MC5403 ADVANCED DATABASE AND DATA MINING

DETAILED SYLLABUS

OBJECTIVES:

- To understand the underlying principles of Relational Database Management System.
- To Understand Data mining principles and techniques and Introduce DM as a cutting-Edge business intelligence
- To learn to use association rule mining for handling large data
- To understand the concept of classification for the retrieval purposes
- To know the clustering techniques in details for better organization and retrieval of data
- To identify Business applications and Trends of Data mining

UNIT I RELATIONAL MODEL

Data Model – Types of Data Models: – Entity Relationship Model – Relational Data Model – Mapping Entity Relationship Model to Relational Model – Structured Query Language – Database Normalization – Transaction Management.

UNIT II DATA MINING & DATA PREPROCESSING

Introduction to KDD process – Knowledge Discovery from Databases - Need for Data Pre-processing – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization and Concept Hierarchy Generation.

UNIT III ASSOCIATION RULE MINING

Introduction - Data Mining Functionalities - Association Rule Mining - Mining Frequent Itemsets with and without Candidate Generation - Mining Various Kinds of Association Rules - Constraint-Based Association Mining.

UNIT IV CLASSIFICATION & PREDICTION

Classification vs. Prediction – Data preparation for Classification and Prediction – Classification by Decision Tree Introduction – Bayesian Classification – Rule Based Classification – Classification by Back Propagation – Support Vector Machines – Associative Classification – Lazy Learners – Other Classification Methods – Prediction – Accuracy and Error Measures.

UNIT V CLUSTERING

Cluster Analysis: - Types of Data in Cluster Analysis – A Categorization of Major Clustering Methods – Partitioning Methods – Hierarchical methods – Density-Based Methods – Grid Based Methods – Model-Based Clustering Methods – Clustering High- Dimensional Data – Constraint-Based Cluster Analysis – Outlier Analysis.

REFERENCES:

1. Berson, Alex & Smith, Stephen J, Data Warehousing, Data Mining, and OLAP, TMH Pub. Co. Ltd, New Delhi, 2012

2. G. K. Gupta, "Introduction to Data Mining with Case Studies", Easter Economy Edition, PrenticeHall of India, 2006

3. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques" Second Edition, Elsevier, Reprinted 2008

4. K.P. Soman, ShyamDiwakar and V. Ajay, "Insight into Data mining Theory and Practice", Easter Economy Edition, Prentice Hall of India, 2006

5. Marakas, George M, Modern Data Warehousing, Mining, and Visualiza Visualization, Pearson Education, 2011

6. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Pearson Education, 2007

7. Prabhu Data Warehousing, PHI Learning Private Limited, New Delhi, 2012.

8. Ponniah, Paulraj, Data Warehousing Fundamentals, John Wiley & Sons, New Delhi, 2011.

9. R. Elmasri, S.B. Navathe, "Fundamentals of Database Systems", Addison-Wesley, 2011