Govt. Polytechnic Hamirpur (H.P.)

Lecture Planning (Theory)

Branch : **Information Technology**

Subject : **PE-4 (Data Science)**

Teacher : **Pankaj Thakur**

Semester : **5th**

Session : **Aug-Dec 2025**

Classroom : **CBLH3**

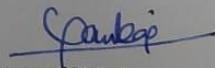| Sr. No.	No. of Lectures Planned	Chapter/ Unit Description	Detail of Contents	Reference/ Resources
1.	12	Introduction	Motivation, Importance, Definitions, Kind of Data, Data Mining Functionalities, Kinds of Patterns, Classification of Data Mining Systems, Data Mining Task Primitives, Integration of A Data Mining System with A Database or Data Warehouse System, Major Issues in Data Mining, Types of Data Sets and Attribute Values, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity. PREPROCESSING: Data Quality, Major Tasks in Data Preprocessing, Data Reduction, Data Transformation and Data Discretization, Data Cleaning and Data Integration	R1
2.	14	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, Data Cube Computation.	R1
3.	14	Patterns, Associations and Correlations	Mining Frequent Patterns, Associations and Correlations: Basic Concepts, Efficient and Scalable Frequent Itemset Mining Methods, Pattern Evaluation Methods, Applications of frequent pattern and associations. Frequent Patterns and Association Mining: A Road Map, Mining Various Kinds of Association Rules, Constraint-Based Frequent Pattern Mining, Extended Applications of Frequent Patterns.	R1
4.	12	Classification	Basic Concepts, Decision Tree Induction, Bayesian Classification Methods, Rule-Based Classification, Model Evaluation and Selection, Techniques to Improve Classification Accuracy: Ensemble Methods, Handling Different Kinds of Cases in Classification, Classification by Neural Networks, Support Vector Machines, Pattern-Based Classification, Lazy Learners (or Learning from Your Neighbors).	R1
5.	12	Cluster Analysis	Basic Concepts of Cluster Analysis, Clustering Structures, Major Clustering Approaches, Partitioning Methods, Hierarchical Methods, Density-Based Methods, Model-Based Clustering, Why outlier analysis, Identifying and handling of outliers, Outlier Detection Techniques. WEB MINING: Basic concepts of web mining, different types of web mining, PAGE RANK Algorithm, HITS Algorithm	R1

Course Outcomes:

After completing this course the students will be able to :

CO-1. Understand the terminology of data warehousing and data mining.

CO-2. Appreciate the strengths and limitations of various data warehousing models.

CO-3. Describe various methodologies used in data mining and data warehousing.

CO-4. Compare different approaches of data warehousing and data mining.

Teaching Resources:

R1: Data Mining: Concepts and Techniques by Jiawei Han et al., Elsevier

R2: Data Mining Introductory and Advanced Topics by M H Dunham, Pearson Edn.

R3: Data Warehousing by Amitesh Sinha, Thomson Learning, India

R4: Top Ten Algorithms in Data Mining by Xingdong Wu, CRC Press, UK.

01/08/2025
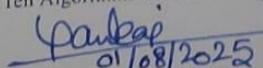
Signature of Teacher with Date

Signature of H.O.D.