BIT 357

DATA WAREHOUSING & DATA MINING

Unit – I

Introduction: What is Data Mining, Data Mining Functionalities, Classification of Data Mining Systems, Major Issues in Data Mining.

Data Preprocessing: Preprocessing, Descriptive Data Summarization, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation.

Unit – II

Data Warehouse and OLAP Technology: What is Data Warehouse, A Multidimensional Data Model, Data Warehouse Architecture and Implementation, from Data Warehousing to Data Mining.

Mining Frequent Patterns, Associations Rules: Basic Concepts, Efficient and Scalable Frequent Item Set Mining Methods, Mining Various kinds of Association Rules.

Unit – III

Classification and Prediction: Introduction, Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, Rule based Classification, Classification by Back Propagation, Support Vector Machines, Associative classification, Other classification Methods, Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor.

Unit – IV

Cluster Analysis: Introduction, 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, Outlier Analysis.

Mining Streams, Time-Series, and Sequence Data: Mining Data Streams, Mining Time-Series Data, and Mining Sequence Patterns in Transactional Databases and Biological Data:

SCETW

Unit – V

Mining Object, Spatial, Multimedia, Text, and Web Data: Multidimensional Analysis and Descriptive Mining of Complex Data Objects, Spatial Data Mining, Multimedia Data Mining, Text Mining, Mining the World Wide Web.

Suggested Reading:

1. Han J & Kamber M, Data Mining: Concepts and Techniques, Third Edition, Elsiver, 2011.
2. Pang-Ning Tan, Michael Steinback, Vipin Kumar, Introduction to Data Mining, Pearson Education, 2008.
3. Arun K Pujari, Data mining Techniques, Second Edition, University Press,2001.
4. Margaret H Dunham, S.Sridhar, Data mining: Introductory and Advanced Topics, Pearson Education, 2008.
5. Humphires, Hawkins, Dy, Data Warehousing: Architecture and Implementation, Pearson Education, 2009.
6. Anahory, Murray, Data Warehousing in the Real World, Pearson Education, 2008.
7. Kargupta, Joshi,etc., Data Mining: Next Generation Challenges and Future Directions, Prentice Hall of India Pvt Ltd, 2007.

SCETW