M. Sc. – Semester – IV
Session 2021-22

Assignment List

Paper I: Data Mining

Unit I

1) What is data mining? Explain Motivating challenges for data mining and also explain by data cane mine.
2) Explain different data mining task.
3) What is data? Explain different types of dataset.
4) What is data quality and also explain issues related to data collection and application.
5) Explain data processing in brief.
6) Write a note on measures of data similarity and dissimilarity.

Unit II

1) Write a note on different techniques used for data exploration using statistic.
2) What is data visualization? Explain different techniques used for data visualization.
3) What is OLAP? Explain OLAP in detail for multidimensional data analysis.
4) What is classification? Explain different classification techniques.
5) What is decision tree? Explain how decision tree can build.
6) Explain methods for comparing classifiers.
7) Explain how performance of classifier can evaluate.
8) Write a note on: Explain issues for model overfitting

UNIT III

1. Differentiate between Naive Bayes classifier and Rule–based classifier.
2. Write a note on support based processing.
3. Write a note on FP Growth Algorithm.
4. Explain nearest neighbour classifier in detail.
5. Explain the following:
   (i) Tree Projection.
   (ii) Pattern Evaluation.

UNIT IV

1. What is cluster? Explain different types of cluster.
2. What is Anomaly detection? Differentiate between outlier and anomaly detection.
3. Explain following algorithm:
   I. Rock
   II. K-Means
4. Discuss characteristics of spatial data sets.
5. Explain clustering algorithm with Hierarchical clustering.