



Register Number:

Date: 17-11-2020

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE -27
MSC Computer Science -III SEMESTER
END SEMESTER EXAMINATION: NOVEMBER 2020
CS 9118 –DATAWAREHOUSING AND DATAMINING

Time – 2.5 hours

Max Marks-70

I. Answer any SEVEN of the following.

7X10=70

1. a) Define datamart and enterprise datawarehouse. [2]
b) Explain the threetier architecture of datawarehouse with a neat diagram. [8]
2. Design a star schema, snowflake schema and fact constellation schema for the datawarehouse that consists of the following four dimensions (Time, Item, Branch and Location). Include the appropriate measures required for the schemas. [10]
3. a) Explain multi dimensional data model. Give an example. [5]
b) Explain various OLAP operations with example. [5]
4. a) What is factless fact table? Give an example. [3]
b) What is data discretization? Give an example. [3]
c) Explain the steps in ETL process. [4]
5. a) Write the difference between designing a data warehouse and an OLAP cube. [5]
b) What is metadata in datawarehouse? Explain the types of metadata. [5]
6. a) Outline the need for data cleaning. [2]
b) Consider the following data for the attribute AGE:
4,8,15,21,21,24,34,28,25. Perform smoothing by bin means and bin boundaries using a bin size of 3. [5]
c) Why we need data transformation? Mention the ways by which data can be transformed. [3]
7. a) What is decision tree? [2]
b) Elucidate the steps in the decision tree classification process with an example. [8]

8. a)What are the different types of data in cluster analysis? [2]
b)Describe the different hierarchical methods in cluster analysis. [8]
- 9.a) Discuss the application of data mining in risk analysis. [5]
b)Explain how multilayer feed forward network is used to predict the class labels of tuples. [5]
