

2024
MCS111 MACHINE LEARNING

Full Marks: 100

Time: Three hours

Answer any FIVE questions.

- 1 a) Explain how the k-Nearest Neighbors (k-NN) method works in classification. What are the advantages and limitations of using k-NN? 10
b) What is logistic regression, and how does it differ from linear regression? Describe an example where logistic regression would be more suitable than linear regression. 10
- 2 a) Describe the K-means clustering algorithm and explain how it groups data points into clusters. What are some limitations of K-means? 10
b) What is Principal Component Analysis (PCA) and how is it used for dimensionality reduction? Provide an example of a situation where PCA might be useful. 10
- 3 a) Explain the purpose of evaluating machine learning algorithms. Describe two common metrics used to evaluate classification models. 10
b) What is ensemble learning, and how does the Random Forest algorithm work as an ensemble method? 10
- 4 a) What is sparse modeling, and why is it useful in machine learning? Give an example of a scenario where sparse modeling might be applied. 10
b) Explain the importance of modeling sequence or time-series data. What challenges are unique to time-series data compared to regular datasets? 10
- 5 a) Explain the concept of online learning in machine learning and discuss how it can be applied to large-scale data. 10
b) Explain the concepts of Semi-supervised Learning (SSL) and Active Learning (AL). Provide an example for each. 10
- 6 a) Discuss the recent trends in machine learning techniques that are shaping the future of artificial intelligence. 10
b) Discuss some commonly used classification algorithms, such as decision trees, support vector machines, and k-nearest neighbors, and how they are applied to IoT data. 10
