B.Sc. Computer Science

SEMESTER – VI

Course Code	Course Title	H	C	I	E	T
17U6DMC16	Data Mining	5	5	25	75	100

Objectives:

- To impart the knowledge of Data warehousing and Data Mining.
- Learning the concepts of various data mining methods and techniques.

Unit - I **Total Hours: 75**

Introduction & Data Preprocessing **(15 Hours)**

Introduction to Data Mining – Data preprocessing: An overview - Data Cleaning - Data Integration - Data reduction – Data transformation and Data discretization.

Unit - II

Data warehousing and online analytical processing **(15 Hours)**

Data warehouse: Basic concepts – Data Warehouse modeling: Data Cube and OLAP – Data Warehouse Implementation – Data generalization by attribute-oriented induction.

Unit - III

Mining Frequent, Associations and correlations **(15 Hours)**

Basic concepts - Frequent Itemset Mining methods – Advanced Pattern mining: A road map - Pattern mining in multilevel, multidimensional space – Constraint-based frequent pattern mining.

Unit - IV

Classification **(15 Hours)**

Basic concepts - Decision tree Induction- Bayes classification methods- Rule based classification – Classification: Advanced methods - Classification by Back propagation.

Unit - V

Cluster analysis **(15 Hours)**

asic Concepts and methods: Cluster analysis- Partitioning methods - Hierarchical methods – Density based methods – Data Mining Trends and Research Frontiers: Data Mining Applications.

Text Book:

"Data Mining concepts and Techniques" – Jiawei Han, Micheline Kamber, Jian Pei -Third Edition - Morgan Kaufmann Publishers, New Delhi.

Chapters:

Unit – I: 1.1, 1.2, 3.1, 3.2, 3.3, 3.4, 3.5.
Unit – II: 4.1, 4.2, 4.4, 4.5
Unit – III: 6.1, 6.2, 7.1, 7.2, 7.3
Unit – IV: 8.1, 8.2, 8.3, 8.4, 9.2
Unit – V: 10.1, 10.2, 10.3, 10.4, 13.3.

Reference Books:

1. "Data Mining Introductory and Advanced topics" - Margaret Dunham –Prentice Hall 2003.
2. "Principles of Data Mining" - Heikki Mannila and Padhraic Smyth – MIT Press Fall 2000.