DATA WAREHOUSING & DATA MINING

Credits: 4 **Semester: IV**

Subject Code: DS20403 **No. of Lecture Hours: 60**

Objectives:

- To establish what Data Warehousing is and why Data Warehousing concepts and techniques are proving crucial for firms faced with growing banks of structured and unstructured data.
- To introduce the basic concepts and techniques of Data Mining.

Outcomes: The students will be able to

CO1: To understand the concepts of Data Warehousing and its importance

CO2: Analyze different Data Mining methods using algorithms

CO3: Explain the classification of data

CO4: Express clustering methods using algorithms.

CO5: Identify importance of Text Mining and related algorithms

UNIT –I

12hrs

1. Introduction to Data Warehousing 1
2. Fundamentals and Data Characteristics 1
3. Data Warehouse Components. Building Data Marts and Data Warehouse 2
4. ETL 2
5. Logical Data Modelling Schemes and Designs for Dimensional Modelling 2
6. OLAP 2

UNIT- II **12hrs**

1. Introduction: Motivation, Data Warehousing and Data Mining 1
2. Technologies 1
3. Data Models 1
4. Data Warehousing and OLAP: User's Perspective 1
5. Data Mining: User's Perspective, Related Disciplines 1
6. Other Issues, Future Trends. 1
7. Frequent Pattern Matching: Introduction, Basic Problem Definition 1
8. Mining Association Rules, Applications 1

9.	Variations, Interestingness, FIM algorithms	2
10.	Incremental Mining, Conciseness of Results, Sequential Rules	2

UNIT -III

12hrs

1.	Classification: Introduction, Basic Problem Definition, Applications	3
2.	Evaluation of Classifiers, Other Issues	2
3.	Classification Techniques	2
4.	Optimal Classification Algorithms	3
5.	Regression	2

UNIT-IV

12hrs

1.	Clustering: Introduction, Basic Problem Definition, Applications	3
2.	Measurement of Similarity	2
3.	Evaluation and Classification of Clustering Algorithms	2
4.	Partitioning Methods, Hierarchical Methods	2
5.	Density-Based Methods, Grid-Based Methods, Outlier Detection	3

UNIT-V

12hrs

1.	Introduction to Text Data Mining	1
2.	General and Functional Architecture	1
3.	Overview of Text Mining	2
4.	Core Mining Operations	2
5.	Pre Processing Techniques	2
6.	Document Collection, Tokenization, Lemmatization, Morphological Analysis	2
7.	Delimiters, Stop Words ,Key Word Search	2