DATA WAREHOUSING AND DATA MINING

(Common for CSE & IT)

Course Code : 15CT1132

L	T	P	C
3	0	0	3

Pre-requisites:

Database Management Systems

Course Outcomes:

At the end of the Course, the Student will be able to:

CO 1 Apply data pre-processing techniques

CO 2 Design data warehouse schema.

CO 3 Discover associations and correlations in given data.

CO 4 Apply classification techniques.

CO 5 Apply clustering techniques.

UNIT-I **(10 Lectures)**

INTRODUCTION:

Data mining-On what kinds of Data, what kinds of patterns can be mined, which kinds of applications are targeted, major issues in Data Mining. DATA PREPROCESSING: An Overview, Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data discretization.

UNIT-II **(10 Lectures)**

DATA WAREHOUSE AND OLAP TECHNOLOGY:

Data Warehouse: Basic concepts, Data Warehouse Modeling: Data Cube and OLAP, Data Warehouse Implementation.

UNIT-III (10 Lectures)

MINING FREQUENT PATTERNS, ASSOCIATION AND CORRELATIONS:

Basic Concepts, Efficient and Scalable Frequent Item set Mining Methods. Mining various kinds of association rules.

UNIT-IV (10 Lectures)

CLASSIFICATION: BASIC CONCEPTS:

What is Classification? What is Prediction? Issues regarding Classification and Prediction. Classification by Decision Tree Induction, Bayesian Classification-Baye's Theorem, Naive Bayesian Classification, Rule-Based Classification-Using IF-THEN Rules for Classification, Rule Extraction from a Decision Tree, Classification by Back propagation.

UNIT-V (10 Lectures)

CLUSTER ANALYSIS: BASIC CONCEPTS:

Cluster analysis, A Categorization of Major Clustering Methods-PartitioningMethods (k-Means, k-Medoids), Hierarchical Methods: Agglomerative Vs Divisive (BIRCH), Density Based Methods (DBSCAN), Grid-Based Methods

(STING). Model-Based Clustering Methods (Expectation-Maximization

TEXT BOOKS:

1. Jlawei Han &MichelineKamber, Data Mining, "Concepts and Techniques", 3rd Edition, Morgan Kaufmann Publishers, 2011.
2. Margaret H Dunham, "Data Mining Introductory and advanced topics", 6thEdition, Pearson Education, 2009.

REFERENCES:

1. Arun K Pujari, "Data Mining Techniques",1st Edition, University Press, 2005.

2. Pang- Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining", 1st Edition, Pearson Education,2012.
3. Sam Aanhory& Dennis Murray, "Data Warehousing in the Real World",1st Edition, Pearson Edn Asia,2008.
4. PaulrajPonnaiah, "Data Warehousing Fundamentals",1st Edition, Wiley student Edition,2007.
5. Ralph Kimball, "The Data Warehouse Life cycle Toolkit",2nd Edition, Wiley student Edition,2005.