MALLA REDDY UNIVERSITY

II Year B Tech CBCS POOL

L/T/P/C 3/0/0/3

(MR22-1CS0148) DATA MINING

COURSE OBJECTIVES:

- 1. To learn data mining concepts
- 2. To understand association rules mining.
- 3. Discuss classification algorithms
- 4. To learn how data is grouped using clustering techniques.

UNIT-I:

Introduction: Fundamentals of data mining, Data Mining Functionalities, Classification of Data Mining systems, Data Mining Task Primitives, Integration of a Data Mining System with a Database or Data Warehouse System, Major issues in Data Mining.

UNIT-II:

Data Preprocessing- Overview, Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization.

UNIT-III:

Data Warehousing and Online Analytical Processing-Data Warehouse: Basic Concepts, Data Warehouse Modeling: Data Cube and OLAP, Data Warehouse Design and Usage, Data Warehouse Implementation, Data Generalization by Attribute-Oriented Induction.

UNIT-IV:

Classification- Basic Concepts, Decision Tree Induction, Bayes Classification Methods: Bayes' Theorem, Na¨ıve Bayesian Classification, Rule-Based Classification: Using IF-THEN Rules for Classification, Rule Extraction from a Decision Tree, Rule Induction Using a Sequential Covering Algorithm, Model Evaluation and Selection, Techniques to Improve Classification Accuracy

UNIT-V:

Cluster Analysis-Cluster Analysis, Partitioning Methods: k-Means, k-Medoids, Hierarchical Methods.

TEXT BOOKS:

 Data Mining- Concepts and -Te3chniques- Jiawei Han, Micheline Kamber, Morgan Kaufmann Publishers, Elsevier, 3rd Edition, 2012.

MALLA REDDY UNIVERSITY (MRU)

REFERENCE BOOKS:

- 1. Data Mining Techniques, Arun K Pujari, 3rd Edition, Universities Press.
- 2. Data Mining, Vikaram Pudi, P Rddha Krishna, Oxford University Press.

COURSE OUTCOMES:

- 1. Ability to perform the pre-processing of data and apply mining techniques on it.
- 2. Ability to identify the association rules, classification and clustering large data sets.
- 3. Ability to solve real world problem s in business and scientific information using data mining.