

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

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: V - 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, &amp; C. It consists of Multiple Choice Questions (MCQ's) &amp; 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. Which of the following algorithms is used to perform candidate elimination in concept learning?(CO1,K2) 1
- (a) K-means
- (b) Version Space
- (c) Decision Tree
- (d) Naive Bayes
- 1-b. Which of the following is a reinforcement learning application?(CO1,K1) 1
- (a) Topic modeling
- (b) Recommendation system
- (c) Pattern recognition
- (d) Image classification
- 1-c. Which decision tree algorithm uses a criterion called 'information gain' to decide the best split for a node?(CO2,K2) 1
- (a) CART
- (b) ID3
- (c) C4.5
- (d) Both b and c
- 1-d. ANNs are less motivated by(CO3,K2) 1
- (a) Chemical Neural networks

- (b) Physical Neural networks
  - (c) Biological neural systems
  - (d) Computational neural networks.
- 1-e. Which of the following is required by K-means clustering?(CO3,K2) 1
- (a) defined distance metric
  - (b) number of clusters
  - (c) initial guess as to cluster centroids
  - (d) all of the mentioned
- 1-f. In K-means clustering, what is the main objective of the algorithm?(CO3,K1) 1
- (a) Minimize the within-cluster variance
  - (b) Maximize the distance between different clusters
  - (c) Find the nearest neighbors for each data point
  - (d) Group data points based on categorical values
- 1-g. Which out of random forests and boosting trees are more prone to overfitting?(CO4,K1) 1
- (a) Random Forest
  - (b) Boosting Trees
  - (c) Both
  - (d) Can't say
- 1-h. Which of the following is true about weight of XGB leaf node.(CO4,K1) 1
- (a) "leaf weight" can be said as the model's predicted output associated with each leaf (exit) node
  - (b) leaf weight" can be said as the actual associated with each leaf (exit) node
  - (c) A or B depend on the situation
  - (d) None of these
- 1-i. \_\_\_\_\_Reinforcement is defined as when an event, occurs due to a particular behavior. (CO5,K1) 1
- (a) Negative
  - (b) Positive
  - (c) neutral
  - (d) None of these
- 1-j. Which of the following is the primary objective of Reinforcement Learning (RL)?(CO5,K1) 1
- (a) To minimize error by using labeled data
  - (b) To learn from feedback provided by the environment
  - (c) To classify data into predefined categories
  - (d) To perform supervised learning on sequential data

2. Attempt all parts:-

2.d.	State the Under and Overfitting.(CO1,K1)	2
2.a.	Describe the process of market basket analysis.(CO2,K2)	2
2.b.	State the primary objective of clustering in machine learning.(CO3,K2)	2
2.c.	Define Bayesian Learning in machine learning?(CO4,K2)	2
2.e.	Describe the agent, environment, and reward in Reinforcement Learning.(CO5,K2)	2

## **SECTION-B** 30

3. Answer any five of the following:-

3-a.	Explain in detail how to implement Find S Algorithm. (CO1,K2)	6
3-b.	What is a hypothesis explain most specific and most general hypothesis?(CO1,K2)	6
3-c.	Explain the AGNES (Agglomerative Nesting) algorithm for hierarchical clustering in detail.(CO2,K2)	6
3-d.	Explain Bayesian Learning and the concept of the Bayes Optimal Classifier.(CO4,K2)	6
3.e.	Discuss the application of Reinforcement Learning (RL) in real-world scenarios such as robotics, healthcare, and autonomous vehicles.(CO5,K2)	6
3.f.	Discuss the fundamental differences between Bagging and Boosting.(CO4,k2)	6
3.g.	Describe in detail all the steps involved in designing a learning system.	6

## **SECTION-C** 50

4. Answer any one of the following:-

4-a.	Describe the different approaches to machine learning, such as symbolic learning, connectionist learning, and evolutionary algorithms.(CO1,K2)	10
4-b.	Explain the process of model building in machine learning. What are the key steps involved, and how do you evaluate and improve the performance,(CO1,K2)	10

5. Answer any one of the following:-

5-a.		10
------	--	----

**Use the ID3 algorithm, explain how you would construct a decision tree for the following dataset?(CO2,K3)**

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

- 5-b. Discuss an example of how association rule learning can be applied in a machine learning task, such as customer segmentation or recommendation systems.(CO2,K3) 10
6. Answer any one of the following:-
- 6-a. Explain the density based clustering with suitable example.. (CO3,K2) 10
- 6-b. How is K-NN used for both classification and clustering, and what role does the distance metric play in clustering tasks?(CO3,K2) 10
7. Answer any one of the following:-
- 7-a. Explain the XGBoost algorithm and why it has become a popular choice for machine learning tasks.(CO4,K3) 10
- 7-b. Discuss the main advantages and limitations of using the Random Forest algorithm for classification and regression tasks.(CO4,K2) 10
8. Answer any one of the following:-
- 8-a. Describe its key components, including the Q-table, action selection, and reward system. Explain the Q-learning algorithm.(CO5,K2) 10
- 8-b. What are some common use cases for reinforcement learning algorithms? (CO5) 10