Printed Page:-04 **Subject Code**:- **ACSML0401N**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: IV - THEORY EXAMINATION (2024 - 2025)

Subject: Machine Learning

Time: 3 Hours **Max. Marks: 100**

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*

1. *This Question paper comprises of* ***three Sections -A, B, & C.*** *It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.*

2. *Maximum marks for each question are indicated on right -hand side of each question.*

3. *Illustrate your answers with neat sketches wherever necessary.*

4. *Assume suitable data if necessary.*

5. *Preferably, write the answers in sequential order.*

6. *No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.*

SECTION-A 20

1. Attempt all parts:-

1-a. What is the primary goal of concept learning in machine learning?(CO1,K1) 1

(a) To generalize from the training data to unseen instances

(b) To memorize the training data

(c) To maximize model complexity

(d) To minimize overfitting

1-b. In machine learning, what does "overfitting" refer to?(CO1,K1) 1

(a) Creating a model that perfectly fits the training data but generalizes poorly to new data.

(b) Creating a model that underestimates the data distribution.

(c) Creating a model that has low variance.

(d) Creating a model that is too simple to capture the data patterns.

1-c. The Apriori algorithm is commonly used for: (CO2,K1) 1

(a) Classification tasks

(b) Regression analysis

(c) Market basket analysis

(d) Neural network training

1-d. A single-layer neural network is often referred to as a:(CO2,K1) 1

(a) Perceptron

(b) Multilayer Perceptron

.

(c) Deep neural network

(d) Support vector machine

1-e. Partitional clustering algorithms, like K-means, aim to: (CO3,K1) 1

(a) Form hierarchical clusters

(b) Partition data into subsets without a predefined number of clusters

(c) Merge existing clusters into larger ones

(d) Discover clusters of varying shapes

1-f. In clustering, what does AGNES (Agglomerative Nesting of Data) primarily involve?(CO3,K1) 1

(a) Splitting data into disjoint subsets

(b) Dividing data into fixed-size clusters

(c) Merging data points into hierarchical clusters

(d) Calculating the variance of data points

1-g. XGBoost is an optimized implementation of:(CO4,K1) 1

(a) Random Forest

(b) K-Nearest Neighbors

(c) Gradient Boosting

(d) Naïve Bayes

1-h. In Bayesian Belief Networks, what are nodes used to represent?(CO4,K1) 1

(a) Probabilistic dependencies between variables

(b) Data points in a dataset

(c) Decision boundaries

(d) Clustering results

1-i. What is Reinforcement Learning?(CO5,K1) 1

(a) A type of supervised learning

(b) A machine learning approach for making predictions

(c) A type of learning where an agent interacts with an environment and learns by receiving rewards and penalties

(d) A form of unsupervised learning

1-j. What is the core idea behind Q Learning in Reinforcement Learning? (CO5,K1) 1

(a) It's a type of deep learning algorithm

(b) Learning to predict the expected cumulative reward for taking an action in a given state

(c) A supervised learning approach

(d) A clustering algorithm

2. Attempt all parts:-

2.a. Name three common approaches in machine learning.(CO1,K1) 2

2.b. Discuss the assumptions of linear regression models and their importance in model 2

.

	interpretation.(CO2,K2)	
2.c.	List different types of clustering algorithms.(CO3,K1)	2
2.d.	What distinguishes XGBoost from other gradient boosting implementations?(CO4,K2)	2
2.e.	Differentiate between model-free and model-based Reinforcement Learning (CO5,K4)	2

SECTION-B 30

3. Answer any five of the following:-

3-a.	Discuss the importance of sensitivity analysis in ensuring robust and interpretable models.(CO1,K2)	6
3-b.	Define underfitting and overfitting in machine learning and how they relate to model generalization? (CO1,K2)	6
3-c.	Compare and contrast the ID3 and C4.5 algorithms in terms of their decision-making criteria and their practical applications.(CO2,K2)	6
3-d.	Describe the core concept behind a support vector machine (SVM) .(CO2,K2)	6
3.e.	Describe the challenges in unsupervised learning algorithms.(CO3,K2)	6
3.f.	Explain the step-by-step process of constructing a Bayesian belief network (BBN) for a real-world problem.(CO4,K2)	6
3.g.	Define value of a policy.(CO5,K1)	6

SECTION-C 50

4. Answer any one of the following:-

4-a.	Explain the fundamental concepts of learning in machine learning, emphasizing differences between supervised and unsupervised learning.(CO1,K2)	10
4-b.	Compare data science and machine learning, highlighting their similarities and differences.(CO1,K2)	10

5. Answer any one of the following:-

5-a.	Differentiate among superviesd, unsupervised and reinforcement learning algorithms.(CO2,K4)	10
5-b.	Explain the process by which decision tree algorithms, such as ID3, construct decision trees from training data.(CO2,K2)	10

6. Answer any one of the following:-

6-a.	What is K-Mode Clustering, and when should it be used?(CO3,K3)	10
6-b.	How do density-based clustering methods like DBSCAN work, and when are they helpful?(CO3,K2)	10

7. Answer any one of the following:-

7-a.	How does the Naïve Bayes Classifier address challenges in text classification like feature independence and data sparsity?(CO4,K2)	10
7-b.	Compare traditional decision trees and C5.0 boosting in terms of accuracy, interpretability, and computational demands. (CO4,K2)	10

.

8. Answer any one of the following:-

8-a.	What is the difference between model-free and model-based Reinforcement Learning with real examples?(CO5,K4)	10
8-b.	Explain the state transition function in Markov Decision Process.(CO5,K2)	10

.