Roll No											

PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF COMMERCE
END TERM EXAMINATION - JUN 2023

Semester : Semester IV - 2021
Course Code : BSE2014
Course Name : Sem IV - BSE2014 - Data Mining and Forecasting
Program : BSE

Date : 12-JUN-2023
Time : 1.00PM - 4.00PM
Max Marks : 100
Weightage : 50%

Instructions:
(i) Read all questions carefully and answer accordingly.
(ii) Question paper consists of 3 parts.
(iii) Scientific and non-programmable calculator are permitted.
(iv) Do not write any information on the question paper other than Roll Number.

PART A

ANSWER ALL THE QUESTIONS **(10 X 2 = 20M)**

1. What are the steps involved in data processing?

(CO2) [Knowledge]

2. How does neural network modeling work, and what are some practical applications of this technique in industry and business?

(CO3) [Comprehension]

3. Explain two Data Mining applications in present business world.

(CO1) [Knowledge]

4. What is classification in data mining, and how is it used in predictive modeling?

(CO3) [Comprehension]

5. How do you handle missing values during data preprocessing?

(CO2) [Knowledge]

6. What is clustering in data mining?

(CO3) [Comprehension]

7. What is the Difference between Data and Information?

(CO1) [Knowledge]

8. Define cluster analysis. Further how is it used in data mining?

(CO4) [Comprehension]

9. Explain ACID Components.

(CO1) [Knowledge]

10. List the common data processing methods used in industry and business?

(CO2) [Comprehension]

PART B

ANSWER ALL THE QUESTIONS **(4 X 10 = 40M)**

11. What are the advantages and disadvantages of using a decision tree algorithm for classification in data mining?

(CO3) [Comprehension]

12. What are the different types of web mining, and how are they used in practice?

(CO4) [Comprehension]

13. What are the major issues in data mining?

(CO1) [Comprehension]

14. Explain the concept of data mining, and describe the key steps involved in the data mining process.

(CO2,CO5) [Comprehension]

PART C

ANSWER ALL THE QUESTIONS **(2 X 20 = 40M)**

15. A real estate company wants to improve its pricing strategy by predicting housing prices in a particular region. The company has a large dataset of historical housing sales data, as well as demographic and economic indicators such as population growth, income levels, and employment rates. The company wants to use data mining techniques to analyze the data and develop a predictive model to forecast future housing prices.

What variables were found to be most significant in predicting housing prices in the studied region, and how were these variables incorporated into the predictive model?

(CO5) [Application]

16. A city government wants to address the issue of housing affordability by implementing a price ceiling on rental units. The price ceiling will limit the amount that landlords can charge for rent, with the goal of making housing more affordable for low-income residents. However, the government wants to ensure that the price ceiling does not have unintended consequences on the housing market, such as a decrease in the quality or quantity of rental units. The government decides to use data analysis techniques to evaluate the effectiveness of the price ceiling policy.

a. What data sources were used to collect information on rental prices, housing quality, and occupancy rates?

b. Were there any other economic variables or factors taken into account during the analysis, such as population growth, income levels, or unemployment rates?

c. Write the Data Analysis Process using the above case.

(CO2) [Comprehension]