Course: BTech **Semester:** 6

Prerequisite: Database Management System, Linear algebra & Statistics

Rationale: This course helps the students to understand different data mining models and data visualization techniques.

Teaching and Examination Scheme

Teaching Scheme					Examination Scheme					Total
Lecture Hrs/Week	Tutorial Hrs/Week	Lab Hrs/Week	Hrs/Week	Credit	Internal Marks			External Marks		
					T	CE	P	T	P	
3	0	0	-	3	20	20	-	60	-	100

SEE - Semester End Examination, **CIA** - Continuous Internal Assessment (It consists of Assignments/Seminars/Presentations/MCQ Tests, etc.)

Course Content

W - Weightage (%) , **T** - Teaching hours

Sr.	Topics	W	T
1	**Introduction to data mining (DM)**Importance of Data Mining, Data Mining-Definition and Functionalities, Classification of Data mining systems, Data mining Architecture, KDD, DM task primitives, Major Issues in Data Mining, Data mining Technologies, Applications of Data Mining.	21	9
2	**Data Pre-processing**Tasks in Data Preprocessing, Reasons of Missing Values & Noisy Data, Cleaning, Integration, Reduction, Transformation and Discretization, Concept Hierarchy Generation.	16	7
3	**Mining Frequent Patterns**Efficient and scalable frequent itemset mining methods, Association Rules, Multidimensional & Multilevel association rules, Generating Association Rules from Frequent Item sets, Interesting Pattern Evaluation Methods.	16	7
4	**Classification & Clustering**Classification vs. prediction, Supervised learning, Approach to Classification: Decision Tree Induction, Unsupervised learning, Cluster Analysis: Partitioning Methods, Hierarchical Methods, Density-Based Methods, Evaluation of Clustering, Outlier Detection. Introduction to analytics tools like Power BI.	21	9
5	**Statistical Representation of Data**Data Quality, Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Histogram Analysis.	10	6
6	**Fundamental of Data Visualization**Introduction to data visualization & analytics, Info-graphic representation of terminologies, DIKW (Data, Information, Knowledge, wisdom) Pyramid, Difference between Analysis and Analytics, Applications of Data Visualization, Applications of Data Analytics	16	7

Reference Books

1.	**Data Mining concepts and Techniques By Jiawei Han, Micheline Kamber | Elsevier**
2.	**Data Mining Techniques By Arun K. Pujari | Universities Press**
3.	**Principles of Statistics** By M. G. Bulmer, Dover Publications Inc.
4.	**Beautiful Visualization** By Noah Iliinsky, Julie Steele Publisher(s): O'Reilly Media, Inc. ISBN: 9781449379865
5.	**Statistics 101: From Data Analysis and Predictive Modeling to Measuring Distribution and Determining Probability, Your Essential Guide to Statistics** By David Borman, Adams Media

Course Outcome

After Learning the Course the students shall be able to:

1. Extract knowledge using data mining techniques
2. Adapt to new data mining tools.
3. Apply the techniques of clustering, classification, association finding, feature selection and visualization to real world data
4. Analyze the dataset and perform Descriptive Statistics.
5. Analyze the dataset and perform an Inferential Statistics.