



End Term Examination (December 2019)

School: School of Engineering

Program: B.Tech Computer Engineering

Course: Data Mining

Course Code: CSC335

Semester: V

MaxMarks: 40

Duration (mins): 90

Answer any FOUR questions. Each Question carries 10 marks.

[40]

Assume Data wherever necessary. Mention the assumptions wherever required.

Draw Diagram / Prototype for each. Mention example wherever necessary.

Q 1. What are the Different phases or process of Data Mining?

Explain with detailed diagram and example of each phase.

Q 2. Explain Predictive Analysis / Modeling using any Data mining Technique or Method

Using example of supervised, unsupervised and semi supervised Learning.

Q3. Discuss any Data mining Tool with its components and benefits to use to add on to any given problem statement.

Q4. Explain the difference between Data Cleaning, Data Pre-processing, and Data Transformation with suitable examples. How 1NF and 2NF normalization can be used to for them. Mention certain Algorithm / Algorithms or methods to handle missing values.

Q5. Which method out of the given two is more appropriate for handling data selection from large data bases (mining for large data):

a. Cartesian product

b. Sub query

Explain with suitable query to illustrate the example for proper justification.
