National Forensics Sciences University, Goa Campus
TA-1 Examination

Program Name – MTECH AIDS **Sem – II** **Date- 10-02-2025**
Subject - Advanced Machine Learning for Cyber Security and Forensics (CTMTAIDS SII P1)
Time- 45 mins **Max. Marks- 25**
Instructions - 1) Answer all questions. 2) Assume suitable data.

Q.1 **Multiple Choice Questions (1 mark each)** 10 marks

1. Which of the following statements is true about bias-variance tradeoff in machine learning?
   A) Increasing model complexity always reduces bias and increases variance
   B) A high-bias model will generally have high variance
   C) Reducing bias can lead to a high-variance model
   D) Variance is independent of the model's complexity

2. Which of the following scenarios is best suited for using semi-supervised learning?
   A) Predicting outcomes for a dataset where both inputs and outputs are labeled
   B) Predicting house prices where most data points are labeled, but some are not
   C) Clustering large datasets without any labeled data
   D) Designing a rule-based system without any learning involved

3. The Apriori algorithm is commonly used to find association rules in market basket analysis. What is the main limitation of the Apriori algorithm?
   A) It can only find high-dimensional associations
   B) It requires the dataset to be clean and without missing values
   C) It generates a large number of candidate itemsets
   D) It can only find strong rules with high support and confidence

4. In Support Vector Machines (SVM), what is the primary objective when finding the optimal hyperplane?
   A) To maximize the margin between the classes
   B) To minimize the classification error
   C) To ensure that all support vectors are correctly classified
   D) To minimize the distance between the hyperplane and the support vectors

5. What is the primary difference between Ridge regression and Lasso regression in terms of regularization?
   A) Ridge regression uses L1 regularization, while Lasso uses L2 regularization
   B) Lasso can shrink coefficients to zero, while Ridge cannot
   C) Lasso regression is more computationally expensive than Ridge regression
   D) Ridge regression has no regularization parameter, while Lasso does

6. Which of the following is the main limitation of the K-means clustering algorithm?
A) It works well only with data that has a Gaussian distribution
B) It requires the number of clusters to be specified in advance
C) It cannot handle data with missing values
D) It only works for data that is linearly separable

7. In Q-learning, which of the following is updated iteratively to determine the best action for the agent?
A) State-action values (Q-values)
B) The learning rate (α)
C) The reward function
D) The policy

8. Which of the following is NOT typically used in supervised learning?
A) Labeled training data
B) Predictive model
C) Unlabeled data
D) Training error minimization

9. Which of the following methods is used to handle categorical variables in machine learning models?
A) Feature scaling
B) One-hot encoding
C) Min-max normalization
D) K-means clustering

10. In decision tree learning, which of the following methods is used to determine the best split at each node?
A) Gini index or entropy
B) Cross-validation
C) Linear regression
D) K-fold validation
Answer: A) Gini index or entropy

Q.2 **Answer any 3 questions (3x5 marks each)** 15 Marks

A. What is Reinforcement Learning and the key concepts. 5 marks

B. What are the Data and Model Related Challenges in ML 5 marks

C. What is Hypothesis class and Version Space 5 marks

D. Write a Formula for MSE, Sigmoid, Tanh, ReLU, Cross-Entropy Loss 5 marks

END