



## DATA MINING

B.Tech. IV Year I Semester								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
CS4102PC	Elective	L	T	P	C	CIA	SEE	Total
		2	0	0	2	25	75	100
Contact classes: 36	Tutorial Classes : NIL	Practical classes : NIL			Total Classes :36			
<b>Prerequisites:</b> <ul style="list-style-type: none"><li>• A course on “Database Management Systems”</li><li>• Knowledge of probability and statistics</li></ul>								

### Course Objectives:

- It presents methods for mining frequent patterns, associations, and correlations.
- It then describes methods for data classification and prediction, and data-clustering approaches.
- It covers mining various types of data stores such as spatial, textual, multimedia, streams.

### Course Outcomes:

- Ability to understand the types of the data to be mined and present a general classification of tasks and primitives to integrate a data mining system.
- Apply pre processing methods for any given raw data.
- Extract interesting patterns from large amounts of data.
- Discover the role played by datamining in various fields.
- Choose and employ suitable datamining algorithms to build analytical applications
- Evaluate the accuracy of supervised and unsupervised models and algorithms.

## COURSE SYLLABUS

### MODULE- I

**Data Mining:** Data-Types of Data-, Data Mining Functionalities- Interestingness Patterns- Classification of Data Mining systems- Datamining Task primitives-Integration of Datamining system with a Data warehouse-Major issues in Data Mining-Data Preprocessing.

---

## **MODULE- II**

**Association Rule Mining:** Mining Frequent Patterns–Associations and correlations –Mining Methods–Mining Various kinds of Association Rules–Correlation Analysis –Constraint based Association mining, Graph Pattern Mining, SPM.

## **MODULE- III**

**Classification:** Classification and Prediction– Basic concepts– Decision tree induction–Bayesian classification, Rule–based classification, Lazylearner.

## **MODULE- IV**

**Clustering and Applications:** Cluster analysis–Types of Data in Cluster Analysis–Categorization of Major Clustering Methods– Partitioning Methods, Hierarchical Methods– Density–Based Methods, Grid–Based Methods, Outlier Analysis.

## **MODULE- V**

**Advanced Concepts:** Basic concepts in Mining data streams–Mining Time–series data—Mining sequence patterns in Transactional databases– Mining Object– Spatial– Multimedia–Text and Webdata – Spatial Datamining–Multimedia Datamining–TextMining–Mining the World Wide Web.

## **TEXTBOOKS:**

1. Data Mining– Concepts and Techniques –Jiawei Han & Micheline Kamber, 3<sup>rd</sup> Edition Elsevier.
2. Data Mining Introductory and Advanced topics – Margaret H Dunham, PEA.

## **REFERENCE BOOK:**

1. Ian H.Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques (Second Edition), Morgan Kaufmann, 2005.