

1/12/23

Time: 3 hours

- N.B. (1) Question one is Compulsory.  
(2) Attempt any 3 questions out of the remaining.  
(3) Assume suitable data if required.

- Q. 1 a) Explain features of Datawarehouse. 05
- b) What is Data Preprocessing? Explain the different methods for the Data integration phase 05
- c) What is hierarchical clustering? Explain divisive clustering 05
- d) Define Metadata and explain the types of metadata 05
- Q. 2 a) Explain association rule mining and multilevel association rules giving example of multidimensional association rules 10
- b) Give Data mining as a step in KDD. Give the architecture of typical Data Mining system 10
- Q. 3 a) Explain Extraction and transformation in ETL process. 10
- b) Illustrate Multidimensional association rules with suitable examples 10
- Q. 4 a) Define classification, issues of classification and explain Naïve bayesian classification with example 10
- b) Find the mean, median, mode, midrange, variance of data 13,15,16,16,19,20,20,21,22,25,26,26,26,30,33,36,40,45,46,52,52,70 10
- Q. 5 a) Explain HITS algorithm and illustrate its working 10
- b) Explain Web structure mining in detail 10

Q. 6 a) What is clustering? Explain k-means clustering algorithm. Suppose the data is {2,4,10,12,3,20,11,25} Consider k=2, cluster the given data using above algorithm 10

b) Illustrate with various operations and examples of OLAP cube 10

\*\*\*\*\*