

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**Question Paper Code: R5F02**

B.E./B.Tech. DEGREE EXAMINATION, NOV 2025

Fifth Semester

Computer Science and Design

R21UCD502 – FOUNDATIONS OF MACHINE LEARNING

(Regulations R2021)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 2 = 20 Marks)

1. Briefly explain the steps used in solving AI problems. CO1-U
2. Why do we need Artificial Intelligence? CO1-U
3. What is the need for probabilistic reasoning in AI? CO1-U
4. Briefly explain Continuous Distribution. CO1-U
5. You are given a small dataset containing information about fruits. Each fruit has a weight and a texture (smooth or bumpy). Based on these features, the goal is to classify the fruits as either an apple or an orange. Find out the label for a fruit with a weight of 160 gms and texture as bumpy. CO2-App

Weight	Texture	Label
150	Bumpy	Orange
170	Bumpy	Orange
140	Smooth	Apple
130	Smooth	Apple

6. Calculate sensitivity and specificity for the following confusion matrix. CO2-App

		Predicted	
		Dog	Not Dog
Actual	Dog	5	1
	Not Dog	1	3

7. What is the constraint of a simple perceptron? Why may it fail with a real-world data set? CO1-U

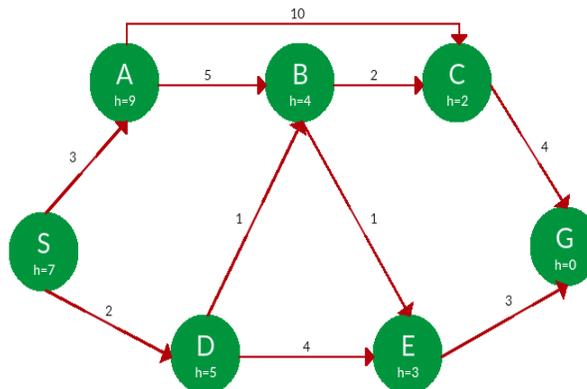
8. What is a linearly inseparable problem? What is the role of the hidden layer? CO1-U
9. Describe how the quality of clustering is measured in the k-means algorithm? CO1-U
10. What is a Dendrogram? Explain its use. CO1-U

PART – B (5 x 16= 80 Marks)

11. (a) You are working on a machine learning project where you need to pre-process the data before training various models. The dataset includes numerical, categorical, and text data, missing values, and outliers. Construct a complete data pre-processing pipeline, detailing each step and the techniques used to handle various data challenges. Explain how you ensure the pre-processing pipeline is scalable and maintainable as the dataset grows. CO3-Ana (16)

Or

- (b) Find various paths between S and G. Based on the time and space complexities, find the optimal path from S to G using the A\* search. Give a proper justification. CO3-Ana (16)



12. (a) In an exam, there were 20 multiple-choice questions. Each question had 4 possible options. A student knew the answer to 10 questions, but the other 10 questions were unknown to him and he chose answers randomly. If the score of the student X is equal to the total number of correct answers, then find out the PMF of X. What is  $P(X > 15)$ ? CO2-App (16)

Or

- (b) An antibiotic resistance test (random variable T) has 1% false positives (i.e. 1% of those not resistant to an antibiotic show positive results in the test) and 5% false negatives (i.e. 5% of those resistant to an antibiotic test negative). Let us assume that 2% of those tested are resistant to antibiotics. Determine the probability that somebody who tests positive is resistant (random variable D). CO2-App (16)

13. (a) Explain Naïve Bayes Classifier with an example. CO1-U (16)  
Or  
(b) Explain the Support Vector Machine (SVM) algorithm with an example. CO1-U (16)
14. (a) Explain Rosenblatt's perceptron model with an example. CO1-U (16)  
Or  
(b) Describe, in detail, the process of adjusting the interconnection weights in a multi-layer neural network. CO1-U (16)
15. (a) Explain the K-Medoids algorithm with an example. CO1-U (16)  
Or  
(b) How the distance between clusters is measured in hierarchical clustering? Explain the use of this measure in deciding when to stop the iteration. CO1-U (16)

