11. (a) What is web mining ? What challenges needs to be pose for discovery of effective resource and knowledge from the web ? 7

(b) Explain term Frequency / Inverse Document Frequency (TF/IDF) scheme for indexing text documents. 6

OR

12. (a) Explain Apriory – based approach for mining frequent subgraphs. 7

(b) What is text mining ? Explain Precision and Recall measures for text retrieval. 6

AQ – 1038

Third Semester M. Sc. (Part – II) Examination

(CBCS Pattern)

COMPUTER SCIENCE

3 MCS 1 : Data Mining and Data Warehousing

P. Pages : 4

Time : Three Hours] [Max. Marks : 80

Note : (1) Assume suitable data wherever necessary.
(2) Illustrate your answer with the help of neat sketches.

1. (a) Explain the concept of data integration and data transformation. 8

(b) Explain data smoothing techniques for removing noisy data. 6

OR

2. (a) What is data mining ? State the importance of data mining. 6

(b) Describe attribute subset selection technique of data reduction with suitable example. 8

3. (a) Describe multidimensional data model. 8

(b) Explain the usage of data warehousing for information processing. 6

OR

4. (a) Describe a three-tier data warehouse architecture. 8

(b) Explain multiway array aggregation for full cube computation. 6

5. (a) Describe "Market Basket Analysis" w.r.t. frequent itemset mining. 6

(b) Describe the method of multilabel association rules mining with example. 7

OR

6. (a) Explain support-confidence framework for mining association rules. 6

(b) Describe the need of correlation analysis in support of association rules mining. 7

7. (a) Explain the following :—

(i) Network Topology.

(ii) Backpropagation. 6

(b) Describe measures for computing classifiers accuracy. 7

OR

8. (a) Explain the following :—

(i) Linear Regression.

(ii) Non-linear Regration. 6

(b) Describe measures for computing predicator errors. 7

9. (a) What is Cluster Analysis ? What types of data usually occur in cluster analysis ? 6

(b) Explain the need of data reduction and transformation techniques in high-dimensional time-series data. 7

OR

10. (a) Explain the purpose of following :—

(i) Partitioning methods.

(ii) Hierarchical methods. 6

(b) Explain K-Medoids with example. 7