

Printed Page:-

Subject Code:- ACSML0601

Roll. No:

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

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (20.. - 20....)

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. The main challenge in machine learning that is addressed by the concept of "bias-variance trade-off"? (CO1, K2) 1
- (a) Achieving perfect accuracy on the training data.
 - (b) Avoiding overfitting.
 - (c) Maximizing model complexity.
 - (d) Minimizing generalization error.
- 1-b. "One-hot encoding" used for in data preprocessing for machine learning? (CO1, K2) 1
- (a) Reducing the dimensionality of data
 - (b) Encoding categorical variables into a binary format
 - (c) Handling missing values
 - (d) Scaling numerical data
- 1-c. In linear regression, what is the primary goal? (CO2, K1) 1
- (a) Minimize the sum of squared errors
 - (b) Maximize the number of features
 - (c) Maximize the complexity of the model
 - (d) All
- 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. K-Nearest Neighbor (K-NN) is primarily used for which type of machine learning task? (CO3, K1) 1
- (a) Clustering
 - (b) Regression
 - (c) Classification
 - (d) Dimensionality reduction
- 1-f. The primary difference between hierarchical and partitional clustering methods? (CO3, K2) 1
- (a) Hierarchical clustering requires a predefined number of clusters, while partitional clustering does not.
 - (b) Hierarchical clustering forms a tree-like structure, while partitional clustering directly partitions the data.
 - (c) Hierarchical clustering is more suitable for mixed data types.
 - (d) Partitional clustering is non-hierarchical.
- 1-g. C5.0 boosting is a technique used in: (CO4, K1) 1
- (a) Decision tree modeling
 - (b) Linear regression
 - (c) Neural networks
 - (d) Clustering
- 1-h. Which ensemble method is known for its use of decision stumps (shallow trees) as base learners? (CO4, K1) 1
- (a) Bagging
 - (b) AdaBoost
 - (c) Random Forest
 - (d) Gradient Boosting
- 1-i. The primary application of Q Learning in Reinforcement Learning is? (CO5, K1) 1
- (a) Image recognition
 - (b) Game-playing agents that learn optimal strategies
 - (c) Natural language processing
 - (d) Stock market prediction
- 1-j. Which component in a Reinforcement Learning system is responsible for evaluating the quality of different actions in a given state? (CO5, K1) 1
- (a) Agent
 - (b) Environment
 - (c) Value function
 - (d) State transition function

2. Attempt all parts:-

- | | | |
|------|---|---|
| 2.a. | Define cross-validation in the context of machine learning. (CO1, K1) | 2 |
| 2.b. | Discuss the assumptions of linear regression models and their importance in model interpretation. (CO2, K2) | 2 |
| 2.c. | Describe the key principles of density-based clustering algorithms. (CO3, K2) | 2 |
| 2.d. | Explain the concept of boosting and its role in improving model performance. (CO4, K2) | 2 |
| 2.e. | Differentiate between model-free and model-based Reinforcement Learning. (CO5, K2) | 2 |

SECTION-B

30

3. Answer any five of the following:-

- | | | |
|------|--|---|
| 3-a. | Discuss data science and machine learning, highlighting their similarities and differences in goals and methodologies. (CO1, K2) | 6 |
| 3-b. | Point out the challenges related to bias in machine learning, including algorithmic and data bias, and strategies for mitigation, with real-world examples. (CO1, K4) | 6 |
| 3-c. | Differentiate between underfitting and overfitting. Give examples. (CO2, K2) | 6 |
| 3-d. | Demonstrate the stepwise process involved in construction of a decision tree with a suitable data set. (Note: create a dummy data set with at least two classes) (CO2, K3) | 6 |
| 3.f. | Illustrate the step-by-step process of constructing a Bayesian belief network (BBN) for a real-world problem. (CO4, K3) | 6 |
| 3.e. | Explain probability based Gaussian Mixture Model with example. (CO3, K4) | 6 |
| 3.g. | Compare and contrast model-free and model-based Reinforcement Learning approaches. (CO5, K4) | 6 |

SECTION-C

50

4. Answer any one of the following:-

- | | | |
|------|--|----|
| 4-a. | Explain "bias", "variance" and "bias-variance tradeoff" in machine learning. Discuss their effects on model performance and methods for balance. (CO1, K4) | 10 |
| 4-b. | Analyze the importance of model interpretability in machine learning and its impact on trust. (CO1, K4) | 10 |

5. Answer any one of the following:-

- | | | |
|------|--|----|
| 5-a. | Compare and contrast the ID3 and C4.5 decision tree algorithms, emphasizing their decision-making criteria and applications. (CO2, K4) | 10 |
| 5-b. | Explain the role of confusion matrices in classification problems. Differentiate precision, recall and F1 Measure. (CO2, K4) | 10 |

6. Answer any one of the following:-

- | | | |
|------|---|----|
| 6-a. | Compare AGNES and DIANA hierarchical clustering. (CO3, K4) | 10 |
| 6-b. | Differentiate between K-means and K-mode clustering in details. (CO3, K4) | 10 |

7. Answer any one of the following:-

- | | | |
|------|--|----|
| 7-a. | Compare traditional decision trees and C5.0 boosting in terms of accuracy, | 10 |
|------|--|----|

interpretability, and computational demands. (CO4, K4)

7-b. Point out the role of weak learners in Gradient Boosting Machines (GBM) and their contribution to creating strong learners. (CO4, K4) 10

8. Answer any one of the following:-

8-a. Explain the Q Learning function in simple terms and how it helps the learning process. (CO5, K2) 10

8-b. Discuss the various used components in reinforcement learning with the help of diagram. (CO5, K2) 10

REG:JAN_JUN-2025