Data Mining Using R-Sem 3

Unit 1: Introduction to Data Mining

1. What is data mining? Explain its importance and applications.
2. Describe the steps in the data mining process.
3. Differentiate between descriptive and predictive data mining.
4. Explain data preprocessing techniques like data cleaning, transformation, and reduction.
5. What is the role of data mining in decision support systems?

Unit 2: Introduction to R

1. What are the features of R? Why is R popular in data mining?
2. Explain R data types: vectors, matrices, lists, and data frames with examples.
3. Write R commands to:
   * Create a vector and perform arithmetic operations
   * Create a data frame and access elements
4. Explain control structures (`if`, `for`, `while`) in R with examples.
5. How is data imported into R? Show with `read.csv()` and `read.table()` functions.

Unit 3: Data Visualization in R

1. What is data visualization? Why is it important in data mining?
2. Explain the use of `plot()`, `hist()`, `boxplot()` in R with syntax.
3. What are different types of charts available in R?
4. Write an R program to plot a bar chart and pie chart.
5. How do you customize graphs in R (titles, colors, labels)?

Unit 4: Data Mining Techniques Using R

1. What is classification? Explain with R example using `rpart` or `caret` package.
2. What is clustering? Perform k-means clustering using R with sample data.
3. Explain association rule mining. Demonstrate `apriori()` using `arules` package.
4. Write R code to normalize a dataset.
5. How is model evaluation done in R? (accuracy, confusion matrix, etc.)

Unit 5: Case Studies and Applications

1. Explain a case study of data mining in retail or healthcare using R.

2. How can R be used for sentiment analysis?

3. Discuss the steps involved in a real-world data mining project using R.

4. How is big data related to data mining and R?

5. Describe how R helps in performing exploratory data analysis (EDA).

- Frequently Asked 10 Marks Questions

* Explain different stages of the data mining process with examples.

* Write a complete R script to perform clustering and visualize results.

* Describe classification and show how it is implemented in R.

* Discuss association rule mining with output interpretation.

* Perform EDA on a dataset using R and explain the findings.