

Data Mining

- 1.1 Course Number: CS458
- 1.2 Contact Hours: 2-0-2 Credits: 8
- 1.3 Semester-offered: 4th Year-Even
- 1.4 Prerequisite: Data Structure, Algorithms, Probability and Statistical Analysis, Database.
- 1.5 Syllabus Committee Member: Dr. Sushum Biswas, Dr. Daya Sagar Gupta & Dr. Gargi Srivastava
2. **Objective:** To give knowledge about fundamentals of data mining
3. **Course Content:**

Unit-wise distribution of content and number of lectures

Unit	Topics	Sub-topic	Lectures
1	Introduction	Data Mining, Motivation, Application, Data Mining—On What Kind of Data?, Data Mining, Functionalities, Data Mining Task Primitives, Major Issues in Data Mining	6
2	Data pre-processing	Descriptive Data Summarization, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation.	6
3	Association Rule	Mining Frequent Patterns, Associations, and Correlations: Basic Concepts and a Road Map, Association Rules, the Apriori Algorithm Classification and Prediction: Classification: Classification, Issues Regarding Classification, Classification by Decision Tree Induction, Bayesian Classification, Rule-Based Classification, Metrics for Evaluating Classifier Performance, Holdout Method and Random Sub sampling	17
4	Prediction	Prediction, Issues Regarding Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor.	6
5	Clustering	Cluster Analysis, Agglomerative versus Divisive Hierarchical Clustering, Distance Measures in Algorithmic, Evaluation of Clustering.	5
		Total	40

4. Readings

4.1 Textbook:

Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and

Techniques (Second Edition), Morgan Kaufmann, 2005, ISBN: 0-12-088407-0.

4.2 Reference books:

“Data Mining: Concepts and Techniques”, Second Edition Jiawei Han and Micheline Kamber.

- 5 Outcome of the Course:** After completion of this course, student will be able to
- 1) Understand the basic concepts of the information retrieval.
 - 2) Analyse the involvement of the information retrieval in modern life style & social media.
 - 3) Apply data pre-processing, indexing, retrieval methods and concepts.
 - 4) Evaluate the effectiveness and efficiency of different information retrieval systems