

Course Code	Course Name	Theory	Practical	Tutorial	Theory	Oral & Practical	Tutorial	Total
ITC602	Data Mining and Business Intelligence	04		--	04	--	--	04

Course Code	Course Name	Examination Scheme						
		Theory Marks				Term Work	Oral & Practical	Total
		Internal assessment			End Sem. Exam			
		Test1	Test2	Avg. of two Tests				
ITC602	Data Mining and Business Intelligence	20	20	20	80	--	--	100

**Course Objectives:** Students will try:

1. To introduce the concept of data Mining as an important tool for enterprise data management and as a cutting edge technology for building competitive advantage.
2. To enable students to effectively identify sources of data and process it for data mining
3. To make students well versed in all data mining algorithms, methods of evaluation.
4. To impart knowledge of tools used for data mining
5. To provide knowledge on how to gather and analyze large sets of data to gain useful business understanding.
6. To impart skills that can enable students to approach business problems analytically by identifying opportunities to derive business value from data.

**Course Outcomes:** Student will be able to:

1. Demonstrate an understanding of the importance of data mining and the principles of business intelligence
2. Organize and Prepare the data needed for data mining using pre preprocessing techniques
3. Perform exploratory analysis of the data to be used for mining.
4. Implement the appropriate data mining methods like classification, clustering or Frequent Pattern mining on large data sets.
5. Define and apply metrics to measure the performance of various data mining algorithms.
6. Apply BI to solve practical problems : Analyze the problem domain, use the data collected in enterprise apply the appropriate data mining technique, interpret and visualize the results and provide decision support.

**Prerequisite:** Database Management System, Advanced Data Management Technology.

**Detailed syllabus:**

<b>Sr. No.</b>	<b>Module</b>	<b>Detailed Content</b>	<b>Hours</b>	<b>CO Mapping</b>
0	Prerequisites	Knowledge of databases, and Data warehousing, OLAP	02	--
I	Introduction to Data Mining	What is Data Mining; Kind of patterns to be mined; Technologies used; Major issues in Data Mining	03	CO1
II	Data Exploration and Data Preprocessing	Types of Attributes; Statistical Description of Data; Data Visualization; Measuring similarity and dissimilarity.  Why Preprocessing? Data Cleaning; Data Integration; Data Reduction: Attribute subset selection, Histograms, Clustering and Sampling; Data Transformation & Data Discretization: Normalization, Binning, Histogram Analysis and Concept hierarchy generation.	09	CO2 CO3
III	Classification	Basic Concepts; Classification methods: 1. Decision Tree Induction: Attribute Selection Measures, Tree pruning. 2. Bayesian Classification: Naïve Bayes Classifier. Prediction: Structure of regression models; Simple linear regression, Multiple linear regression. Accuracy and Error measures, Precision, Recall, Holdout, Random Sampling, Cross Validation.	09	CO4 CO5
IV	Clustering	Cluster Analysis: Basic Concepts; Partitioning Methods: K-Means, K-Medoids; Hierarchical Methods: Agglomerative, Divisive, BIRCH; Density-Based Methods: DBSCAN  What are outliers? Types, Challenges; Outlier Detection Methods: Supervised, Semi Supervised, Unsupervised, Proximity based, Clustering Based.	10	CO4 CO5
V	Frequent Pattern	Market Basket Analysis, Frequent Itemsets, Closed Itemsets, and	10	CO4

	Mining	Association Rules; Frequent Pattern Mining, Efficient and Scalable Frequent Itemset Mining Methods, The Apriori Algorithm for finding Frequent Itemsets Using Candidate Generation, Generating Association Rules from Frequent Itemsets, Improving the Efficiency of Apriori, A pattern growth approach for mining Frequent Itemsets; Mining Frequent itemsets using vertical data formats; Introduction to Mining Multilevel Association Rules and Multidimensional Association Rules; From Association Mining to Correlation Analysis, lift, ; Introduction to Constraint-Based Association Mining.		CO5
VI	Business Intelligence	What is BI? Business intelligence architectures; Definition of decision support system; Development of a business intelligence system using Data Mining for business Applications like Fraud Detection, Clickstream Mining, Market Segmentation, retail industry, telecommunications industry, banking & finance CRM etc.	09	CO6

#### **Text Books:**

1. Han, Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann 3rd Edition.
2. P. N. Tan, M. Steinbach, Vipin Kumar, "Introduction to Data Mining", Pearson Education.
3. Business Intelligence: Data Mining and Optimization for Decision Making by Carlo Vercellis ,Wiley India Publications.
4. G. Shmueli, N.R. Patel, P.C. Bruce, "Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner", 2nd Edition, Wiley India.

#### **References:**

1. Michael Berry and Gordon Linoff "Data Mining Techniques", 2nd Edition Wiley Publications.
2. Michael Berry and Gordon Linoff "Mastering Data Mining- Art & science of CRM", Wiley Student Edition.
3. Vikram Pudi & Radha Krishna, "Data Mining", Oxford Higher Education.

#### **Assessment:**

##### **Internal Assessment for 20 marks:**

Consisting of **Two Compulsory Class Tests**

Approximately 40% to 50% of syllabus content must be covered in First test and remaining 40% to 50% of syllabus contents must be covered in second test.

**End Semester Examination:** Some guidelines for setting the question papers are as:

- Weightage of each module in end semester examination is expected to be/will be proportional to number of respective lecture hours mentioned in the syllabus.
- Question paper will comprise of total **six questions, each carrying 20 marks.**
- **Q.1** will be **compulsory** and should **cover maximum contents of the syllabus.**
- **Remaining question will be mixed in nature** (for example if Q.2 has part (a) from module 3 then part (b) will be from any other module. (Randomly selected from all the modules.)
- Total **four questions** need to be solved.