

BE Semester- 5th (IT) Question Bank

(Data warehousing & Mining)

All questions carry equal marks(10 marks)

Q.1	Draw and explain the architecture of data warehousing.
Q.2	Explain the metadata in the data warehouse.
Q.3	Describe the need for data warehousing.
Q.4	Draw and explain the different schemas of data warehousing.
Q.5	Give the difference between OLTP and OLAP.
Q.6	Explain the different OLAP servers.
Q.7	What is cuboid? Explain various OLAP operations on Data Cube with example.
Q.8	Draw and explain the KDD process.
Q.9	Explain the data mining functionalities.
Q.10	Explain the major issues in data mining.
Q.11	Explain the architecture of a typical data mining system.
Q.12	Explain different types of data on which mining can be performed.
Q.13	Explain the data discretization and data summarization with an example.
Q.14	Explain data cleaning, data integration and transformation in detail.
Q.15	What is noise? Describe the possible reasons for noisy data. Explain the different techniques to remove the noise from data.
Q.16	List and describes the methods for handling the missing values in data cleaning.
Q.17	Write an algorithm for finding frequent item-sets using candidate generation.
Q.18	Describe the list of techniques for improving the efficiency of Apriori-based mining.
Q.19	What is market basket analysis? Explain association rules with confidence and support.
Q.20	Explain the issue regarding the classification and prediction.
Q.21	Explain the different attribute selection measures.
Q.22	Explain the rule based classification and case based reasoning in details.
Q.23	Explain rule based classification by decision tree induction algorithm.
Q.24	Explain the typical requirement of the clustering.
Q.25	Explain k-means and k-medoids algorithms of clustering.
Q.26	Write short note on: 1) Web content mining 2) Web usage mining
Q.27	Write short note on: Spatial mining.
Q.28	What are the challenges for effective resources and knowledge discovery in

	mining the world wide web?
Q.29	Explain the trends in data mining.
Q.30	Explain current application on KDD.
Q.31	Explain why data pre-processing required and explain any one technique in details.
Q.32	Explain support and confidence in association mining with suitable example.
Q.33	Explain hash based Apriori algorithm.
Q.34	Explain difference between partitions based Apriori and Apriori algorithm.
Q.35	Explain regression techniques.
Q.36	Give the difference between K-mean and K-medoid algorithm.
Q.37	What are the goals of a data warehouse?
Q.38	What are the advantages of using a data warehouse?
Q.39	How is attribute – oriented induction performed?
Q.40	List out the advantages and disadvantages of a decision tree.