Code: 20CS3602

III B.Tech - II Semester – Regular Examinations – JUNE 2023

MACHINE LEARNING
(COMPUTER SCIENCE & ENGINEERING)

Duration: 3 hours   Max. Marks: 70

Note: 1. This paper contains questions from 5 units of Syllabus. Each unit carries 14 marks and have an internal choice of Questions.
2. All parts of Question must be answered in one place.

BL – Blooms Level   CO – Course Outcome

			BL	CO	Max. Marks
		UNIT-I			
1		What is a decision tree? Explain ID3 algorithm to train a decision tree. Discuss with an example.	L2	CO1	14 M
		OR			
2	a)	How are machine learning techniques like supervised learning, unsupervised learning, suited to various application domains in the fields of finance, healthcare, retail, and entertainment?	L2	CO1	7 M
	b)	Briefly explain the steps involved in designing a learning system.	L2	CO1	7 M
		UNIT-II			
3		Define perceptron. Demonstrate perceptron training rule using AND operations.	L3	CO2	14 M
		OR			

4	Describe the Multilayer feed forward networks. Explain the feed forward error back propagation learning algorithm to train a multi-layer feed forward neural network.	L3	CO2	14 M

UNIT-III

5	a)	What is confusion matrix? Illustrate Accuracy, Precision, Recall and F1-score which can computed from the confusion matrix for a binary classification problem.	L3	CO2	7 M
	b)	Illustrate Ensembling technique Boosting.	L3	CO2	7 M

OR

6	Illustrate Naive bayes classifier. Discuss with an example.	L3	CO2	14 M

UNIT-IV

7	a)	Suppose you are given a dataset consisting of 5 data points, each with 2 features (x and y), and labeled as either "positive" or "negative". You want to use the k-nearest neighbor algorithm to classify a new data point with features (3, 4) as either "positive" or "negative", using Euclidean distance as the distance metric. The 5 data points and their corresponding labels are: Data point 1: (1, 2), label "negative" Data point 2: (3, 5), label "positive" Data point 3: (2, 1), label "negative" Data point 4: (4, 6), label "positive"	L4	CO4	7 M

		Data point 5: (5, 3), label "positive" What would be the predicted label for the new data point if k=1, k=3, and k=5?			
	b)	Apply the various steps involved in the support vector machine algorithm for finding the optimal hyperplane when the data is linearly separable.	L3	CO2	7 M

OR

8		Compare k-nearest neighbor (KNN) algorithm and the weighted k-nearest neighbor (WKNN) algorithm, and how do these differences impact the accuracy and efficiency of these algorithms in practice?	L4	CO4	14 M

UNIT-V

9	a)	Illustrate Extrinsic method used for measuring the clustering quality.	L3	CO3	7 M
	b)	Demonstrate the K-means clustering algorithm with an example.	L3	CO3	7 M

OR

10		How does the agglomerative hierarchical clustering algorithm work, and explain some common linkage criteria used to determine the distance between clusters at each step of the algorithm.	L3	CO3	14 M