Printed Page:- 03 **Subject Code:- ACSE0515**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: V - 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. Examples of Ordinal data can be (CO1) 1

(a) ID Numbers, eye color, zip codes

(b) Rankings, taste of potato chips, grades, height

(c) Calendar dates, temperatures in Celsius or Fahrenheit, phone numbers

(d) Temperature in Kelvin, length, time, counts

1-b. Identify the type of learning in which labeled training data is used. (CO1) 1

(a) Semi Supervised learning

(b) Supervised Learning

(c) Reinforcement Learning

(d) Unsupervised Learnng

1-c. We can compute the coefficient of linear regression by using (CO2) 1

(a) gradient descent

(b) Normal Equation

(c) both A and B

(d) None of above

1-d. In a logistic regression model, the decision boundary can be (CO2) 1

(a) Linear

(b) Non-linear

(c) Both (A) and (B)

(d) None of these

1-e. Which of the following is NOT True about Ensemble Techniques? (CO3) 1

(a) Bagging decreases the variance of the classifier.

(b) Boosting helps to decrease the bias of the classifier.

(c) Bagging combines the predictions from different models and then finally gives the results.

(d) Bagging and Boosting are the only available ensemble techniques.

1-f. Decision Tree is (CO3) 1

(a) Flow-Chart

(b) Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label

(c) Flow-Chart & Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label

(d) None of the mentioned

1-g. Reinforcement learning is- (CO4) 1

(a) Unsupervised learning

(b) Supervised learning

(c) Award based learning

(d) None

1-h. Which crossover method is not exist in Genetic Algorithm (CO4) 1

(a) Single

(b) Multi

(c) Uniform

(d) Random

1-i. Does gradient boosted trees generally perform better than random forest? (CO5) 1

(a) Yes

(b) No

(c) Can't say

(d) None

1-j. Which of the following is true about Xgboost (CO5) 1

(a) Like any other boosting method, XGB is sensitive to outliers

(b) Unlike LightGBM, in XGB, one has to manually create dummy variable/label before feeding to the model

(c) 1 or 2 both

(d) None of these

2. Attempt all parts:-

2.a. Define the role of Machine Learning in our daily life. (CO1) 2

2.b. Briefly explain the formulas to find RMSE and MSE. (CO2) 2

2.c.	Naive Bias is a classification algorithm or regression algorithm, describe. (CO3)	2
2.d.	Provide an intuitive explanation of what is a Policy in Reinforcement learning. (CO4)	2
2.e.	Explain ensemble technique that is used by gradient boosting trees (CO5)	2

SECTION-B 30

3. Answer any five of the following:-

3-a.	Explain Various Machine Learning Libraries and its importance. (CO1)	6
3-b.	Explain the candidate elimination algorithm. (CO1)	6
3-c.	Can we solve the multiclass classification problems using Logistic Regression? If Yes then How? (CO2)	6
3-d.	Explain overfitting and underfitting. What causes overfitting? (CO2)	6
3.e.	Explain ID3 algorithm with example. (CO3)	6
3.f.	Compare Q-Learning and Policy Gradients methods.(CO4)	6
3.g.	Discuss one real-world scenario where you have used xgboost.(CO5)	6

SECTION-C 50

4. Answer any one of the following:-

4-a.	Explain One hot encoding. What is the problem with One hot encoding (CO1)	10
4-b.	Describe in detail all the steps involved in designing a learning system. (CO1)	10

5. Answer any one of the following:-

5-a.	Explcain confusion matrix. Explain accuracy, precision and recall with example. (CO2)	10
5-b.	Explain when will you use Classification over Regression. Justify your answer with example. (CO2)	10

6. Answer any one of the following:-

6-a.	Explain SVM with their kernel functions. (CO3)	10
6-b.	Differentiate Post-pruning and Pre-pruning methods. Also give the requirements of Post-pruning and Pre-pruning (CO3)	10

7. Answer any one of the following:-

7-a.	Describe PCA in detail. Also explain the work of PCA. (CO4)	10
7-b.	List down the steps of Genetic algorithm with example. (CO4)	10

8. Answer any one of the following:-

8-a.	Explain different types of Memory-Based Collaborative approaches (CO5)	10
8-b.	How are Knowledge-based Recommender Systems different from Collaborative and Content-based Recommender Systems? Discuss applications of Recommender Systems. (CO5)	10