CSH506: Data Mining

Teaching Scheme	Examination Scheme
Lectures: 1 hrs/Week Tutorials: 1 hr/Week Credits: 2	Class Test -6 Marks Teachers Assessment – 3 Marks Attendance – 6 Marks End Semester Exam – 35 marks

Prerequisite: - CSH301 RDBMS

Course Objectives:

1. To understand Data Mining, its origin and applications.
2. To understand types of data and to improve the quality of data and efficiency and the ease of the mining process.
3. Differentiate OnLine Transaction Processing and OnLine Analytical processing
4. Learn Multidimensional schemas suitable for data warehousing along with OLAP operations.
5. To understand how to identify associations among objects and to learn various algorithms to find them.
6. To understand applications and algorithms for Clustering along with methodologies of data mining.

Detailed Syllabus

UNIT 1- Data Mining:
Definition, Data Mining as the Evolution of Information Technology, Knowledge Discovery Process (KDP), Classification of Mining systems, Techniques involved.

UNIT 2- Data Preprocessing:
Needs, Pre-processing data, Data Cleaning, Data integration and transformation, data reduction, discretization, Concept of hierarchy generation.

UNIT 3- Data Warehouse:
Definition, Differences between Operational Database Systems and Data Warehouses, OLTP vs. OLAP, 3 Tier Architecture of Data Warehouse, Concept of ETL.

UNIT 4- Data Warehouse Modeling:
Data Cube- A Multidimensional Data Model, Stars, Snowflakes, and Fact Constellations: Schemas for Multidimensional Data Models, OLAP operation

UNIT 5- Data Mining Techniques:
Introduction to Association Rule and Association Rule Mining, Classification: Decision Tree Induction, K-nearest neighbor, Clustering: Cluster Analysis.

UNIT 6- Data Mining Trends:
Mining Complex Data Types, Methodologies of Data Mining, Data Mining Applications, Web Mining.

Text and Reference Books

1. Data Mining -Concepts and Techniques, Han, Kamber, Harcourt India, 2006.
2. Data Mining Introductory and advanced topics, Margaret H Dunham, Pearson, 2002.
3. Data Mining Techniques, Arjun K. Pujari, University Press, 2001.

Head
Department of Computer Applications
Faculty of Computer Applications
Invertis University, Bareilly (UP)

Registrar
Invertis University
Bareilly

Dean Academics
Faculty of Computer Applications
Invertis University, Bareilly (UP)

Course Outcomes:

After completing the course, students will be able to

1. Understand the concept of data mining and its applications.
2. Understand pre-processing steps to improve the quality of data to ease data mining process.
3. Understand OLTP and OLAP as well as 3 tier architecture of data warehouse.
4. Understand various Multidimensional schemas and to apply OLAP operations.
5. Establish associations among objects by applying various algorithms.
6. Perform cluster analysis and understand the methodologies of data mining.