U.S.N.										

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

Autonomous Institute Affiliated to VTU

October 2024 Supplementary Examinations

Programme: B.E. | **Semester: VI**
Branch: Institutional Elective | **Duration: 3 hrs.**
Course Code: 24AM6OEIML | **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)	Define Machine Learning and illustrate its distinction from traditional programming using block diagrams.	*CO1*	*PO1*	**5**
	b)	Outline the pertinent challenges involved in Machine Learning.	*CO1*	*PO2*	**5**
	c)	Describe the factors influencing step selection in designing a learning system for checkers game, and sketch the final system design.	*CO2*	*PO2*	**10**
		UNIT - II			
2	a)	Illustrate how the concept learning can be used as the task of searching through a large space of hypothesis.	*CO1*	*PO2*	**5**
	b)	Apply Find-S algorithm to determine the maximally specific hypothesis for detecting malignant tumors from MRI scans with the provided readings.	*CO2*	*PO3*	**5**

Size	**Shape**	**Density**	**Tumor**
Small	Regular	Thin	Not-Malignant
Medium	Irregular	Thick	Malignant
Large	Irregular	Thick	Malignant

	c)	Apply Candidate Elimination Algorithm for the given loan approval data to derive the version space:	*CO2*	*PO3*	**10**

Ache	**Breathing-Issue**	**Cough**	**Fatigue**	**Fever**	**Covid**
Severe	Exists	Intense	Present	High	Yes
Severe	Exists	Intense	Absent	High	Yes
Mild	Exists	Light	Present	Low	No
Mild	Exists	Intense	Absent	High	Yes

		UNIT - III			
3	a)	Distinguish Linear and Logistic regression techniques.	*CO1*	*PO2*	**5**

		b)	Illustrate the working of Linear Support Vector Machine.	*CO2*	*PO2*	**5**
		c)	For the given data, Apply KNN with k=5 and classify the new instance (80,40) to a specific class.	*CO2*	*PO3*	**10**

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

OR

	4	a)	Justify the importance of choosing the right value for k in k-nearest neighbor with an example.	*CO1*	*PO2*	**4**
		b)	Researchers in a lab are studying a chemical reaction for a new compound. The data records the mass of the compound over time and it is as follows.	*CO2*	*PO3*	**8**

Time Unit (x)	5	7	12	16	20
Mass(y)	40	120	180	210	240

Apply the Simple Linear Regression and predict the mass of compound at time unit 10 and 15 respectively.

		c)	The dataset of promotion decisions for five employees is given below.	*CO2*	*PO3*	**8**

Years of experience	1	2	3	5	6	7
Yes (1) / No (0)	0	0	0	1	1	1

Apply Logistic Regression with optimizer **z = - 10 + 3 * years** to:

i. Compute the probability of promotion for an employee with 4 years of experience.

ii. Determine the minimum years of experience needed for an employee to have over 97% probability of promotion.

UNIT - IV

	5	a)	Illustrate the general structure of decision tree with suitable example.	*CO1*	*PO1*	**5**
		b)	Find the entropy for the given probabilities.	*CO2*	*PO2*	**5**

P1	P2	P3	P4
0.1	0.2	0.3	0.4

<table>
<tr><td></td><td></td><td>c)</td><td>Design a decision tree for the given dataset using the Iterative Dichotomiser (ID3) algorithm.
<table>
<tr><th>Instance</th><th>A1</th><th>A2</th><th>Classification</th></tr>
<tr><td>1</td><td>False</td><td>True</td><td>$</td></tr>
<tr><td>2</td><td>False</td><td>True</td><td>$</td></tr>
<tr><td>3</td><td>False</td><td>False</td><td>+</td></tr>
<tr><td>4</td><td>True</td><td>False</td><td>$</td></tr>
<tr><td>5</td><td>True</td><td>True</td><td>+</td></tr>
<tr><td>6</td><td>True</td><td>True</td><td>+</td></tr>
</table></td><td>CO2</td><td>PO3</td><td>10</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - V</td><td></td><td></td><td></td></tr>
<tr><td></td><td>6</td><td>a)</td><td>Illustrate the interactions of Reinforcement Learning components during the learning process.</td><td>CO1</td><td>PO1</td><td>4</td></tr>
<tr><td></td><td></td><td>b)</td><td>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.</td><td>CO2</td><td>PO3</td><td>8</td></tr>
<tr><td></td><td></td><td>c)</td><td>Create a dendrogram by merging clusters based on minimum distance for the data points (17, 21, 24, 41, 26, 42) and update the proximity matrix using hierarchical clustering.</td><td>CO2</td><td>PO3</td><td>8</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>Compare Supervised and Unsupervised Learning techniques.</td><td>CO1</td><td>PO2</td><td>4</td></tr>
<tr><td></td><td></td><td>b)</td><td>Generate clusters for the numbers provided using the Density Based Scan (DBSCAN) algorithm, ensuring each cluster contains at least 4 elements and setting the epsilon (ε) parameter to 2.5.
<table>
<tr><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td></tr>
<tr><td>(3,7)</td><td>(4,6)</td><td>(5,5)</td><td>(6,4)</td><td>(7,3)</td><td>(6,2)</td><td>(7,2)</td><td>(8,4)</td></tr>
</table></td><td>CO2</td><td>PO2</td><td>8</td></tr>
<tr><td></td><td></td><td>c)</td><td>Employ the Apriori algorithm to derive association rules from the provided dataset by assuming a minimum support of 40% and a minimum confidence of 70%.
1 -- {Orange, Mango, Apple, Banana}
2 -- {Grapes, Kiwi, Mango}
3 -- {Apple, Banana}
4 -- {Grapes, Mango}
5 -- {Apple, Banana, Orange}</td><td>CO2</td><td>PO3</td><td>8</td></tr>
</table>

******