RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

Artificial Intelligence and Data Science, VI-Semester

Departmental Elective – AD 603 (A) Data Mining and Warehousing

COURSE OBJECTIVES

1. Student should understand the value of Historical data and data mining in solving real-worldproblems.
2. Student should become affluent with the basic Supervised and unsupervised learning algorithmscommonly used in data mining.
3. Student develops the skill in using data mining for solving real-world problems.

Unit-I: Data Warehousing: Introduction, Delivery Process, Data warehouse Architecture, DataPre-processing: Data cleaning, Data Integration and transformation, Data reduction. Data warehouseDesign: Data warehouse schema, Partitioning strategy Data warehouse Implementation, Data Marts,Meta Data, Example of a Multidimensional Data model. Introduction to Pattern Warehousing.

Unit-II: OLAP Systems: Basic concepts, OLAP queries, Types of OLAP servers, OLAP operations etc.Data Warehouse Hardware and Operational Design: Security, Backup And Recovery.

Unit-III: Introduction to Data& Data Mining :Data Types, Quality of data, Data Pre-processing, Similaritymeasures, Summary statistics, Data distributions, Basic data mining tasks, Data Mining V/sknowledge discovery in databases. Issues in Data mining. Introduction to Fuzzy sets and fuzzylogic.

Unit-IV: Supervised Learning: Classification: Statistical-based algorithms, Distance-based algorithms,Decision tree-based algorithms, Neural network-based algorithms, Rule-based algorithms,Probabilistic Classifiers

Unit-V: Clustering & Association Rule mining: Hierarchical algorithms, Partitional algorithms, Clusteringlarge databases – BIRCH, DBSCAN, CURE algorithms. Associationrules: Parallel and distributedalgorithms such as Apriori and FP growth algorithms.

Books Recommended:

Text Books:

1. Pang – ningTan , Steinbach & Kumar, “Introduction to Data Mining”, Pearson Edu, 2019.
2. Jaiwei Han, MichelineKamber, “Data Mining : Concepts and Techniques”, Morgan KaufmannPublishers.

Reference Books:

1. Margaret H. Dunham, “Data Mining : Introductory and Advanced topics”, Pearson Edu., 2009.

2. Anahory& Murray, “Data Warehousing in the Real World”, Pearson Edu., 2009.

COURSE OUTCOMES

After completion of this course, the students would be able to:

CO1. Understand the need of designing Enterprise data warehouses and will be enabled to approachbusiness problems analytically by identifying opportunities to derive business.

CO2. Compare and contrast, various methods for storing & retrieving data from different data Sources/repository.

CO3. Ascertain the application of data mining in various areas and Pre-process the given data andvisualize it for a given application or data exploration/mining task

CO4. Apply supervised learning methods to given data sets such as classification and its various types.

CO5. Apply Unsupervised learning methods to given data sets such as clustering and its various types.

CO6 Apply Association rule Mining to various domains