Hall Ticket No | | | | | | | | | | | Question Paper Code: ACAC03

INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)
Dundigal-500043, Hyderabad

B.Tech IV SEMESTER END EXAMINATIONS (REGULAR/SUPPLEMENTARY) - AUGUST 2023
Regulation: UG-20
FOUNDATIONS OF MACHINE LAERNING

Time: 3 Hours (CSE(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)) **Max Marks: 70**

Answer ALL questions in Module I and II
Answer ONE out of two questions in Modules III, IV and V
All Questions Carry Equal Marks
All parts of the question must be answered in one place only

MODULE – I

1. (a) How does the process of machine learning work? Mention the common learning problems and scenarios where machine learning techniques are applied. [BL: Understand| CO: 1|Marks: 7]

   (b) Imagine you are working on a project to develop a recommendation system for an e-commerce platform. The goal is to suggest personalized product recommendations to users based on their browsing history and purchase behavior. Considering the different types of learning, which learning approach (supervised, unsupervised, reinforcement, or semi-supervised) would be most suitable for this scenario? Justify your choice by explaining how the selected learning approach aligns with the problem requirements and the available data. [BL: Apply| CO: 1|Marks: 7]

MODULE – II

2. (a) What are the fundamental principles and key characteristics that differentiate linear and non-linear classification algorithms in machine learning? [BL: Understand| CO: 2|Marks: 7]

   (b) Consider a scenario where you have collected data on patients with a particular disease, including variables such as age, gender, and medical test results. How could logistic regression be applied to predict the probability of disease progression or treatment success? Discuss the assumptions of logistic regression and how they apply in this medical context. [BL: Apply| CO: 2|Marks: 7]

MODULE – III

3. (a) Write about error-correcting output codes (ECOC) in ensemble learning. How do they enable the classification of multi-class problems using binary classifiers? [BL: Understand| CO: 3|Marks: 7]

   (b) In the context of voting in ensemble learning, what challenges might arise when dealing with models that have differing levels of performance or expertise? How could you address these challenges to improve the overall ensemble performance? [BL: Apply| CO: 3|Marks: 7]

4. (a) Explain the bagging technique in ensemble learning, specifically focusing on how random forest combines multiple decision trees to make predictions and mitigate overfitting. [BL: Understand| CO: 4|Marks: 7]

   (b) Boosting algorithms like Adaboost aim to improve the overall ensemble performance by iteratively focusing on misclassified instances. However, what challenges might arise when dealing with noisy or mislabeled data? How could you mitigate these challenges in the boosting process? [BL: Apply| CO: 4|Marks: 7]

MODULE – IV

5. (a) Discuss the purpose of clustering in data analysis, and how does it differ from other types of machine learning techniques? [BL: Understand| CO: 5|Marks: 7]

   (b) Locally linear embedding (LLE) aims to preserve local relationships in the data. However, what potential issues or limitations might arise when applying LLE to datasets with complex or nonlinear structures? How could these limitations be addressed? [BL: Apply| CO: 5|Marks: 7]

6. (a) Outline the concept of factor analysis and how it can be used for clustering by capturing the latent factors underlying the observed variables? [BL: Understand| CO: 5|Marks: 7]

   (b) Consider a dataset of customer transactions at a grocery store, with features such as purchase amount and frequency. How would you use K-means clustering to segment customers into different groups based on their shopping behaviors? [BL: Apply| CO: 5|Marks: 7]

MODULE – V

7. (a) Write the fundamental concept behind neural networks, and how do they differ from traditional computing models? [BL: Understand| CO: 6|Marks: 7]

   (b) Consider a dataset of animals, labeled as either "mammal" or "reptile," with features like weight and body temperature. Using the K-nearest neighbors algorithm, how would you classify a new animal as either a "mammal" or a "reptile" based on its weight and body temperature? [BL: Apply| CO: 6|Marks: 7]

8. (a) Differentiate between support vector machines (SVMs) used for linear and non-linear classification. [BL: Understand| CO: 6|Marks: 7]

   (b) When comparing neural networks and support vector machines, what are the key differences in terms of model interpretability, scalability, and computational efficiency. How would these factors influence the choice between the two approaches in practical applications? [BL: Apply| CO: 6|Marks: 7]