network addressing scheme (IP Address 135.126.0.0/16) that will supply the **minimum number of subnets**.



Calculate the following

- Extra subnets required for 100% growth.
- Total number of subnets needed.
- Number of addresses needed for 100% growth in the largest subnet.

OR

9.b)

An ISP is granted a block of addresses starting with 120.100.0.0/16 (65,536 addresses). The ISP needs to distribute these addresses to three groups of customers as follows:

- The first group has 64 customers; each needs 256 addresses.
- The second group has 128 customers; each needs 128 addresses.
- The third group has 128 customers; each needs 64 addresses.

Design the sub-blocks and find out how many addresses are still available after these allocations.

10.a)

Discuss the limitations of wireless Adhoc Network. How the Hidden/Exposed terminal problem affects the communication in wireless LANs? Write the solution for it?

OR

10.b)

Elaborate the asymmetric key cryptography with neat block diagram. Discuss about the cyber security best practices followed in any organizations with suitable scenario. How to prevent cyber threat and attacks in a network. Suggest your recommendations.


Final Assessment Test - April 2025

Course: MCSE605L - Machine Vision

Slot: D2+TD2

Class NBR(s): 2175

Max. Marks: 100

Time: Three Hours

- KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer ALL Questions

(10 X 10 = 100 Marks)

1. a) Discuss the fundamental steps in image processing. [6]
 b) Delineate Sampling and Quantization with example. [4]
2. Identify the type of noise in the following image of size 5x5 and explain the different types of filters (mean, min, max and median) to remove the noise with necessary equations.

10	10	10	9	9
10	9	10	10	10
10	12	10	11	11
10	10	10	255	10
10	12	12	12	11

10, 10, 10, 10, 10,
 10, 10, 10, 10, 11,
 12, 12, 12, 255

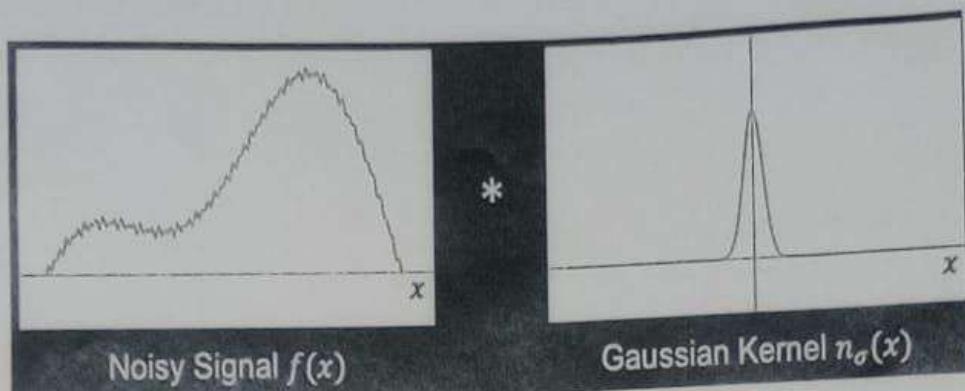
3. Perform split and merge segmentation on the following image of size 8x8 also discuss the steps involved in it.

 $f(x,y) =$

6	5	6	6	7	7	6	6
6	7	6	7	5	5	4	7
6	6	4	4	3	2	5	6
5	4	5	4	2	3	4	6
0	3	2	3	3	2	4	7
0	0	0	0	2	2	5	6
1	1	0	1	0	3	4	4
1	0	1	0	2	3	5	4

4.

Illustrate the convolution process using the given signal and kernel with necessary equations. Further discuss the role of convolution in the spatial domain and frequency domain.



5.

Discuss the essential steps involved in camera calibration using the Direct Linear Transformation (DLT) method. Explain the process of extracting intrinsic and extrinsic camera parameters.

6.

Identify suitable technique to recognize the following template in the given image and explain the steps involved. Describe the input image in detail.



7.

When multiple predicted bounding boxes overlap with the ground truth, how does the IoU metric determine the best bounding box, and what strategies can be used to refine the selection process? Explain with necessary equations.

8.

Define optical flow and describe its significance in computer vision for analyzing motion patterns in video sequences. Explain the concept of dense and sparse optical flow, highlighting their respective advantages and applications.

Feature extraction
processing
Object recognition
pre
classification

9.a)

- 4-16
4-14
- Apply k-means clustering algorithm and find the clusters where the Medicine A and C are the cluster head, if we have 5 objects (5 types of medicines) and each object have two attributes or features as shown in the table below.

9 + 9

Object	Attribute 1 (X) Weight Index	Attribute 2 (Y) pH
Medicine A	1	1
Medicine B	1	0
Medicine C	0	2
Medicine D	2	4
Medicine E	3	5

OR

- 9.b) A bank wants to automate its loan approval process based on applicants' profiles. The bank considers three categorical factors: Employment Status (Employed/Unemployed), Credit Score (Good/Poor), and Existing Debt (Yes/No). Using historical loan data, the bank aims to predict whether a new applicant's loan should be approved or rejected. Apply KNN classifier and assume K = 3.

Dataset:

Applicant	Employment	Credit Score	Existing Debt	Loan Status
A1	Employed	Good	No	Approved
A2	Unemployed	Good	No	Approved
A3	Employed	Poor	No	Approved
A4	Employed	Good	Yes	Approved
A5	Unemployed	Poor	No	Rejected
A6	Employed	Poor	Yes	Rejected
A7	Unemployed	Poor	Yes	Rejected
A8	Employed	Good	No	Approved
A9	Unemployed	Good	Yes	Approved
A10	Employed	Good	No	Approved

- 10.a) Calculate the distance measure in the given image between two pixels p and q.
- Euclidean Distance
 - City block distance
 - Chess board distance
 - M-Adjacency distance

where

- $v = [1, 2]$
- $v = [1, 2, 3, 4, 5, 6]$

where the point p is at (1,1) and q is at (5,5)

2	3	2	6	1
6	2	3	6	2
5	3	2	3	5
2	4	3	5	2
4	5	2	3	6

OR

- 10.b) Perform Histogram equalization for the following data, plot the original image histogram and equalized image histogram.

Gray Level	0	1	2	3	4	5	6	7
No. of pixels	800	1000	850	650	300	250	100	150

Ca/D/TY

4100

0.11

2.25

1.79



Time: Three Hours Max. Marks: 100
➤ KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
➤ DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer ALL Questions

(10 X 10 = 100 Marks)

In a real-world application like an e-commerce platform, how do the key components of a DBMS work together to ensure efficient data storage, retrieval, and management? Can you explain the role of each component in handling tasks such as processing customer orders, managing inventory, and generating reports?

- a. Schedule: R1(A), W1(B), R2(B), W2(A), W3(C), R3(A) [5]
b. Schedule: W1(A), R2(A), W2(B), R1(B), W3(C), R3(C) [5]

 - i. Draw the precedence graph for the given schedule.
 - ii. Identify whether the schedule is conflict serializable or not.
 - iii. Present the equivalent serial schedules.

Passenger		
<u>PId</u>	Pname	Age
1001	Akshay Kumar	43
1005	Vijay	40
1050	Aishwarya	55
1080	Bhumika	35
1090	Anushka	25
1050	Pavi	22

Let us assume that in our parallel database system we have 3 processors P0, P1, P2 and n disks D0, D1, D2,... where we partition our data. Partition attribute is age and its range vector v [24, 40]. Apply various partitioning strategies proposed to manage the data distribution into multiple processors evenly and illustrate each strategy in detail.

4. Explain how the given query is decomposed and processed in a distributed database. Show how normalization, analysis, elimination of redundancy and rewriting is done with the given query.

Table Name: ASG

Attributes:

- ENO (Employee Number): unique identifier for each employee.
- PNO (Project Number): unique identifier for each project.
- DUR (Duration): duration of the project.
- RESP (Responsibility): responsibility of the employee in the project.

SELECT ENO FROM ASG WHERE (NOT (PNO = "Sales Expansion")
AND (PNO = "Sales Expansion" OR DUR = 12) AND NOT (DUR = 12)) OR
RESP = "Analyst";

Auto
by
some
Detection
classification

- 5/ a) A social media platform wants to develop a feature that can automatically generate captions for user-uploaded images. What technique can be used to develop this feature and ensure that the generated captions are accurate and relevant? [5]
- b) Develop a real-time object recognition system for self-driving cars that can detect and classify road objects, such as cars, pedestrians, and road signs, using automatic image analysis. What are the key challenges and limitations of such a system? [5]

6.

```
<courses>
<course id="1">
    <title>Introduction to Computer Science</title>
    <instructor>Professor Smith</instructor>
    <credits>3</credits>
</course>
<course id="2">
    <title>Data Structures and Algorithms</title>
    <instructor>Professor Johnson</instructor>
    <credits>4</credits>
</course>
</courses>
```

Specify the following views as XQuery:

- a) Display the titles of all courses with 3 credits. [2.5]
- b) Give the instructor of the course with the title "Data Structures and Algorithms". [2.5]
- c) Display the credits of the first course. [2.5]
- d) Insert a new supplier record. [2.5]

A company is migrating its on-premises database to a cloud-based database. What are the key considerations for selecting a suitable data storage system on the cloud? How would you design a data representation strategy for the cloud-based database? What are the potential challenges with data partitioning in the cloud?

8. a) Compare the key features of RBAC and MAC, and explain how RBAC provides [5] a more practical and flexible approach to access control.

- b) Consider the relational schema: [5]

Customers

Attribute Name	Data Type	Description
---	---	---
Cid	int	Customer ID
Cname	varchar(15)	Customer name
Address	varchar(30)	Address
Phone	varchar(10)	Phone number

Data New
Backend platform
other
port
conn

Orders

Attribute Name	Data Type	Description
---	---	---
Oid	int	Order ID
Cid	int	Customer ID
Order_date	date	Order date
Total	decimal(10, 2)	Total amount

User A creates the tables Customers and Orders and inserts values into them. Write SQL queries to grant the following permissions:

- User B can view only the Cname and Address attributes of the Customers table.
- User C can view only the Oid and Order_date attributes of the Orders table.

- 9.a) Design an EER diagram to capture the requirements of a food delivery platform that connects customers with local restaurants. The platform should allow customers to browse restaurants, view menus, place orders, and track their deliveries. Every restaurant has a unique ID, name, address, and cuisine type. Menus consist of dishes with unique IDs, names, descriptions, prices, and ingredients. Customers have unique IDs, names, email addresses, and delivery addresses. Each order has a unique ID, customer ID, restaurant ID, order date, and total amount. Each order can have multiple order items, which specify the dish ID, quantity, and price. The platform should also allow customers to leave reviews for restaurants and dishes.

State any assumptions you have that affects your design. Clearly indicate primary keys and the cardinalities with which an entity participates in a relationship with a pair of (minimum, maximum) value.

9.b)

OR

Suppose we have a database for an investment firm, consisting of the following attributes:

B – Broker,

O – Office of a broker

I – Investor

S – Stock

Q – Quantity of stock owned by an investor

D – Dividend paid by a stock.

Hence, the overall schema is $R = (B, O, I, S, Q, D)$.

Assume that the following functional dependencies are required to hold on this database.

$FD = \{I \rightarrow B, IS \rightarrow Q, B \rightarrow O, S \rightarrow D\}$

List all the candidate keys for R and Normalize up to 3NF.

10.a)

Consider a block size of 512 bytes, a block pointer of size 8 bytes, and a search key value of 10 bytes for a primary index and 5 bytes for a clustering index on an ordered field. The size of each entry in the data file is 50 bytes with 200,000 records, and the record pointer is of size 6 bytes. The clustering index has 1,000 zip code values.

- i) Find the blocking factor for the data file, primary index, and clustering index. [4]
- ii) Find the number of blocks required for the data file, primary index, and clustering index. [4]
- iii) For the primary and clustering index, calculate the number of accesses needed. [2]

OR

10.b)

Schema:

Customer(CustId, Name, Address)

Order(OrdId, CustId, OrderDate)

OrderItem(OrdId, ItemId, Quantity)

SQL Query:

```
SELECT CustId, Name FROM Customer c, Order o, OrderItem oi  
WHERE c.CustId = o.CustId AND o.OrdId = oi.OrdId AND  
OrderDate >'01-Jan-2020' AND Quantity > 10
```

Construct an Initial Query Tree and Optimized Query Tree using Heuristic Query Optimization Techniques.

↔↔↔ BH/D/TY ↔↔↔



V/D/TY

Reg. No: 21MCE0006



VIT

Vellore Institute of Technology

Final Assessment Test - April 2025.

Course: MCSE602L - Machine Learning

Class NBR(s): 2112/2116

Slot: A2

Time: Three Hours

Max. Marks: 100

- KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
➤ DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer All Questions
(10 X 10 = 100 Marks)

1. Consider the below mentioned data set:

Age	Gender	Cough	Fever	Headache	Disease Prediction
Young	Male	Yes	Yes	No	Yes
Young	Female	Yes	Yes	Yes	Yes
Middle-aged	Male	No	Yes	Yes	No
Senior	Female	Yes	No	Yes	No
Senior	Male	Yes	Yes	No	Yes
Middle-aged	Female	No	No	No	No
Young	Male	No	No	Yes	No
Senior	Female	Yes	Yes	Yes	Yes

Apply Candidate Elimination to find the version space (set of hypotheses) that explains the relationship between symptoms (cough, fever, and headache) and disease prediction based on the patient's age, gender, and symptoms.

2.

Explain how clean and well-prepared data can significantly impact the performance of machine learning models. Additionally, discuss the potential risks associated with using raw, unprocessed data in machine learning, supported by a relevant real-world case study.

1+	21.28	9+	18.06	6.25	14.06
1+	25+	4+	22.56	2.25+	7.56
4+	16+	1+	4.56	1.56	0.25

Page 1 of 4



Shot on moto g31

7 May 2025, 8:12 pm

3.

Match the following

Col A	Col B
i. High Bias	a. The balance that needs to be maintained between model complexity and simplicity to minimize error
ii. High Variance	b. Evaluates features based on statistical techniques or metrics, independently of the machine learning model
iii. Overfitting	c. The model is too simple, and assumptions or approximations may not capture the data well.
iv. Underfitting	d. The model is very sensitive to the training data and may perform poorly on unseen data.
v. Bias-Variance Tradeoff	e. The model fits the noise and random fluctuations in the training data, leading to poor generalization
vi. Filter Method	f. The model fails to capture the underlying patterns in the data, leading to poor performance.
vii. Wrapper Method	g. Feature selection technique where features are evaluated based on their performance with a specific machine learning model
viii. Backward Elimination	h. A feature selection method where features are removed one at a time, starting with the least useful feature, based on model performance
ix. Forward Selection	i. Starts with an empty set of features and adds the most useful feature at each step to improve model performance
x. Recursive Feature Elimination (RFE)	j. Starts with an empty set of features and adds the most useful feature at each step to improve model performance

0.25 42.25 30.25

12.25 + 26 20.25

6.25 +

VB, LV
VB, HV over
HB, LV over
HB, HV over

4. A dataset contains 1,000 samples, and a classification model is trained with a subset of features. After training by X model the following confusion matrix is obtained. Calculate the following metrics based on the confusion matrix: Precision, Recall (Sensitivity), Specificity, F1-score, and Accuracy.

	Predicted Positive (P)	Predicted Negative (N)
Actual Positive (P)	250	50
Actual Negative (N)	30	670

- 5.a) Design a neural network with the following architecture: 1 input layer containing 2 neurons (Input 1 and Input 2), 1 hidden layer with 2 neurons, and 1 output layer with a single neuron. Consider initial random weight as 0.8, bias values as 0.6, a leaning rate of 0.1 and sigmoid activation function using the above values to solve a binary classification problem based on and logical function.

Input 1	Input 2	Output (y)
0	0	0
0	1	0
1	0	0
1	1	1

OR

(No. 425) f

- 5.b) A single layer perceptron (SLP) is a feed-forward network based on a threshold transfer function. SLP is a simplest type of artificial neural networks and can only classify linearly separable cases with a binary target (1, 0). Classify the following data using perceptron from neural networks.

Width	Height	Classification
0.452	2.322	0
3.121	3.641	1
2.230	0.120	0
1.110	1.101	0

6. Imagine you have a dataset consisting of images of 5 different types of fruits, with each image corresponding to a specific fruit variety. Apply Random forest classifier to classify these images in to their respective categories.

TP FN
FP TN

7.a) Find the three clusters after 2 epoch for the following eight examples using the k-means algorithm and Euclidean distance $A_1=(2,10)$, $A_2=(2,5)$, $A_3=(8,4)$, $A_4=(5,8)$, $A_5=(7,5)$, $A_6=(6,4)$, $A_7=(1,2)$, $A_8=(4,9)$. Suppose that the initial seeds (centers of each cluster) are A_1 , A_4 and A_7 .

OR

7.b) Consider the following distance matrix. Apply the Single Link - Hierarchical Clustering technique.

	P1	P2	P3	P4	P5	P6
P1	0.00	0.24	0.22	0.37	0.34	0.23
P2	0.24	0.00	0.15	0.20	0.14	0.25
P3	0.22	0.15	0.00	0.15	0.28	0.11
P4	0.37	0.20	0.15	0.00	0.29	0.22
P5	0.34	0.14	0.28	0.29	0.00	0.39
P6	0.23	0.25	0.11	0.22	0.39	0.00

And also show the results by drawing a dendrogram.

8. Consider a 2D dataset, use the BIRCH algorithm to cluster the data. Assume the following, CF Tree is initialized with a branching factor of 2, Threshold (T) for the CF Tree is set to 3, Process the points in the dataset in sequence (A to F). At each step, construct the CF Tree and classify the points into clusters.

Point	X Coordinate	Y Coordinate
A	1.0	2.0
B	1.5	1.8
C	3.0	3.5
D	8.0	8.0
E	8.5	8.5
F	9.0	9.0

9. Explain Federated Learning and its importance in privacy-preserving machine learning. How does federated learning work in the context of online classification tasks, and what are its main challenges in terms of model convergence, communication efficiency, and data privacy?

10. As an AI engineer, how do you perceive machine learning playing a crucial role in various industries? What future directions would you propose to advance AI concepts in order to promote sustainable global development, and what are the justifications behind your suggestions? Explain it in detail.

◆◆◆◆ V/D/TY ◆◆◆◆

25 p.4

9 p.16

1 p.6 4 9 p.4

4 p.9

16 p.36

1 p.9

16 p.36

W/D/TY



VIT

Vallibel Institute of Technology

Final Assessment Test - April 2025

Reg. No: 24MC5006

Course: MCSE612L - Cyber Security

Class NBR(s): 3999

Slot: B1+TB1

Time: Three Hours

Max. Marks: 10

- > KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- > DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer ALL Questions
(10 X 10 = 100 Marks)

1. Discuss the CIA Triad (Confidentiality, Integrity, and Availability) in cybersecurity. Provide examples of how each principle can be compromised in an organizational network (e.g., data breaches, insider threats, malware, and DDoS attacks). Analyze the impact of these threats and propose effective mitigation strategies.
2. Explain how User Behavior Analytics (UBA) enhances traditional penetration testing by improving anomaly detection and Advanced Persistent Threat (APT) identification. Propose a hybrid penetration testing approach integrating UBA to enable real-time threat detection and response.
3. Apply the concept of the software vulnerability lifecycle by explaining its phases. Develop a security approach that integrates secure coding practices, threat intelligence, and automated patching to mitigate vulnerabilities effectively.
4. Propose a comprehensive strategy for securing remote access in a hybrid workforce environment. Compare how Zero Trust Architecture (ZTA) mitigates risks more effectively than traditional VPN solutions, focusing on endpoint security and cloud security.
5. Evaluate the effectiveness of Zenmap in cybersecurity operations by analyzing its strengths and limitations. Discuss its role in vulnerability assessment and network reconnaissance. Provide a Python-based example demonstrating how Nmap can be used to scan a network for open ports and potential vulnerabilities.
6. Analyze the strengths and limitations of Red Team and Blue Team approaches in modern cybersecurity defense. Design a hybrid "Purple Team" framework that incorporates continuous security testing and real-time automation to improve cyber resilience.
7. Describe the key challenges in addressing cybercrime and explain how legal, technical, and societal factors affect cybersecurity efforts. How do cybercrime tactics influence cybersecurity, and what measures can help protect against them?
8. Explain the major cybersecurity challenges organizations encounter in the digital era and discuss how technological advancements influence the evolution of cyber threats.

Page 1





9.a) Classify different types of cyber attackers based on their motives. Examine how cyber attacks can be systematically classified based on their attack surface, exploitation methods, and overall consequences.

OR

9.b) Analyze the limitations of legacy cybersecurity systems and explain how modern security frameworks address these challenges. Support your discussion with a real-world example of behavior-based malware analysis.

10.a) What is access control in cybersecurity, and how do various access control models enhance data protection? Illustrate Role-Based Access Control (RBAC) with a Python script for managing user permissions.

OR

10.b) Explain how modern web browsers can be configured to defend against phishing attacks, malicious extensions, and data leaks. Provide an example of a Python-based approach for blocking suspicious URLs.

↔↔↔ W/D/TY ↔↔↔

Page 2 of 2

