

[Total No. of Questions: 09]

[Total No. of Pages: 02]

Uni. Roll No.

Program: B.Tech. (Batch 2018 onward)

Semester: 6

Name of Subject: Introduction to Machine Learning

Subject Code: PCIT-114

Paper ID: 17206

EVENING

13 JUN 2023

Time Allowed: 03 Hours

Max. Marks: 60

NOTE:

- 1) Parts A and B are compulsory
- 2) Part-C has Two Questions Q8 and Q9. Both are compulsory, but with internal choice
- 3) Any missing data may be assumed appropriately

Part – A

[Marks: 02 each]

Q1.

- a) Define machine learning and traditional programming and highlight the differences between the two approaches.
- b) Name and briefly define three different regression techniques used in machine learning.
- c) What are the main types of classification algorithms used in machine learning? Provide examples of each type.
- d) Elaborate on the concept of clustering and mention two different clustering algorithms used in machine learning.
- e) 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?
- f) What is Overfitting, and How Can You Avoid It?

Part – B

[Marks: 04 each]

- Q2.** Describe the working principles of simple linear regression and multiple linear regression. Explain when and how these techniques are used in machine learning.
- Q3.** Discuss the concept of fuzzy logic and explain the steps involved in the fuzzy logic process, including fuzzification and defuzzification.
- Q4.** Compare and contrast the decision tree classification and random forest classification algorithm. Explain their differences and applications in machine learning.

Q5. Assume we have a dataset of patients of some hospital. The dataset contains various features such as age, blood pressure, cholesterol level, and body mass index for each patient. Our goal is to classify whether a new visiting patient has any of the following diseases: diabetes, heart disease, or Alzheimer's. Elaborate best way for the same.

EVENING

Q6. How to handle missing or corrupted data in a dataset?

Q7. Investigate the advantages and limitations of the K-means clustering algorithm and the expectation-maximization algorithm in machine learning. Compare their performance on different types of datasets and discuss their practical implications

Part – C

[Marks: 12 each]

Q8. Explain the Confusion Matrix with Respect to Machine Learning Algorithms. What Is a Predicted Positive and Predicted Negative and How Are They Significant?

	Predicted Positive	Predicted Negative
Actual Positive	65	10
Actual Negative	15	10

- Calculate the accuracy, precision, recall, and F1 score for this classification model.
- Discuss the limitations of using a confusion matrix as a standalone evaluation metric for classification models.

OR

What is the difference between

- Classification and Clustering
- Supervised and unsupervised machine learning
- Classification and Regression

Give examples of using clustering to solve real-life problems

Q9. Discuss the concept of classification in machine learning. Compare and contrast the working principles and applications of at least four different classification algorithms. Provide real-world examples where each algorithm can be applied and highlight their strengths and limitations.

OR

Evaluate the role of fuzzy logic in machine learning. Discuss the principles of fuzzy logic, including fuzzy sets and membership functions. Explain how fuzzy logic can be applied in real-world scenarios and the advantages it offers over traditional binary logic. Provide case studies or examples to illustrate your points.
