COURSE TITLE	**DATA MINING TECHNIQUES**
COURSE CODE	**05MD0204**
COURSE CREDITS	**5**

Course Outcomes: After completion of this course, student will be able to:

1 Students will have the basic idea of data mining and other related concepts
2 Students will be aware of various data mining techniques
3 Students will be able to apply classification algorithms
4 Students will be able to apply clustering algorithms and association rules
5 Students will be able to understand advanced data mining techniques

Pre-requisite of course: Basic concepts of databases and statistics

Teaching and Examination Scheme

Theory Hours	Tutorial Hours	Practical Hours	ESE	IA	CSE	Viva	Term Work
4	1	0	50	30	20	0	0

Contents : Unit	Topics	Contact Hours
1	**Introduction and related concepts** Introduction to subject and syllabus, Basic data mining tasks, Data mining vs knowledge discovery in DB, Data mining issues / metrics, Social implications, Data mining from database perspective, Database / OLAP systems, Fuzzy sets and fuzzy logic, Information retrieval, DSS, dimensional modeling , Data warehousing , OLAP, Web search engines, Statistics, Machine learning, Pattern matching	10
2	**Data mining techniques** Introduction , A statistical perspective, Similarity measures, Decision trees, Neural networks, Genetic algorithms, Class test - 1	6
3	**Classification** Introduction, Statistical based algorithms -1, Statistical based algorithms – 2, Distance based algorithms, Decision tree based algorithms – 1, Decision tree based algorithms – 2, Neural network based algorithms -1, Neural network based algorithms – 2, Rules based algorithms, Combining techniques	9

Contents : Unit	Topics	Contact Hours
4	**Clustering and Association rules** Introduction to clustering, Similarity and distance measures, Outliers, Hierarchical algorithms – 1, Hierarchical algorithms -2, Partitional algorithms – 1, Partitional algorithms – 2, Clustering large databases – 1, Clustering with categorical attributes, Association rules Introduction, Large itemsets, Basic algorithms – 1, Basic algorithms – 2, Parallel and distributed algorithms, Comparing approaches, Incremental rules,, Advanced association rule techniques, Measuring the quality of rules	14
5	**Web, Spatial and Temporal mining** Web mining Introduction, Web content mining,, Web structure mining,, Web usage mining – 1, Web usage mining – 2, Spatial mining Introduction, Spatial data overview,, Spatial data mining primitives, Generalization and specialization,, Spatial rules, Spatial classification algorithm,, Spatial clustering algorithms, Temporal mining Modeling temporal events,, Time series,, Pattern detection,, Sequences,, Temporal association rules, Class test - 2	14
Total Hours		**53**

Textbook :

1 Data Mining : Introductory and Advanced Topics, , Margaret Dunham, Pearson Education, 2003

References:

1 Data Mining : Practical Machine Learning Tools and Techniques, Data Mining : Practical Machine Learning Tools and Techniques, Ian H. Witten, Eibe Frank, Mark A. Hall, Morgan Kaufmann, 2016

2 Data Mining : The Textbook, Data Mining : The Textbook, Charu C. Aggarwal, Springer, 2016

3 Data Mining : Concepts and Techniques, Data Mining : Concepts and Techniques, Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann, 2011

Suggested Theory Distribution:

The suggested theory distribution as per Bloom's taxonomy is as follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process

Distribution of Theory for course delivery and evaluation					
Remember / Knowledge	**Understand**	**Apply**	**Analyze**	**Evaluate**	**Higher order Thinking**
15.00	25.00	25.00	20.00	15.00	

Instructional Method:

1 Presentation

Instructional Method:

2 Video tutorials

3 case study

Supplementary Resources:

1 www.nptel.ac.in/courses/106105174