

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

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

B.Tech

SEM: VI - 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) What is machine learning?[CO1,K1]

1

- (a) A subset of artificial intelligence
- (b) A type of computer programming language
- (c) The study of human learning
- (d) None of the above

(1.2) Which of the following is NOT a type of machine learning?[CO1,K1]

1

- (a) Supervised learning (b) Traditional programming
- (c) Unsupervised learning (d) Reinforcement learning

(1.3) In machine learning, what task involves assigning an input data point to a category or label?[CO2,K1]

1

- (a) Classification (b) Regression
- (c) Clustering (d) Association

(1.4) What is the fundamental building block of a neural network?[CO2,K1]

1

- (a) Node (b) Edge (c) Neuron (d) Leaf

(1.5) What is the primary goal of clustering in machine learning?[CO3,K1]

1

- (a) Classification (b) Data compression
- (c) Dimensionality reduction (d) Grouping similar data points

(1.6) Which clustering algorithm is known for forming clusters of arbitrary shapes and is sensitive to noise?[CO3,K1]

1

- (a) K-Means clustering (b) K-Mode clustering
(c) DBSCAN (d) Expectation Maximization (EM)
- (1.7) What is Bayesian Learning primarily concerned with?[CO4,K1] 1
(a) Unsupervised learning (b) Supervised learning
(c) Probabilistic reasoning (d) Feature engineering
- (1.8) In ensemble methods like Bagging, what is the primary goal?[CO4,K1] 1
(a) Reduce overfitting (b) Increase bias
(c) Reduce variance (d) Boost the accuracy of a single model
- (1.9) 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
- (1.10) In Reinforcement Learning, what is the primary objective of the learning task?[CO5,K1] 1
(a) Minimize the error between predicted and actual outputs
(b) Maximize the cumulative reward over time
(c) Discover hidden patterns in the data
(d) Optimize feature selection
2. Attempt all parts:-
- (2.1) What does bias refer to in machine learning?[CO1,K1] 2
(2.2) What is the role of the dependent variable in regression analysis, and how does it differ from the target variable in classification?[CO2, K2] 2
(2.3) Explain the concept of "Iterative distance-based clustering"? [CO3,K2] 2
(2.4) How would you apply the bootstrapping technique to estimate the accuracy of a machine learning model on a given dataset?[CO4,K3] 2
(2.5) In what types of problems is Q Learning commonly applied?[CO5,K2] 2

SECTION-B

30

Answer any one of the following:-

- (3.1) How would you apply the steps of designing a learning system to build a machine learning solution for a real-world problem (e.g., house price prediction or spam detection)? [CO1,K2] 6
(3.2) Summarize the history of machine learning, highlighting significant milestones and developments.[CO1,K2] 6

Answer any one of the following:-

- (3.3) Explain the main differences between classification and regression problems in machine learning. Provide examples to illustrate each? [CO2,K2] 6

- (3.4) Given the following training dataset, apply a decision tree algorithm to construct the first split using an appropriate measure (e.g., information gain or Gini index). Show all calculations and determine the best attribute for splitting?[CO2,K3] 6

Outlook	Temperature	Humidity	Play
Sunny	Hot	High	No
Sunny	Hot	High	No
Overcast	Hot	High	Yes
Rain	Mild	High	Yes
Rain	Cool	Normal	Yes

Answer any one of the following:-

- (3.5) Compare and contrast the hierarchical clustering methods AGNES (Agglomerative Nesting) and DIANA (Divisive Analysis).[CO3,K2] 6
- (3.6) Explain K-Mode Clustering method in detail.[CO3,K2] 6

Answer any one of the following:-

- (3.7) Given a dataset with high variance and noise, how would you apply either bagging or boosting to improve model performance? Explain which method you would choose and how you would implement it in this situation.?[CO4,K3] 6

ID	Feature1	Feature2	Class
1	2	3	A
2	3	4	A
3	10	15	B
4	12	18	B
5	2	2	A
6	11	17	B
7	3	3	A
8	9	14	B

- (3.8) Given the following dataset, apply the Naïve Bayes classifier to predict whether a new instance belongs to class Yes or No using Bayes' theorem.?[CO4,K3] 6

Outlook	Humidity	Play
Sunny	High	No
Sunny	High	No
Overcast	High	Yes
Outlook	Humidity	Play
Rain	High	Yes
Rain	Normal	Yes
Sunny	Normal	Yes

Answer any one of the following:-

- (3.9) Explain the role and significance of the state transition function in Markov Decision Processes. (CO5, K2) 6

(3.10)	Demonstrate how Q-values are optimized in Reinforcement Learning? (CO5, K2)	6															
SECTION-C		50															
4.	Answer any <u>one</u> of the following:-																
(4.1)	Explain the fundamental concepts of learning in machine learning, emphasizing differences between supervised and unsupervised learning.[CO1,K2]	10															
(4.2)	What characterizes a "well-defined learning problem" in machine learning, and why is it crucial for model development? Provide examples.[CO1,K2]	10															
5.	Answer any <u>one</u> of the following:-																
(5.1)	In multiple linear regression, how do independent variables help in predicting the dependent variable? Provide an example to substantiate your point. (CO2, K3)	10															
(5.2)	Describe logistic regression model along with its advantages and disadvantages.[CO2,K2]	10															
6.	Answer any <u>one</u> of the following:-																
(6.1)	Given the following data points, apply the K-means clustering algorithm with K = 2. Show the steps of selecting initial centroids, assigning clusters, and updating centroids until convergence.? [CO3,K3]	10															
<table> <tr> <th>Point</th><th>X</th><th>Y</th></tr> <tr> <td>A</td><td>1</td><td>2</td></tr> <tr> <td>B</td><td>2</td><td>1</td></tr> <tr> <td>C</td><td>4</td><td>5</td></tr> <tr> <td>D</td><td>5</td><td>4</td></tr> </table>			Point	X	Y	A	1	2	B	2	1	C	4	5	D	5	4
Point	X	Y															
A	1	2															
B	2	1															
C	4	5															
D	5	4															
(6.2)	How do DBSCAN and OPTICS clustering methods handle noisy data and irregular clusters?[CO3,K2]	10															
7.	Answer any <u>one</u> of the following:-																
(7.1)	Analyze how Gradient Boosting Machines (GBM) work by breaking down their sequential learning process. Then compare GBM with another ensemble method (such as Random Forest) in terms of error reduction, model building strategy, and handling of bias and variance. Also, analyze which real-world scenarios are better suited for GBM and justify your reasoning.[CO4,K4]	10															
(7.2)	Analyze the working of XGBoost by breaking down its key components such as regularization, tree pruning, handling missing values, and parallel processing. Then compare how these features improve the standard gradient boosting algorithm in terms of speed, accuracy, and overfitting control.[CO4,K4]	10															
8.	Answer any <u>one</u> of the following:-																
(8.1)	In Reinforcement Learning, what's the main goal, and how does it differ from other types of learning?[CO5,K2]	10															
(8.2)	Explain various real-world examples of Reinforcement Learning.[CO5,K2]	10															