IT6702 **DATA WAREHOUSING AND DATA MINING** **L T P C**
3 0 0 3

OBJECTIVES:

The student should be made to:

- Be familiar with the concepts of data warehouse and data mining,
- Be acquainted with the tools and techniques used for Knowledge Discovery in Databases.

UNIT I DATA WAREHOUSING **9**

Data warehousing Components –Building a Data warehouse — Mapping the Data Warehouse to a Multiprocessor Architecture – DBMS Schemas for Decision Support – Data Extraction, Cleanup, and Transformation Tools –Metadata.

UNIT II BUSINESS ANALYSIS **9**

Reporting and Query tools and Applications – Tool Categories – The Need for Applications – Cognos Impromptu – Online Analytical Processing (OLAP) – Need – Multidimensional Data Model – OLAP Guidelines – Multidimensional versus Multirelational OLAP – Categories of Tools – OLAP Tools and the Internet.

UNIT III DATA MINING **9**

Introduction – Data – Types of Data – Data Mining Functionalities – Interestingness of Patterns – Classification of Data Mining Systems – Data Mining Task Primitives – Integration of a Data Mining System with a Data Warehouse – Issues –Data Preprocessing.

UNIT IV ASSOCIATION RULE MINING AND CLASSIFICATION **9**

Mining Frequent Patterns, Associations and Correlations – Mining Methods – Mining various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining – Classification and Prediction - Basic Concepts - Decision Tree Induction - Bayesian Classification – Rule Based Classification – Classification by Back propagation – Support Vector Machines – Associative Classification – Lazy Learners – Other Classification Methods – Prediction.

UNIT V CLUSTERING AND TRENDS IN DATA MINING **9**

Cluster Analysis - Types of Data – Categorization of Major Clustering Methods – K-means– Partitioning Methods – Hierarchical Methods - Density-Based Methods –Grid Based Methods – Model-Based Clustering Methods – Clustering High Dimensional Data - Constraint – Based Cluster Analysis – Outlier Analysis – Data Mining Applications.

TOTAL: 45 PERIODS

OUTCOMES:

After completing this course, the student will be able to:

- Apply data mining techniques and methods to large data sets.
- Use data mining tools
- Compare and contrast the various classifiers.

TEXT BOOKS:

1. Alex Berson and Stephen J.Smith, "Data Warehousing, Data Mining and OLAP", Tata McGraw – Hill Edition, Thirteenth Reprint 2008
2. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Third Edition, Elsevier, 2012.

REFERENCES:

1. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Person Education, 2007.
2. K.P. Soman, Shyam Diwakar and V. Aja, "Insight into Data Mining Theory and Practice", Eastern Economy Edition, Prentice Hall of India, 2006.
3. G. K. Gupta, "Introduction to Data Mining with Case Studies", Eastern Economy Edition, Prentice Hall of India, 2006.
4. Daniel T.Larose, "Data Mining Methods and Models", Wiley-Interscience, 2006.