

Roll No .....

**CS-8203 (GS)**

**B.E. VIII Semester**

Examination, November 2019

**Grading System (GS)**

**Data Mining and Knowledge Discovery**

*Time : Three Hours*

**Maximum Marks : 70**

**Note:** i) Attempt any five questions.

ii) All questions carry equal marks.

1. What are the decision trees? How these are useful in classifying data? What are its pros and cons as compare to other classification techniques?

2. Discuss the following

i) Star Schema

ii) Snowflake Schema

State which schema is better and why?

3. Write short notes on Two of the following:

i) Gini index

ii) Naive based method

iii) Density based method

4. Discuss in detail the application of data mining for financial data analysis? Give suitable data flow diagram.

5. What do you mean by data reduction techniques? Discuss attribute subset selection method with the help of suitable example.

6. Explain Pincer Search algorithm in detail. Demonstrate the working of this algorithm by taking suitable example dataset.

7. a) Discuss about prediction method using cluster analysis with an example.

b) Explain split algorithm based on the Gini index.

8. Write short notes on Two of the following:

i) Metadata Repository

ii) Confusion Matrix

iii) Data Cleaning

\*\*\*\*\*