

Note: All questions are compulsory.

Diagrams are mandatory.

Use of calculator is allowed.

Q.1 Attempt any three of the following:

15 M

- What is data warehousing? Explain the need of data warehousing.
- Define the features of data warehousing.
- How to implement hardware and software parallel processing?
- Explain the components of the Data warehouse.
- What is metadata and why is it important?
- What is data mart? Explain the difference between data warehouse and data marts.

Q.2 Attempt any three of the following:

15 M

- What is Dimension modeling? Explain benefits of dimension modeling.
- Define Star Schema. what is inside the dimension table.
- Write note on:
 - Primary Keys
 - Surrogate Keys
 - Foreign Keys
- What is a snowflake schema? Explain its advantages and disadvantages.
- Define the basic steps of data transformation.
- What do you mean by data loading? Also explain four modes of it.

Q.3 Attempt any three of the following:

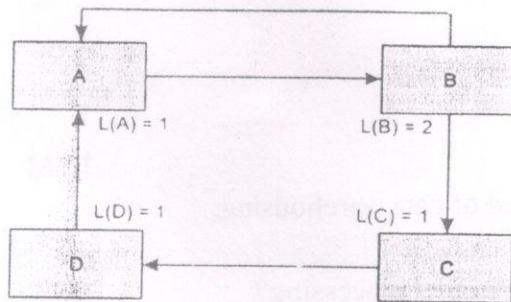
15 M

- Explain the concept of data mining with its applications.
- Explain the process of Data Mining.
- Differentiate between Data Mining and Machine Learning.
- Define classification with the working of classification.
- What is data preprocessing? Also explain its methods.
- What is entropy explained in brief?

Q.4 Attempt any three of the following:

15 M

- Explain the applications of cluster analysis.
- Explain the working of HITS Algorithm.
- What is K-means clustering? Explain working of the K-means algorithm.
- Calculate page rank of below web pages. Assume initial damping factor (d) is 0.7.



- e. Explain the concept of Web Content mining.
- f. Explain the working of search engines.

Q.5 Attempt any three of the following:

15 M

- a. Explain the concept of Association Rule Mining.
- b. Explain the working of Naive algorithms.
- c. Write note on:
 - I. Simple Storage
 - Ii. Horizontal Storage
 - Iii. Vertical Storage
- d. Explain the Apriori algorithm.
- e. What is the FP-Growth algorithm?
- f. Consider the sales database given below.

TID	ITEMS
1	Bread, Milk
2	Bread, Egg, Milk
3	Jam, Bread, Eggs
4	Eggs, Jam, Milk, Bread

Calculate:

- I. Support(Bread)
- Ii: Support(Bread, Eggs)
- Iii. Confidence(Bread, Milk)
- Iv. Confidence(Jam, Bread)
- V. Support(Jam).