U.S.N.										

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

Autonomous Institute Affiliated to VTU

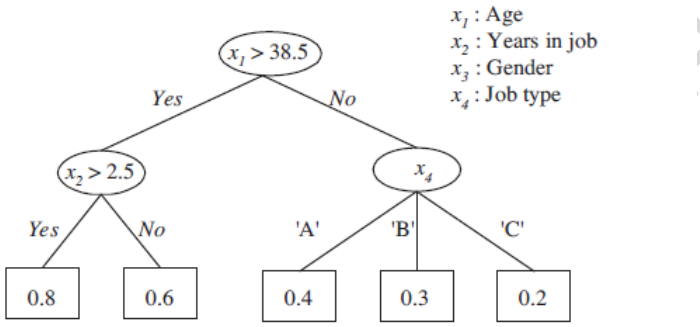
January / February 2025 Semester End Main Examinations

Programme: B.E. **Semester: V**
Branch: Information Science and Engineering **Duration: 3 hrs.**
Course Code: 23IS5PCMLG **Max Marks: 100**
Course: Machine Learning

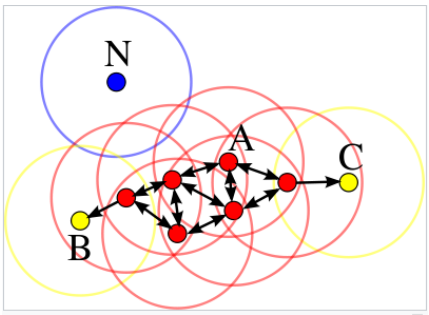
Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

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)	Bring out the differences between Supervised and Unsupervised learning.	CO1	PO1	8
	b)	Calculate Accuracy, Precision, Recall, and F1 Score for the Confusion Matrix depicted in Table. 1b on Water Shortage in Schools. Also suggest which metric would not be a good evaluation parameter here and why? **Table. 1b: Confusion Matrix** **Water Shortage in School** / **Actual: YES (Water Shortage)** / **Actual: NO (No Water Shortage)** **Prediction: YES**: 75, 5 **Prediction: NO**: 5, 15	CO1	PO1	8
	c)	An optimized model will be sensitive to the patterns in the data, but at the same time will be able to generalize to new data. In this, **both the bias and variance should be low.** Illustrate the bias-variance trade-off to achieve an effective balance between bias and variance for optimal performance with the help of a diagram.	CO1	PO1	4
		OR			
2	a)	What are the key differences between instance-based learning and model-based learning in terms of how they generalize from data and handle unseen examples?	CO1	PO1	6
	b)	Summarize the various gradient descent approaches and provide the solution to the following scenario: Suppose you are training a machine learning model on a dataset with 1 million samples. You notice that using the entire dataset to compute the gradient during each iteration is too slow, but computing gradients for each individual sample introduces high variability in the updates. What gradient descent approach would you choose, and why?	CO1	PO1	10

		c)	Given the below scenario, determine whether it is an example of supervised or unsupervised learning. i) The task in vision recognition that aims to understand and categorize an image as a whole under a specific label., ii) The process of separating markets or customers into smaller, more manageable groups based on shared characteristics., iii) Financial organizations spot fraudulent transactions, iv) Recommendation systems- The method by which data is collected varies greatly depending on the type of products or services sold.	*CO1*	*PO1*	**4**
			UNIT - II			
	3	a)	Interpret the significance of extracting Rules from Trees. Additionally, write the Rules for the tree given in fig. 3a.  	*CO2*	*PO1*	**8**
		b)	How do Regression Trees differ from Classification Trees? Justify with an example.	*CO2*	*PO1*	**6**
		c)	Explain the following terms with relevant equations: i. Entropy ii. Gini Index iii. Misclassification Error	*CO2*	*PO1*	**6**
			OR			
	4	a)	Write the pseudocode for constructing the Classification Trees.	*CO2*	*PO1*	**8**
		b)	Compare and contrast Univariate and Multivariate Trees.	*CO2*	*PO1*	**6**
		c)	Elucidate the concept of Pruning and its types.	*CO2*	*PO1*	**6**
			UNIT - III			
	5	a)	Consider the dataset given in Table. 5a, where the "Weather" attribute has values such as Sunny, Overcast, and Rainy, and the "Play" attribute has values Yes or No. Apply the Naive Bayes classifier to predict the likelihood of playing (Play) given the weather condition is Overcast. **Table 5a. Weather Dataset for Play Decision**	*CO3*	*PO2*	**10**

Sl. No.	Weather	Play	Sl. No.	Weather	Play
1.	Sunny	NO	8.	Sunny	NO
2.	Sunny	NO	9.	Sunny	YES
3.	Overcast	YES	10.	Rainy	YES
4.	Rainy	YES	11.	Sunny	YES
5.	Rainy	YES	12.	Overcast	YES
6.	Rainy	NO	13.	Overcast	YES
7.	Overcast	YES	14.	Rainy	NO

		b)	Explain BRUTE-FORCE MAP LEARNING algorithm with necessary assumptions and equations.	*CO2*	*PO1*	**10**
			OR			
	6	a)	Why is Bayesian Learning important? List the features of Bayesian Learning methods. Also, mention the challenges faced in practical implementation.	*CO2*	*PO1*	**10**
		b)	Describe the Minimum Description Length principle with equations.	*CO2*	*PO1*	**10**
			UNIT - IV			
	7	a)	Apply K-means clustering for the following data points: P1(2, 10), P2(2, 5), P3(8, 4), P4(5, 8), P5(7, 5), P6(6, 4), P7(1, 2), P8(4, 9). Consider P1(2, 10), P4(5, 8) and P7(1, 2) as initial centroids. Use Euclidean distance and show the clusters formed up to two iterations.	*CO3*	*PO2*	**10**
		b)	Elaborate the working principle behind Ensemble methods. Also, explain voting classifier with suitable diagram & python code.	*CO2*	*PO1*	**10**
			OR			
	8	a)	Imagine you are working with a dataset containing information about various environmental factors affecting air quality, such as temperature, humidity, particulate matter levels, carbon dioxide levels, and wind speed. Your goal is to identify the most critical factors influencing air quality. i) Which dimensionality reduction technique could you apply to extract the most significant features from the dataset and reduce its complexity? ii) Design a python program using SKLearn library to achieve the goal.	*CO3*	*PO2*	**10**
		b)	i) Explain the steps involved in the DBSCAN algorithm. ii) Identify and illustrate the terms w.r.t DBSCAN for the following diagram. 	*CO2*	*PO1*	**10**
			UNIT - V			
	9	a)	Outline the problems or scenario where Neural networks can be appropriate.	*CO2*	*PO1*	**5**

		b)	i) Derive the Gradient Descent rule and explain the conditions under which it is applied. ii) Differentiate between Gradient Descent and Stochastic Gradient Descent.	***CO1***		**10**
		c)	Represent the sigmoid threshold unit with a neat diagram and explain how it works.	***CO2***	***PO1***	**5**
			OR			
	10	a)	Define Perceptron with a diagrammatic representation.	***CO2***	***PO1***	**5**
		b)	Illustrate the Back propagation algorithm.	***CO2***	***PO1***	**10**
		c)	How did the Biological Neural system inspire Artificial Neural Networks? Justify your answer by listing the similarities and differences.	***CO2***	***PO1***	**5**

****