U.S.N.										

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

Autonomous Institute Affiliated to VTU

February 2025 Semester End Main Examinations

Programm: B.E.	**Semester: IV**
Branch: Artificial Intelligence and Machine Learning	**Duration: 3 hrs.**
Course Code: 24AM4PCIML	**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.

<table>
<tr><th></th><th></th><th>UNIT - I</th><th>CO</th><th>PO</th><th>Marks</th></tr>
<tr><td>1</td><td>a)</td><td>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</td><td>CO1</td><td>PO1</td><td>6</td></tr>
<tr><td></td><td>b)</td><td>Write Find S algorithm and derive most specific hypotheses for the given data by applying Find S algorithm.
<table>
<tr><th>Origin</th><th>Manufacturer</th><th>Color</th><th>Decade</th><th>Type</th><th>Example type</th></tr>
<tr><td>Japan</td><td>Honda</td><td>Blue</td><td>1980</td><td>Economy</td><td>Positive</td></tr>
<tr><td>Japan</td><td>Toyota</td><td>Green</td><td>1970</td><td>Sports</td><td>Negative</td></tr>
<tr><td>Japan</td><td>Toyota</td><td>Blue</td><td>1990</td><td>Economy</td><td>Positive</td></tr>
<tr><td>USA</td><td>Chryster</td><td>Red</td><td>1980</td><td>Economy</td><td>Negative</td></tr>
<tr><td>Japan</td><td>Honda</td><td>White</td><td>1980</td><td>Economy</td><td>Positive</td></tr>
<tr><td>Japan</td><td>Toyota</td><td>Green</td><td>1980</td><td>Economy</td><td>Positive</td></tr>
<tr><td>Japan</td><td>Honda</td><td>Red</td><td>1980</td><td>Economy</td><td>Negative</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>8</td></tr>
<tr><td></td><td>c)</td><td>Write and prove the Version space representation theorem.</td><td>CO1</td><td>PO1</td><td>6</td></tr>
<tr><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>a)</td><td>Differentiate between supervised and unsupervised Machine Learning.</td><td>CO1</td><td>PO1</td><td>6</td></tr>
<tr><td></td><td>b)</td><td>Describe the Candidate Elimination algorithm and apply the same for the given dataset to find generic and specific boundary.
<table>
<tr><th>Sky</th><th>Temp</th><th>Humid</th><th>Wind</th><th>Water</th><th>Forest</th><th>Output</th></tr>
<tr><td>Sunny</td><td>Warm</td><td>Normal</td><td>Strong</td><td>Warm</td><td>Same</td><td>Yes</td></tr>
<tr><td>Sunny</td><td>Warm</td><td>High</td><td>Strong</td><td>Warm</td><td>Same</td><td>Yes</td></tr>
<tr><td>Rainy</td><td>Cold</td><td>High</td><td>Strong</td><td>Warm</td><td>Change</td><td>No</td></tr>
<tr><td>Sunny</td><td>Warm</td><td>High</td><td>Strong</td><td>Cool</td><td>Change</td><td>Yes</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>8</td></tr>
<tr><td></td><td>c)</td><td>Illustrate the concept of modeling inductive systems by equivalent deductive systems.</td><td>CO2</td><td>PO2</td><td>6</td></tr>
</table>

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 - II			
	3	a)	Apply ID3 algorithm to construct a decision tree on the given COVID-19 dataset. Justify the importance of Occum's razor law in decision trees.	*CO3*	*PO3*	**10**

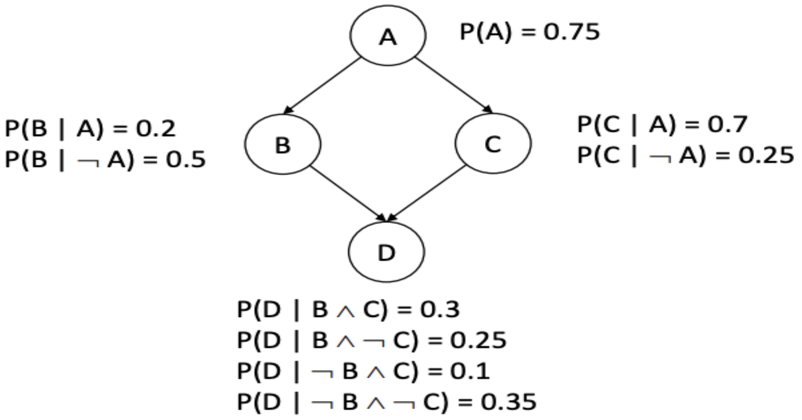
ID	**Fever**	**Cough**	**Breathing Issues**	**Infected**
1	no	no	no	no
2	yes	yes	yes	yes
3	yes	yes	no	no
4	yes	no	yes	yes
5	yes	yes	yes	yes
6	no	yes	no	no
7	yes	no	yes	yes
8	yes	no	yes	yes
9	no	yes	yes	yes
10	yes	yes	no	yes
11	no	yes	no	no
12	no	yes	yes	yes
13	no	yes	yes	no
14	yes	yes	no	no

		b)	Apply K- Nearest Neighbor algorithm for the dataset and classify the instance: <Brightness = 80, Saturation = 40> where K = 5.	*CO3*	*PO3*	**6**

Brightness	40	50	60	10	70	60	25
Saturation	20	50	90	25	70	10	80
Class	Red	Blue	Blue	Red	Blue	Red	Blue

		c)	Distinguish between linear support vector machines and nonlinear support vector machines.	*CO1*	*PO1*	**4**
			OR			
	4	a)	Find best hyperplane which separates positive and negative class vectors and plot the same. Positive vectors: (1, 1), (-1, 1), (-1, -1), (1, -1). Negative vectors: (2, 0), (0, 2), (-2, 0), (0, -2).	*CO2*	*PO2*	**10**
		b)	Derive the gradient descent rule for a locally weighted regression to the target function.	*CO1*	*PO1*	**5**
		c)	Illustrate the general structure of decision tree with suitable example.	*CO1*	*PO1*	**5**
			UNIT - III			
	5	a)	Consider the dataset and find whether the object with attribute Confident = Yes, Sick = No will Fail or Pass, using Bayesian classification.	*CO3*	*PO3*	**8**

Confident	**Studied**	**Sick**	**Result**
Yes	No	No	Fail
Yes	No	Yes	Pass
No	Yes	Yes	Fail
No	Yes	No	Pass
Yes	Yes	Yes	Pass

		b)	Describe the working of Estimation Maximization (EM) algorithm with a diagram.	*CO1*	*PO1*	**4**
		c)	Provide an outline of any two ensemble learning techniques commonly used in machine learning, highlighting the fundamental principles and advantages of each.	*CO1*	*PO1*	**8**
			OR			
	6	a)	Consider the following Bayesian network. A, B, C, and D are Boolean random variables. If we know that A is true, what is the probability of D being true?  	*CO3*	*PO3*	**6**
		b)	Illustrate the idea of stacking in ensemble learning with a neat diagram. Explain how it combines different models to improve overall prediction accuracy.	*CO1*	*PO1*	**7**
		c)	Derive an equation for the concept of Least Squared Error Hypothesis.	*CO1*	*PO1*	**7**
			UNIT - IV			
	7	a)	Cluster the values (1, 3, 9, 11, 2, 19, 29, 10, 24) into two groups using k-means clustering with initial centroids M1 = 3 and M2 = 10.	*CO2*	*PO2*	**8**
		b)	Elaborate on the advantages and disadvantages of spectral clustering.	*CO1*	*PO1*	**6**
		c)	Outline the anomaly detection techniques in machine learning with suitable examples.	*CO1*	*PO1*	**6**
			OR			
	8	a)	Apply hierarchical agglomerative clustering using single linkage to form clusters on [17, 21, 24, 41, 26, 42]. Conclude your answer with a dendrogram.	*CO2*	*PO2*	**8**
		b)	Illustrate the Importance of elbow method during clustering the dataset using K-means algorithm.	*CO3*	*PO3*	**6**
		c)	Write the procedure of silhouette technique of clustering.	*CO1*	*PO1*	**6**
			UNIT - V			
	9	a)	Consider the two-dimensional data points: (2, 1), (3, 5), (4, 3), (5, 6), (6, 7), (7, 8). Compute the principal components using Principle Component Analysis (PCA) algorithm.	*CO2*	*PO2*	**12**

		b)	Illustrate the process of Sequential backward subset selection algorithm.	*CO3*	*PO3*	**08**
			OR			
	10	a)	Enumerate the steps involved in Linear Discriminant analysis with suitable equations.	*CO1*	*PO1*	**10**
		b)	Illustrate the following: i. Kernel Principle Component Analysis. ii. Locally Linear Embedding.	*CO3*	*PO3*	**10**

******