Printed Page:- 04 **Subject Code**:- **ACSML0401N**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: IV CARRY OVER THEORY EXAMINATION - AUGUST 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-a. Which of the following are common classes of problems in machine learning? (CO1) 1

(a) Regression

(b) Classification

(c) Clustering

(d) All of the above

1-b. What is Machine Learning? (CO1) 1

(a) Artificial Intelligence

(b) Deep Learning

(c) Deep Statistics

(d) Both a and b

1-c. Explain the term outliers in detail. (CO2) 1

(a) It is the main trend of our dataset

(b) Extreme datapoints in our dataset

(c) It is a regression technique

(d) Values that are correlated to eachother

1-d. In general, to have a well-defined learning problem, we must identity which of the following (CO2) 1

(a) The class of tasks

(b) The measure of performance to be improved

(c) The source of experience

(d) All of the above

1-e. What is the minimum no. of variables/ features required to perform clustering? (CO3) 1

(a) 0

(b) 1

(c) 2

(d) 3

1-f. Decision tree is a flowchart like_____ (CO3) 1

(a) leaf structure

(b) tree structure

(c) steam

(d) none of these

1-g. _____ is an ensemble learning method which trains multiple models independently in parallel. (CO4) 1

(a) Bagging

(b) Boosting

(c) Perceptron

(d) None of these

1-h. Which of the following statements is true about the learning rate alpha in gradient descent? (CO4) 1

(a) If alpha is very small, gradient descent will be fast to converge. If alpha is too large, gradient descent will overshoot

(b) If alpha is very small, gradient descent can be slow to converge. If alpha is too large, gradient descent can be slow too

(c) If alpha is very small, gradient descent can be slow to converge. If alpha is too large, gradient descent will overshoot

(d) None of these

1-i. Which of the following is an application of reinforcement learning? (CO5) 1

(a) Topic modeling

(b) Recommendation system

(c) Pattern recognition

(d) Image classification

1-j. Hidden Markov Model is used in- (CO5) 1

(a) Supervised learning

(b) Unsupervised learning

(c) Reinforcement learning

(d) All of the above

2. Attempt all parts:-

2.a. Discuss some useful applications of machine learning. (CO1) 2

2.b. Explain Dependent and Independent Variable. (CO2) 2

2.c. Explain complete linkage in Hierarchical Clustering. (CO3) 2

2.d. Explain Conditional probability. (CO4) 2

2.e. Write the application areas of Reinforcement Learning. (CO5) 2

SECTION B **30**

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

3-a. What is a hypothesis? Explain most specific and most general hypothesis? (CO1) 6

3-b. How do you design a checkers learning problem? (CO1) 6

3-c. Differentiate between Logistic Regression and Linear Regression Algorithm. (CO2) 6

3-d. How Apriori algorithm is different from FP Growth algorithm? (CO2) 6

3.e. Explain Density Based clustering with an example. (CO3) 6

3.f. Distinguish between Bagging and Boosting. (CO4) 6

3.g. How Reinforcement Learning is different Semi Supervised Learning ? (CO5) 6

SECTION C **50**

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

4-a. Explain find –S algorithm with given example. Give its application. (CO1) 10

Time	Weather	Temperature	Go for Walk	Humidity	Wind
Morning	Sunny	Warm	Yes	Mild	Strong
Evening	Rainy	Cold	No	Mild	Normal
Morning	Sunny	Moderate	Yes	Normal	Normal
Evening	Sunny	Cold	Yes	High	Strong

4-b. Define Machine Learning. Differentiate between Over fitting and Under fitting. Explain with the help of diagram (CO1) 10

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

5-a. Explain all the types of Regression methods in detail. (CO2) 10

5-b. Build decision tree using ID3 algorithm for following dataset (CO2) 10

Day No.	Outlook	Temp	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Predict class label for:

S. No.	Outlook	Temp	Humidity	Wind	Play Tennis
D15	Sunny	Hot	Normal	Weak	

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

6-a. Explain Hierarchical Clustering. Consider the given dataset in the form of similarity matrix and explain how clustering can be done for the same.(CO3) 10

Observations	P1	P2	P3	P4	P5
P1	0				
P2	9	0			
P3	3	7	0		
P4	6	5	9	0	
P5	11	10	2	8	0

6-b. What do you mean by clustering? Explain iterative distance based clustering with suitable example. (CO3) 10

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

7-a. Differentiate between Random Forest and C4.5 (CO4) 10

7-b. What is 'training Set' and 'test Set' in a Machine Learning Model? How Much Data Will You Allocate for Your Training, Validation, and Test Sets? (CO4) 10

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

8-a. Explain the key elements of Reinforcement Learning in detail. (CO5) 10

8-b. Explain the Q function and Q Learning Algorithm assuming deterministic rewards and actions with example. (CO5) 10