Printed Page:- 05 **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 (2023 - 2024)

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. In machine learning, what are well-defined learning problems?(CO1) 1

(a) Problems that are easy to solve

(b) Problems with clear, structured data

(c) Problems that are mathematically solvable

(d) Problems where learning is feasible and meaningful

1-b. What is the primary goal when designing a learning system? (CO1) 1

(a) Minimizing data collection

(b) Minimizing generalization error

(c) Maximizing model complexity

(d) Maximizing predictive accuracy

1-c. Regression is a supervised learning task used for: (CO2) 1

(a) Categorizing data points into classes

(b) . Predicting numerical values

(c) Grouping data points into clusters

(d) None

1-d. In linear regression, what is the primary goal? (CO2) 1

(a) Minimize the sum of squared errors

(b) Maximize the number of features

(c) Maximize the complexity of the model

(d) All

1-e. In K-means clustering, how are clusters formed? (CO3) 1

(a) By iteratively selecting the nearest neighbors

(b) By randomly grouping data points

(c) By maximizing intra-cluster similarity and minimizing inter-cluster similarity

(d) By performing hierarchical merging of clusters

1-f. In hierarchical clustering, what does AGNES stand for? (CO3) 1

(a) Agglomerative Nesting of Data

(b) Adaptive Grouping of Entities

(c) Association of Geographical Nodes

(d) Aggregated Numerical Segmentation

1-g. In the Naïve Bayes Classifier, why is it called "naïve"? (CO4) 1

(a) It is easy to implement

(b) It assumes feature independence

(c) It is a simple model

(d) It only works with categorical data

1-h. What is the primary difference between boosting and bagging in ensemble methods? (CO4) 1

(a) Boosting uses randomization, while bagging does not

(b) Bagging combines weak learners sequentially, while boosting does not

(c) Bagging combines weak learners independently, while boosting combines them sequentially

(d) Boosting always outperforms bagging

1-i. Which learning model is fundamental in Reinforcement Learning and represents the interaction between an agent and its environment? (CO5) 1

(a) Markov Decision Process (MDP

(b) Logistic Regression

(c) k-Nearest Neighbors

(d) Principal Component Analysis (PCA

1-j. What is one of the key advantages of Reinforcement Learning? (CO5) 1

(a) It requires labeled training data

(b) It can handle only supervised learning tasks

(c) It allows agents to learn from interactions with their environment

(d) It relies solely on static rules and heuristics

2. Attempt all parts:-

2.a. Discuss the need of data preprocessing. (CO1) 2

2.b. In classification, what is the purpose of a decision boundary, and how is it related to the choice of algorithms? (CO2) 2

2.c. List different types of clustering algorithms. (CO3) 2

2.d. When would you use Bayesian Belief Networks in a real-world problem? (CO4) 2

2.e. Provide a real-world example of Reinforcement Learning in practice. (CO5) 2

SECTION B **30**

3. Answer any <u>five</u> of the following:-

3-a. Discuss the importance of sensitivity analysis in ensuring robust and interpretable models.(CO1) 6

3-b. Describe the concepts of bias and variance in machine learning and strategies for balancing them.(CO1) 6

3-c. Describe the core concept behind a support vector machine (SVM). (CO2) 6

3-d. Explain Apriori algorithm to find association rules from a transactional dataset.(CO2) 6

3.e. Discuss the role of clustering in unsupervised machine learning. (CO3) 6

3.f. Explore the strengths and weaknesses of the Naïve Bayes Classifier.(CO4) 6

3.g. Discuss the concept of the Q Learning function in Reinforcement Learning. (CO5) 6

SECTION C **50**

4. Answer any <u>one</u> of the following:-

4-a. Apply Find-S Algorithm to convert most specific hypothesis to most general hypothesis on given dataset. (CO1) 10

example	citations	size	inLibrary	price	editions	buy
1	some	small	no	affordable	many	no
2	many	big	no	expensive	one	yes
3	some	big	always	expensive	few	no
4	many	medium	no	expensive	many	yes
5	many	small	no	affordable	many	yes

4-b. Apply candidate elimination algorithm to find consistent hypothesis on given dataset. (CO1) 10

5. Answer any <u>one</u> of the following:-

5-a. Describe the concept of overfitting in regression models and strategies to address it.(CO2) 10

5-b. Apply ID3 on the given dataset to classify the data in either YES or NO. (CO2) 10

Instance	a1	a2	a3	Classification
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	Yes

6. Answer any <u>one</u> of the following:-

6-a. What is K-Nearest Neighbor (K-NN) clustering, and when is it useful? (CO3) 10

6-b. Apply the DBSCAN algorithm to the given data points and create the clusters with minPts = 4 and epsilon (ε) = 1.9 10

Dataset:

P1: (3,7), P2: (4, 6), P3: (5,5), P4: (6,4), P5: (7, 3), P6: (6, 2), P7: (7,2), P8: (8,4), P9: (3, 3), P10: (2,6) (CO3)

7. Answer any <u>one</u> of the following:-

7-a. What is the role of Bayesian Belief Networks in capturing complex probabilistic dependencies among variables? (CO4) 10

7-b. In what scenarios is the Bayes Optimal Classifier particularly useful for classification tasks? (CO4) 10

8. Answer any <u>one</u> of the following:-

8-a. What is Reinforcement Learning, and how is it used in real world applications ? (CO5) 10

8-b. Discuss examples of Q-Learning applications and the types of problems it typically solves.(CO5) 10