Reg. No. | | | | | | | | | | | |

BCACAEN 601

Sixth Semester B.C.A. Degree Examination, June/July 2024
(NEP 2020) (2023 – 24 Batch Onwards)
FUNDAMENTALS OF DATA SCIENCE

Time : 2 Hours Max. Marks : 60

Note : *Answer* ***any six*** *questions from Part – **A** and* ***one full*** *question from* ***each*** *Unit in Part – **B**.*

PART – A

(6×2=12)

1. a) What is Data Mining ?
   b) What is supervised learning and unsupervised learning ?
   c) Define concept hierarchy. Give an example.
   d) What is Data Mart ? List its types.
   e) What are frequent patterns ? What is the use of frequent pattern mining ?
   f) What are categorical attributes and quantitative attributes ?
   g) What is Decision Tree ?
   h) Expand CLARA and ROCK.

PART – B

Unit – I

2. a) Define KDD process. Explain the different stages of KDD.
   b) Explain various application areas of Data Mining. **(6+6)**

3. a) Explain the following Data Mining techniques :
      i) Verification model
      ii) Discovery model.
   b) List and explain the issues and challenges in Data Mining. **(6+6)**

P.T.O.

Unit – II

4. a) Define and explain data warehouse.
   b) Explain any four OLAP operations.
   c) Explain any two numerosity reduction techniques. **(4+4+4)**

5. a) Define Measure. Explain different categories of measures.
   b) Explain any four ways of handling missing values.
   c) Explain the different steps involved in data transformation. **(4+4+4)**

Unit – III

6. a) Explain the classification of frequent pattern mining.
   b) Explain any three pruning strategies of mining closed frequent item sets. **(6+6)**

7. a) Explain Apriori algorithm.
   b) What is constrain based mining ? Explain the constraints included in constrain based mining. **(6+6)**

Unit – IV

8. a) Explain any four criteria which are used for comparing classification and prediction methods.
   b) Explain IF-THEN rules for classification.
   c) Explain DBSCAN and OPTICS. **(4+4+4)**

9. a) Explain any four requirements of clustering in data mining.
   b) Explain K-means Partition Algorithm.
   c) Explain four cases of cost function for K-mediod clustering. **(4+4+4)**