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: Institutional Elective **Duration: 3 hrs.**

Course Code: 24AM6OEMLG **Max Marks: 100**

Course: Introduction to 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)		Explain any four applications of machine learning.	CO1	PO1	**6**
	b)		Identify and justify the suitable machine learning techniques for the following scenarios: i. Medical diagnosis where patient symptoms and test results are labeled with the correct disease. ii. Anomaly detection in network traffic to identify unusual behavior or potential security threats. iii. Training a robot to navigate a maze. iv. Handwritten digit recognition from images labeled with the correct digit.	CO2	PO2	**8**
	c)		Differentiate between supervised and unsupervised machine learning techniques.	CO1	PO1	**6**
			OR			
2	a)		Explain perspectives and issues of machine learning.	CO1	PO1	**6**
	b)		Analyze and elaborate the following in the design of checkers game learning problem: i. Choosing the training experience. ii. Choosing the target function. iii. Choosing a Representation for the Target Function.	CO2	PO2	**8**
	c)		Formulate following problems as well-posed learning problems by clearly defining the learning task, performance measure and training experience. i. Identify spam emails ii. A self-driving car	CO1	PO1	**6**
			UNIT - II			
3	a)		Explain the following terms: i. Consistent hypothesis ii. General Boundary iii. Specific Boundary iv. Version Space	CO1	PO1	**8**

	b)	i. Write Find-S algorithm. ii. Use Find-S algorithm to determine most specific hypotheses for the given data. i.	CO2	PO2	**12**

Origin	Manufacturer	Color	Decade	Type	Example type
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chryster	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive
Japan	Toyota	Green	1980	Economy	Positive
Japan	Honda	Red	1980	Economy	Negative

		OR			
4	a)	Assume a problem has totally six features out of which one feature has three possible values and the remaining five features have two possible values. Compute the following: i. Distinct instances ii. Syntactically distinct hypothesis iii. Semantically distinct hypothesis	CO1	PO1	**8**
	b)	Write Candidate Elimination Algorithm and identify the Version Space for the given dataset (Target concept=Enjoy sports)	CO2	PO2	**12**

Sky	AirTemp	Humidity	Wind	Water	Forecast	Enjoy sports
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes

		UNIT - III			
5	a)	Write K-Nearest Neighbor algorithm for approximating discrete valued functions.	CO1	PO1	**5**
	b)	For the given positively labeled data points (3,1) (3,-1) (6,1) (6,-1) and negatively labeled data points (1,0) (0,1) (0,-1) (-1,0), apply Support Vector Machine (SVM) algorithm and obtain the hyperplane that classifies the given data points.	CO2	PO2	**10**
	c)	Distinguish between linear and logistic regression models.	CO2	PO1	**5**
		OR			
6	a)	Outline on instance-based learning method. Write any two real applications where instance-based learning method used.	CO1	PO1	**5**

<table>
<tr><td></td><td></td><td>b)</td><td>Apply K-Nearest Neighbor algorithm for the given dataset to classify new instance of data:
Credit score (CS)=785 and Age=35, consider K=3.
<table>
<tr><td>Example</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>CS</td><td>742</td><td>655</td><td>484</td><td>366</td><td>824</td><td>201</td><td>421</td></tr>
<tr><td>Age</td><td>28</td><td>24</td><td>50</td><td>60</td><td>40</td><td>25</td><td>28</td></tr>
<tr><td>Eligible</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>10</td></tr>
<tr><td></td><td></td><td>c)</td><td>Differentiate between linear and non-linear support vector machine.</td><td>CO2</td><td>PO2</td><td>5</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - IV</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>Illustrate Reduced error pruning and Rule Post pruning with suitable examples.</td><td>CO1</td><td>PO1</td><td>8</td></tr>
<tr><td></td><td></td><td>b)</td><td>I. Write ID3 algorithm.
II. Construct the decision trees to represent the following boolean functions:
i. A ∧ ¬B
ii. A V [B ∧ C]</td><td>CO2</td><td>PO2</td><td>12</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>8</td><td>a)</td><td>With suitable justification, provide the characteristics of the problems where decision tree learning is generally best suited.</td><td>CO1</td><td>PO1</td><td>8</td></tr>
<tr><td></td><td></td><td>b)</td><td>Apply ID3 algorithm to construct the decision tree for the following data:
<table>
<tr><th>Day</th><th>A1</th><th>A2</th><th>A3</th><th>Classification</th></tr>
<tr><td>1</td><td>True</td><td>Hot</td><td>High</td><td>No</td></tr>
<tr><td>2</td><td>True</td><td>Hot</td><td>High</td><td>No</td></tr>
<tr><td>3</td><td>False</td><td>Hot</td><td>High</td><td>Yes</td></tr>
<tr><td>4</td><td>False</td><td>Cool</td><td>Normal</td><td>Yes</td></tr>
<tr><td>5</td><td>False</td><td>Cool</td><td>Normal</td><td>Yes</td></tr>
<tr><td>6</td><td>True</td><td>Cool</td><td>High</td><td>No</td></tr>
<tr><td>7</td><td>True</td><td>Hot</td><td>High</td><td>No</td></tr>
<tr><td>8</td><td>True</td><td>Hot</td><td>Normal</td><td>Yes</td></tr>
<tr><td>9</td><td>False</td><td>Cool</td><td>Normal</td><td>Yes</td></tr>
<tr><td>10</td><td>False</td><td>Cool</td><td>High</td><td>No</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>12</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - V</td><td></td><td></td><td></td></tr>
<tr><td></td><td>9</td><td>a)</td><td>Explain with neat diagram, agent interacting with environment during reinforcement learning.</td><td>CO1</td><td>PO1</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>For the transaction dataset,
<table>
<tr><th>Transaction ID</th><th>Items Purchased</th></tr>
<tr><td>1</td><td>Bread, Cheese, Egg, Juice</td></tr>
<tr><td>2</td><td>Bread, Cheese, Juice</td></tr>
<tr><td>3</td><td>Bread, Milk, Yogurt</td></tr>
<tr><td>4</td><td>Bread, Juice, Milk</td></tr>
<tr><td>5</td><td>Cheese, Juice, Milk</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>10</td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td>Assuming a minimum level of support min_sup = 50% and a minimum level of confidence min_conf = 75%:
i. Compute frequent itemsets using the Apriori algorithm. For each iteration show the candidate and acceptable frequent itemsets.
ii. Derive the strong association rules using itemsets of (i) with provided minimum support and confidence values.</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>c)</td><td>Illustrate the anomaly detection in unsupervised clustering.</td><td>CO1</td><td>PO1</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>10</td><td>a)</td><td>Explain the following terms:
i. Hierarchical clustering
ii. Single Link clustering
iii. Complete Link clustering</td><td>CO1</td><td>PO1</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>Consider the datapoints (1.0, 1.0), (5.0, 7.0) as initial centroids and classify the given data into 2 clusters by applying K means clustering technique for two iterations.
<table>
<tr><td>A</td><td>1.0</td><td>1.5</td><td>3.0</td><td>5.0</td><td>3.5</td><td>4.5</td><td>3.5</td></tr>
<tr><td>B</td><td>1.0</td><td>2.0</td><td>4.0</td><td>7.0</td><td>5.0</td><td>5.0</td><td>4.5</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>10</td></tr>
<tr><td></td><td></td><td>c)</td><td>Differentiate between reinforcement learning and unsupervised learning.</td><td>CO1</td><td>PO1</td><td>4</td></tr>
</table>

****