Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE – SEMESTER- VII EXAMINATION-SUMMER 2023

Subject Code: 3170924 Date: 28/06/2023

Subject Name: AI and 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.**
4. **Simple and non-programmable scientific calculators are allowed.**

			Marks
Q.1	**(a)**	Clearly state and explain the importance of Artificial Intelligence in the healthcare and financial sector.	**03**
	(b)	State difference between supervised and unsupervised machine learning with examples.	**04**
	(c)	With the help of a schematic diagram, explain the concept of multi-layer perceptron neural network. Why do data scientists use back-propagation algorithm to train the neural network?	**07**
Q.2	**(a)**	What is semi-supervised learning? How is semi-supervised learning different from supervised and unsupervised learning?	**03**
	(b)	Explain the concept of decision trees for classification in machine learning setup. What is entropy and information gain w.r.t. a decision tree?	**04**
	(c)	With the help of a schematic diagram, explain the concept of Support vector machines (SVM). In case of a non-linear dataset, explain the role of kernel functions along with its advantages.	**07**
		OR	
	(c)	With the help of a schematic diagram, explain the concept of K-means clustering. How KNN is different from K-means clustering? Name the popular distance metrics used in KNN algorithm.	**07**
Q.3	**(a)**	Consider a scenario where a data scientist must choose between using Random Forest (a type of tree-based algorithm) and Support vector machine (SVM) for a classification task. Which algorithm should the data scientist use and why?	**03**
	(b)	While using 10000 numeric samples for a regression task, a data scientist faces a situation where the *mean squared error (MSE)* of the target variable in the validation phase is more than that in the training phase. Suggest possible remedies to address the situation.	**04**
	(c)	Explain the concept of Genetic algorithm with the help of a block diagram.	**07**
		OR	

Q.3 **(a)** Define the following terms w.r.t. fuzzy logic: **03**

(i) Membership function
(ii) Core
(iii) Crossover point

(b) What is Reinforcement learning? How it is different from Supervised learning? State examples where Reinforcement learning is used. **04**

(c) What is the importance of activation function in neural networks? Explain the concept of forward-phase and backward-phase in the weight-update algorithm used in neural networks with a suitable diagram. **07**

Q.4 **(a)** What is a confusion matrix? With the help of an example, explain the elements of a confusion matrix for a binary classification problem. **03**

(b) Explain the concept of Bias-Variance tradeoff in machine learning. How do we address the problem of over-fitting in a regression task? **04**

(c) Explain the concept of linear regression in machine learning. State key differences between univariate and multivariate linear regression. **07**

OR

Q.4 **(a)** What is the key difference between Linear regression and Logistic regression? State the importance of sigmoid function in machine learning. Draw the sigmoid curve. **03**

(b) Explain the concept of Fuzzy Logic Control with an example and a suitable block diagram. **04**

(c) What do you mean by Recommendation system(s)? State key differences between content-based filtering and collaborative filtering algorithm. Write five major applications of the recommendation system. **07**

Q.5 **(a)** Some patient features are expensive to collect (e.g., brain scans) whereas others are not (e.g., temperature). Therefore, we have decided to first ask our classification algorithm to predict whether a patient has a disease, and if the classifier is 80% confident that the patient has a disease, then we will do additional examinations to collect additional patient features In this case, which classification methods do you recommend: neural networks, decision tree, or naive Bayes? Justify your answer in one or two sentences. **03**

(b) What is fuzzification and defuzzification? Explain the meaning of universe of discourse. **04**

(c) Explain the concept of Cross-validation (CV) in machine learning. State different CV techniques used in a regression task. **07**

OR

Q.5 **(a)** Define the following terms: **03**

(i) Accuracy in ML
(ii) Attribute in ML
(iii) Decision boundary for SVM

(b) What is feature engineering in machine learning? How do you apply it in the process of modelling? **04**

(c) State key differences between Machine Learning and Deep Learning. Is Artificial neural network (ANN) a machine learning algorithm or deep learning algorithm? Justify your answer. **07**