U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

June 2025 Semester End Main Examinations

Programme: B.E. **Semester: VI**

Branch: Computer Science & Engineering **Duration: 3 hrs.**

Course Code: 23CS6PCMAL / 22CS6PCMAL **Max Marks: 100**

Course: Machine Learning

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

		UNIT - I	*CO*	*PO*	Marks
1	a)	List and explain the different types of Machine Learning (ML) algorithms with an example for each.	*CO1*	*PO2*	**9**
	b)	Identify the type of ML algorithm to be used for the following scenario: i) Classification of email as spam or not ii) Photo-hosting services, such as Google Photos iii) Allowing a robot to walk in various unknown terrains iv) You have a lot of data about your blog's visitors v) Algorithm relies on a similarity measure to make predictions	*CO1*	*PO2*	**5**
	c)	List and explain various key purposes of Machine learning.			**6**
		OR			
2	a)	Analyze the following learning problem and write task T, Performance measure P and Training experience E. i. Face Recognition Problem ii. Handwriting Recognition	*CO1*	*PO2*	**5**
	b)	Mention the advantages of writing a function for data preparation than doing manual data preparation.	*CO1*	*PO1*	**5**
	c)	Write brief note on: i) Preparing a dataset to perform a Machine learning task ii) Online learning	*CO1*	*PO1*	**10**

			UNIT - II			
	3	a)	Construct a Regression tree using the following data which consists of 10 data instances and three attributes. Result is the target attribute. **S.No. / Assessment / Assignment / Project / Result (%)** 1. Good Yes Yes 95 2. Average Yes No 70 3. Good No Yes 75 4. Poor No No 45 5. Good Yes Yes 98 6. Average No Yes 80 7. Good No No 75 8. Poor Yes Yes 65 9. Average No No 58 10. Good Yes Yes 89	*CO1*	*PO5*	**10**
		b)	Consider the below dataset, apply multiple linear regression for the below table using matrix method. Target variable is **y**. **x1 / x2 / y** 60 22 140 62 25 155 67 24 159 70 20 179 71 15 192 72 14 200 75 14 212 78 11 215	*CO1*	*PO3*	**10**
			OR			
	4	a)	Apply the linear regression model for the following dataset, where the week and number of working hours spent by a student in a library are tabulated. Based on the dataset, predict the number of hours that will be spent by the student in the 7^{th} and 9^{th} week. **Week** 1 2 3 4 5 Hours Spent 12 18 22 28 35	*CO1*	*PO3*	**10**
		b)	Explain the working of the ID3 algorithm for decision tree construction. Discuss how entropy and information gain are used in attribute selection. Illustrate the algorithm with a step-by-step example using a small dataset.	*CO1*	*PO1*	**10**

<table>
<tr><td></td><td></td><td></td><td colspan="3">UNIT - III</td><td></td><td></td><td></td></tr>
<tr><td></td><td>5</td><td>a)</td><td>For the following confusion matrix, calculate Accuracy, Precision, Recall, F1 score for each class.
<table>
<tr><th rowspan="2">True class</th><th colspan="3">Predicted Class</th></tr>
<tr><th>1</th><th>2</th><th>3</th></tr>
<tr><td>1</td><td>8</td><td>2</td><td>0</td></tr>
<tr><td>2</td><td>1</td><td>9</td><td>0</td></tr>
<tr><td>3</td><td>1</td><td>2</td><td>7</td></tr>
</table></td><td>CO1</td><td>PO3</td><td>10</td></tr>
<tr><td></td><td></td><td>b)</td><td>Define cross-validation. What is its importance and when should Stratified K-Fold be preferred over K-Fold? Please provide a suitable example.</td><td>CO1</td><td>PO1</td><td>10</td></tr>
<tr><td></td><td></td><td></td><td colspan="3">OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>6</td><td>a)</td><td>Suppose 10000 patients get tested for flu. Out of them, 9000 are actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For the healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the precision and recall for the data.</td><td>CO2</td><td>PO5</td><td>8</td></tr>
<tr><td></td><td></td><td>b)</td><td>Write short notes on:
i. t-test
ii. McNemar’s test</td><td>CO1</td><td>PO1</td><td>8</td></tr>
<tr><td></td><td></td><td>c)</td><td>Differentiate between Area-Under-Curve and Receiver-Operating-Characteristic.</td><td>CO2</td><td>PO2</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td colspan="3">UNIT - IV</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>Consider the features X1and X2 with following datapoints. Compute the principal component using PCA algorithm.
<table>
<tr><td>Features</td><td>E1</td><td>E2</td><td>E2</td><td>E3</td></tr>
<tr><td>X1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td>X2</td><td>4</td><td>6</td><td>8</td><td>10</td></tr>
</table></td><td>CO1</td><td>PO3</td><td>12</td></tr>
<tr><td></td><td></td><td>b)</td><td>Explain the Stacking ensembled method with a neat diagram.</td><td>CO1</td><td>PO1</td><td>8</td></tr>
<tr><td></td><td></td><td></td><td colspan="3">OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>8</td><td>a)</td><td>Explain bagging and boosting ensemble methods with a neat diagram and example.</td><td>CO2</td><td>PO1</td><td>10</td></tr>
<tr><td></td><td></td><td>b)</td><td>Use K-Means clustering to cluster the following data into two groups. Assume cluster centroids are m1=2 and m2=4. The distance function used is Euclidean distance {2, 4, 10, 12, 3, 20, 30, 11, 25}. Run for four iterations.</td><td>CO1</td><td>PO3</td><td>10</td></tr>
<tr><td></td><td></td><td></td><td colspan="3">UNIT - V</td><td></td><td></td><td></td></tr>
<tr><td></td><td>9</td><td>a)</td><td>Explain Markov Decision Process with a neat diagram and give example.</td><td>CO1</td><td>PO1</td><td>10</td></tr>
<tr><td></td><td></td><td>b)</td><td>Explain Active RL and Generalization in Reinforcement Learning (RL).</td><td>CO1</td><td>PO1</td><td>10</td></tr>
<tr><td></td><td></td><td></td><td colspan="3">OR</td><td></td><td></td><td></td></tr>
</table>

<table>
<tr>
<td></td>
<td>10</td>
<td>a)</td>
<td>Consider the scenario:

Agent is in state S3
Takes action a2
Lands in state S4
Gets reward R=5
Learning rate α=0.6
Discount factor γ=0.9

Q-table before update
<table>
<tr><th></th><th>a1</th><th>a2</th><th>a3</th><th>a4</th></tr>
<tr><td>S3</td><td>3</td><td>6</td><td>1</td><td>2</td></tr>
<tr><td>S4</td><td>0</td><td>4</td><td>5</td><td>8</td></tr>
</table>
Use the Q-learning update rule (equation) to update the Q(S3, a2). Show solving steps completely and clearly.</td>
<td>CO2</td>
<td>PO3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b)</td>
<td>Differentiate between Model based and model free reinforcement with an example.</td>
<td>CO2</td>
<td>PO2</td>
<td>10</td>
</tr>
</table>

****