

Classification by Decision tree induction, Bayesian classification, Rule based classification, backpropagation through Neural Networks, Genetic Algorithm, Support Vector Machines, Prediction: linear and non-linear regression techniques.

MODULE-4: MINING TIME SERIES DATA AND DATA STREAMS

Mining Time series Data, Periodicity Analysis for time related sequence data, Similarity search in Time-series analysis; Mining Data Streams, Methodologies for stream data processing and stream data systems, Frequent pattern mining in stream data, Classification of dynamic data streams.

MODULE-5: ADVANCED MINING APPLICATIONS

Web Mining, Web page layout structure; mining web link structure, content and usage patterns; Recent trends in Distributed Warehousing and Data Mining, Class Imbalance Problem; Graph Mining; Social Network Analysis

Course Outcomes:

1. The students will be able to understand basic concepts of data warehouse and data mining, techniques and applications
2. The students will be able to understand the techniques to extract patterns from transactional database using Association and Apriori algorithms
3. The students will be able to understand different clustering techniques and will be able to cluster data sets
4. The students will be able to classify data set into different classes and acquire the knowledge to make predications based on classified data
5. The students will be able to understand and analyze time series data
6. The students will be able to understand types of web mining viz. content, structure and usage mining. Web content mining in detail.
7. The students can extend the Graph mining algorithms to Web mining
8. Students will understand advance applications of data mining

REFERENCES

CODE: PEC-CS-D-601

SUBJECT NAME: DATA MINING

NO OF CREDITS: 3

B.TECH 6th SEMESTER

L T P

3 0 0

SESSIONAL: 25

THEORY EXAM: 75

TOTAL : 100

Pre-requisites: Database Management System

Course Objectives:

1. To familiarize the students with the basic roadmap of data mining and various data mining techniques.
2. To introduce the techniques of frequent pattern mining and Clustering
3. To acquaint students with classification and prediction techniques in data mining.
4. To introduce students with time series data and data streams
5. To introduce various advance mining applications areas like web mining, social network analysis etc.

MODULE-1: INTRODUCTION

Introduction to Data Warehousing, Architecture, Data warehouse schemas, OLAP operations, KDD process, Data Mining: Predictive and Descriptive models, Data Mining primitives and Applications

MODULE-2: FREQUENT PATTERN MINING AND CLUSTERING

Mining frequent patterns, association and correlations; Association Rule Mining, Sequential Pattern Mining concepts, Cluster Analysis – Types of Data in Cluster Analysis, Partitioning methods, Hierarchical Methods; Transactional Patterns and other temporal based frequent patterns,

MODULE-3: CLASSIFICATION AND PREDICTION

