GUJARAT TECHNOLOGICAL UNIVERSITY
PGDDS – SEMESTER –I-EXAMINATION – SUMMER-2022

Subject Code: 1618003   **Date: 13-07-2022**

Subject Name: Fundamentals of Machine Learning

Time: 10:30 AM to 01:00 PM   **Total Marks: 70**

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**

			MARKS
Q.1	**(a)**	What is machine learning? What are key tasks of machine learning?	**03**
	(b)	Discuss candidate-elimination algorithm.	**04**
	(c)	Compare Supervised, Unsupervised and Semi supervised Learning.	**07**
Q.2	**(a)**	Define following terms in decision tree. 1) Leaf node 2) Entropy 3) Information gain	**03**
	(b)	Briefly explain the properties of Gini Impurity in decision tree.	**04**
	(c)	List down the advantages of the Decision Trees.	**07**

OR

	(c)	List out the disadvantages of the Decision Trees.	**07**
Q.3	**(a)**	If it takes one hour to train a Decision Tree on a training set containing 1 million instances, roughly how much time will it take to train another Decision Tree on a training set containing 10 million instances?	**03**
	(b)	Why do we require Pruning in Decision Trees? Explain.	**04**
	(c)	Describe the structure of an artificial neuron. How is it similar to a biological neuron? What are its main components?	**07**

OR

Q.3	**(a)**	What is threshold activation function in ANN?	**03**
	(b)	Design a multi-layer perceptron to implement A XOR B.	**04**
	(c)	Describe, in details, the process of adjusting the interconnection weights in a multi-layer neural network.	**07**
Q.4	**(a)**	Explain the concept of Prior, Posterior, and Likelihood with an example.	**03**
	(b)	Discuss application of ANN in face recognition.	**04**
	(c)	Explain, in details, the backpropagation algorithm. What are the limitations of this algorithm?	**07**

Q.4 **(a)** What is Naïve Bayes classifier? Why is it named so? **03**

(b) Briefly explain optimal Bayes classifier. **04**

(c) A drug test (random variable T) has 1% false positives (i.e., 1% of those not taking drugs show positive in the test), and 5% false negatives (i.e., 5% of those taking drugs test negative). Suppose that 2% of those tested are taking drugs. Determine the probability that somebody who tests positive is actually taking drugs (random variable D). **07**

Q.5 **(a)** Explain Minimum Description Length Principle. **03**

(b) Write two strength and weakness of Bayesian classifier. **04**

(c) Explain how Naïve Bayes classifier is used in text classification. **07**

<div align="center">OR</div>

Q.5 **(a)** Explain overfitting and underfitting. **03**

(b) Discuss a scenario of banking system where machine learning can be helpful. **04**

(c) What are Bayesian Belief networks? Where are they used? Can they solve all types of problems? **07**

<div align="center">**********</div>