**M. Sc. – Semester – IV**

**Session 2021-22**

**Question Set**

**Paper I: Data Mining**

**Unit I**

**10 or 8 marks questions:**

1. Discuss Data mining in detail. Also explain origins of data mining.
2. What is Association Rule Discovery ? Explain in detail.
3. Explain following :— (i) Attribute Transformation. (ii) Density calculation for data–set.
4. Explain sequential pattern discovery of data mining and spatial data mining
5. What is factor subset selection ? What are the advantages of it ? Give any one method for it.
6. What are the tasks of data mining ? Discuss in short.
7. Explain dimensionality reduction with one example.
8. Define : (i) Association Rule Discovery (ii) Sequential Pattern Discovery Give one example of each.
9. Explain the need of Data Mining with its scientific viewpoint.
10. Discuss in brief the challenges of Data Mining.
11. Explain in detail Association Rule Discovery.
12. Explain Data Preprocessing in detail.
13. What is data mining ? Explain motivating challenges for data mining.
14. How association rule discovery and clustering helps data mining process ? Discuss.
15. Write short note on data preprocessing and dimensionality reduction.
16. Explain different types of data sets involved in data mining.
17. What are association rules ? Explain sequential pattern discovery with example.
18. Explain the commercial and scientific viewpoint of data in data mining.
19. Write notes on : (i) Discretization and Binarization (ii) Challenges of Data Mining.
20. Explain in detail types of data sets and data quality.
21. Define the following and give example of each : (i) Association Rule Discovery. (ii) Sequential Pattern Discovery.
22. Elaborate the application areas and challenges of Data mining.
23. Explain dimensionality reduction with one example.
24. What are the characteristics of good subset and attributes ? Give any one subset selection method.
25. What is Data Mining ? Explain concept of Commercial viewpoint and Scientific viewpoint.
26. Write notes on : (i) Data Quality (ii) Aggregation.
27. What is clustering ? Explain in detail.
28. Explain the term Attribute Transformation and Density.

**4 or 5 marks questions:**

1. What is feature subset selection ?
2. What are the types of datasets ? Discuss with example.
3. Discuss Discretization and Binarization.
4. Explain sampling data in detail.
5. Explain Attribute Transformation in detail.
6. What are data mining tasks ? Explain.
7. Explain Regression in detail.

**Important topic:**

1. Association Rule Discovery
2. sequential pattern discovery
3. subset selection
4. tasks of data mining
5. Data Preprocessing
6. Clustering
7. Discretization and Binarization
8. Challenges of Data Mining
9. dimensionality reduction
10. Data Quality
11. Aggregation
12. types of datasets
13. Regression

**Unit-II**

**8 or 10 Mark :**

1. What are the data exploration techniques ? Explain.
2. Explain any one algorithm for data classification. Give one example.
3. Explain Data Exploration Techniques.
4. Write notes on :

(i) Box Plots

(ii) Scatter Plots.

1. What is OLAP ? Explain OLAP in detail for multidimensional analysis.
2. What are the different measures of location required for data mining ?
3. What are Data Mining classification ? Explain.
4. What are different measures of spread in Data Mining ?
5. Explain following :

(i) Decision tree.

(ii) ROC curve

1. What are the mathematical measures of data mining?
2. Explain tree induction process in detail.
3. Explain the term ROC curve and discuss interval for

accuracy.

1. What are the practical issues of classification ?
2. Explain different visualization techniques.
3. What are range and variance ? Explain.
4. Write a note on Bask Concepts and Decision Trees.
5. Explain ROC curve with example.
6. Explain :-
7. Histograms
8. Contour Plots.
9. What is OLAP ? How OLAP is useful in datamining ? Explain different OLAP operations.
10. **or 5 Mark** :
11. Explain the working of OLAP.
12. Explain Matrix Plot.
13. Write a note on comparing performance of any two algorithms.
14. What are OLAP Operation ?
15. What is frequency and mode ?
16. Explain measures of Node Impurity.
17. What are different practical issues involved in classification of data mining ?

**UNIT III**

**Summer 2016**

**EITHER**

(a) Explain Data Mining Association Analysis in detail.

10

(b) Differentiate between Naive Bayes classifier and

Rule–based classifier. 10

**OR**

(c) Write a note on :—

(i) Maximal frequent Horible Closed Itemset.

(ii) Tree Project 10

(d) What are the subjective Interestingness Measures ? 10

Explain?

**Winter 2017**

**EITHER**

(a) Explain use of ANN for data classification. 8

(b) Write a note on support based processing. 8

**OR**

(c) Give a note on Naive Bayes classifier. 8

(d) Write a note on FP Growth Algorithm.8

**Summer 2018**

**EITHER**

(A) Explain Instance-Based Classifiers. 10

(B) Write notes on :

(i) Naive Bayes Classifier

(ii) Artificial Neural Networks (ANN). 10

**OR**

(C) Explain Rule-Based Classifier. 10

(D) What is Pattern Evaluation ? Explain Rule Generation. 10

**Winter 2018**

**EITHER**

(A) What is Naive Bayes classifier ? 10

(B) Write notes on :

(i) FP Growth algorithm

(ii) Support Based Pruning. 10

**OR**

(C) Differentiate between association rule discovery and association rule mining. 10

(D) Explain nearest neighbour classifier in detail. 10

**Winter 2019**

**EITHER**

**(**a) Explain use of Artificial Neural Networks (ANN) for data classification. 8

(b) Explain the working of FP Growth Algorithm. 8

**OR**

(c) Write a note on Naive Bayes classifier. 8

(d) Explain the following :

(i) Tree Projection.

(ii) Pattern Evaluation. 8

**Summer 2019**

**EITHER**

(a) Explain Rule-Based classifier. 8

(b) Explain the concept of Association analysis. 8

**OR**

(c) Write notes on : 8

(i) Rule Generation

(ii) Tree Projection.

(d) Explain the properties of a good measure. 8

**4 marks questions**

1. What are the properties of a good measure ?
2. Explain Tree Projection Rule.
3. What is rule-based classifier ?
4. Explain the concept of Hash Tree.
5. What is instance based classifier ?

**UNIT IV**

**8 and 10 marks question.**

1. What is cluster analysis? Explain application of cluster analysis.
2. Write an algorithm for SNN clustering?
3. What is Anomaly detection ? Differentiate between outlier and anomaly detection.
4. Explain following algorithm:
5. Rock
6. K-Means
7. Discuss characteristics of Spatial data sets.
8. Discuss the graph-based clustering with example.
9. Write note on:
10. Rock
11. Jarvis Patrik clustering
12. Explain Shared Nearest Neighbour Approach.
13. Explain clustering algorithm with Hierarchical clustering.
14. Explain Density based clustering in detail.
15. What are the limitations on current merging schemes? Explain in detail.
16. What are the anomaly detection Schemes?
17. What is cluster? Explain different types of cluster.
18. What are clustering algorithm? Explain any one algorithm.

**4 or 5 marks question.**

1. What are the different characteristic of spatial data sets?
2. Explain k-means algorithm.
3. Explain hierarchical clustering.
4. Explain SNN cluster algorithm.
5. What is LOF approach.
6. What are the application of cluster analysis?