.

Printed Page:-3 **Subject Code:- BCSML0401**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
B.Tech
SEM: IV - THEORY EXAMINATION (2025 - 2026)
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:-

(1.1) Regression predicts: (CO1, K1) 1

(a) Categories
(b) Continuous values
(c) Clusters
(d) Rewards

(1.2) Candidate Elimination algorithm is used for: (CO1, K1) 1

(a) Clustering
(b) Version space
(c) Regression
(d) RL

(1.3) CART stands for (CO2, K1) 1

(a) Classification and Regression Trees
(b) Clustering Algorithm Rule Technique
(c) Classification Algorithm Random Trees
(d) None

(1.4) Activation function in NN introduces (CO2, K1) 1

(a) Linearity
(b) Non-linearity
(c) Simplicity
(d) Error

(1.5) Codewords in ECOC represent (CO3, K1) 1

(a) Features
(b) Classes
(c) Samples
(d) Distances

(1.6) Random Forest builds trees using (CO3, K1) 1

(a) Full dataset
(b) Bootstrap samples
(c) Only test data
(d) Reduced dataset

(1.7) Bellman equation is used for (CO4, K1) 1

(a) Optimization
(b) Clustering
(c) Regression
(d) Sampling

(1.8) Markov property means (CO4, K1) 1

(a) Past matters
(b) Only present matters
(c) Future matters
(d) None

(1.9) Population in GA refers to (CO5, K1) 1

(a) Single solution
(b) Set of solutions
(c) Features
(d) Classes

(1.10) High mutation rate leads to (CO5, K1) 1

(a) Stability
(b) Random search
(c) Convergence
(d) No change

2. Attempt all parts:-

(2.1) Explain version space. (CO1, K2) 2

(2.2) Explain Multiple Linear Regression. (CO2, K2) 2

(2.3) Describe updating of weights in AdaBoost. (CO3, K2) 2

(2.4) Discuss dynamic programming in RL (CO4, K2) 2

(2.5) Describe roulette wheel selection. (CO5, K2) 2

SECTION-B 30

Answer any one of the following:-

(3.1) Explain the concept learning task with an example. (CO1, K2) 6

(3.2) Discuss PAC learning with ε and δ parameters. (CO1, K2) 6

Answer any one of the following:-

(3.3)	Differentiate between Linear Regression and Logistic Regression. (CO2, K4)	6
(3.4)	Examine the Apriori algorithm and the generation of association rules. (CO2, K4)	6

Answer any one of the following:-

(3.5)	Analyze the role of weak learners and weight updating in AdaBoost. (CO3, K4)	6
(3.6)	List advantages and limitations of stacking compared to voting methods. (CO3, K4)	6

Answer any one of the following:-

(3.7)	Analyze policy evaluation and policy improvement with examples. (CO4, K4)	6
(3.8)	Examine Monte Carlo methods and compare with temporal-difference learning. (CO4, K4)	6

Answer any one of the following:-

(3.9)	Analyze the relationship between hypothesis space and search efficiency in GA. (CO5, K4)	6
(3.10)	Investigate the impact of selection pressure on convergence speed and solution quality. (CO5, K4)	6

SECTION-C 50

4. Answer any one of the following:-

(4.1)	Analyze the Reinforcement Learning framework, including the agent, environment, reward, and policy. (CO1, K4)	10
(4.2)	Compare supervised, unsupervised, and reinforcement learning in terms of data, approach, and applications. (CO1, K4)	10

5. Answer any one of the following:-

(5.1)	Explain the concept of multicollinearity in multiple linear regression and methods to handle it. (CO2, K2)	10
(5.2)	Describe the complete process of model evaluation in regression using error metrics such as MSE, RMSE, and R^2. (CO2, K2)	10

6. Answer any one of the following:-

(6.1)	Examine the concept of bootstrap sampling and its role in bagging technique. (CO3, K4)	10
(6.2)	Analyze advantages of GMM over traditional clustering methods. (CO3, K4)	10

7. Answer any one of the following:-

(7.1)	Clarify the distinction in the exploration-exploitation trade-off and explain the Epsilon-Greedy strategy. (CO4, K4)	10
(7.2)	Compare Monte Carlo learning and Temporal Difference learning in terms of methodology, advantages, and limitations. (CO4, K4)	10

8. Answer any one of the following:-

(8.1)	Discuss integration of Genetic Algorithms with other machine learning techniques such as neural networks or clustering. Analyze benefits and challenges in solving complex real-world problems. (CO5, K4)	10
(8.2)	A recommendation system uses GA but produces inaccurate results. Analyze the role of fitness function and design an improved evaluation strategy. (CO5, K4)	10