Printed Page:- 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 (2022 - 2023)

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 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 Which of the Following Step is not involved in Data Cleaning (CO1) 1

(a) Removal of unwanted observations

(b) Fixing Structural errors

(c) Binarization Data

(d) Handling missing data

1 _____ calculated as the sum of absolute differences between the coordinates of data point and centroid of each class. (CO2) 1

(a) Euclidean Distance

(b) Feature Scaling

(c) Manhattan Distance

(d) Minkowski Distance

1 What is Regression? (CO2) 1

(a) It is a technique to predict values

(b) It is a technique to find outliers

(c) It is a technique to fix data

(d) It is a Machine Learning algorithm

1 In a naive Bayes algorithm, when an attribute value in the testing record has no example in the training set, then the entire posterior probability will be zero. (CO3) 1

(a) True

(b) FALSE

(c) Can’t determined

(d) None of these

1 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 Reinforcement learning is- (CO4) 1

(a) Unsupervised learning

(b) Supervised learning

(c) Award based learning

(d) None

1 The multi-armed bandit problem is a generalized use case for- (CO4) 1

(a) Reinforcement learning

(b) Supervised learning

(c) Unsupervised learning

(d) All of the above

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

(a) Yes

(b) No

(c) Can't say

(d) None

1	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	Explain the candidate elimination algorithm. (CO1)	6
3	Explain well posed learning system with example. (CO1)	6
3	Explain overfitting and underfitting. What causes overfitting? (CO2)	6
3	Explain how would you train Linear Regression model using Gradient Descent. (CO2)	6
3.e.	Explain ID3 algorithm with example. (CO3)	6
3.f.	Compare Q-Learning and Policy Gradients methods.(CO4)	6
3.g.	Explain some advantages of using Neighborhood-based approaches for Recommender Systems (CO5)	6

SECTION C 50

4. Answer any one of the following:-

4	Describe in detail all the steps involved in designing a learning system. (CO1)	10
4	Compare supervised ,Unsupervised and Re-enforcement learning with example (CO1)	10

5. Answer any one of the following:-

5	What is confusion matrix? Explain accuracy, precision and recall with example. (CO2)	10
5	Explain when will you use Classification over Regression. Justify your answer with	10

example. (CO2)

6. Answer any one of the following:-

6	Explain SVM with their kernel functions. (CO3)	10
6	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	Descibe about PCA. What does a PCA do? (CO4)	10
7	Discuss the different machine learning tools and applications. (CO4)	10

8. Answer any one of the following:-

8	Explain different types of Memory-Based Collaborative approaches (CO5)	10
8	Describe LightGBM and also compare LightGBM vs XGBoost. (CO5)	10