

**DATA MINING & BUSINESS INTELLIGENCE (Major Elective – I)**  
**Semester I (Computer Engineering)**  
**SUB CODE: MECE106-A**

**Teaching Scheme (Credits and Hours)**

Teaching scheme				Total Credit	Evaluation Scheme					Total Marks
L	T	P	Total		Theory		Mid Sem Exam	CIA	Pract.	
Hrs	Hrs	Hrs	Hrs		Hrs	Marks	Marks	Marks	Marks	
04	00	02	06	05	3	70	30	20	30	150

**LEARNING OBJECTIVES:**

The objective of this course is

- Understand the basic concept and the process of data mining
- Understand the basic concept of business intelligence
- Identify the potential benefits, risks, and range of organizational and managerial issues associated with a successful implementation of a business intelligence system.
- Explain the need for a data integration process, data profiling, data cleansing and data enhancement, and their contribution to adding value to data.
- Learn how to use business intelligence to solve business problems
- Distinguish between the concepts of knowledge discovery and creation, and data mining, and select appropriate data mining tools and techniques to implement a business intelligence strategy and applications of data mining.
- Be proficient with leading data mining & business intelligence software

**OUTLINE OF THE COURSE:**

Unit No	Topics
1	Introduction to Data Mining
2	Data Preprocessing
3	Data Mining Primitives
4	Association Rule Mining
5	Classification and Prediction
6	Cluster Analysis
7	Business Intelligence and its Impact
8	Business Intelligence Capabilities

**Total hours (Theory): 60**

**Total hours (Practical): 30**

**Total hours: 90**

**DETAILED SYLLABUS:**

<b>Sr. No</b>	<b>Topic</b>	<b>Lecture Hours</b>	<b>Weight age (%)</b>
1	<b>Introduction to Data Mining:</b> <ul style="list-style-type: none"><li>➤ Importance of Data Mining, Data Mining functionalities, Classification of Data mining systems, Data mining architecture, Major Issues in Data Mining, Applications of Data Mining, Social impacts of data mining.</li></ul>	04	10
2	<b>Data Preprocessing:</b> <ul style="list-style-type: none"><li>➤ Data Preprocessing, Data cleaning, Data Integration and Transformation, Data reduction, Discretization and Concept Hierarchy Generation.</li></ul>	06	10
3	<b>Data Mining Primitives:</b> <ul style="list-style-type: none"><li>➤ Languages and System Architectures, Concept description: Characterization and Comparison, Analytical Characterization, Mining Class Comparison</li></ul>	08	10
4	<b>Association Rule Mining:</b> <ul style="list-style-type: none"><li>➤ Mining of Single dimensional Boolean association rules, Multilevel association rules and Multidimensional association rules, Correlation Analysis, Constraint based association Mining</li></ul>	08	10
5	<b>Classification and Prediction:</b> <ul style="list-style-type: none"><li>➤ Basic issues regarding classification and predication, Classification by Decision Tree, Bayesian classification, classification by back propagation, Associative classification, Prediction, Classifier accuracy</li></ul>	08	15
6	<b>Cluster Analysis:</b> <ul style="list-style-type: none"><li>➤ Basic issues, clustering using partitioning methods, Hierarchical methods, Density based methods, Grid based methods and model based methods, Algorithms for outlier analysis</li></ul>	08	15
7	<b>Business Intelligence and its Impact:</b> <ul style="list-style-type: none"><li>➤ Introduction, Factors driving Business Intelligence, Role of Data, Information, and Knowledge in Data Warehousing, mining and DSS, BI in Organizations, Obstacles in BI</li></ul>	06	10
8	<b>Business Intelligence Capabilities:</b> <ul style="list-style-type: none"><li>➤ Concept of data mart, data warehousing and data mining , data visualization and presentation, Designing physical database, Deploying and supporting DW/BI system, BI Architecture – spread sheets, concept of dashboard, OLAP, decision engineering, LIS</li><li>➤ BI Tools – concept of dashboard</li><li>➤ BI Application in various domains</li><li>➤ BI Analytics (discriminant analysis and logistic regression, cluster analysis, principle component analysis )</li></ul>	12	20

## **INSTRUCTIONAL METHOD AND PEDAGOGY (Continuous Internal Assessment (CIA) Scheme)**

- At the start of course, the course delivery pattern, prerequisite of the subject will be discussed.
- Lectures will be conducted with the aid of multi-media projector, black board, OHP etc.
- Attendance is compulsory in lecture and laboratory which carries 10 marks in overall evaluation.
- One internal exam will be conducted as a part of internal theory evaluation.
- Assignments based on the course content will be given to the students for each unit and will be evaluated at regular interval evaluation.
- Surprise tests/Quizzes/Seminar/tutorial will be conducted having a share of five marks in the overall internal evaluation.
- The course includes a laboratory, where students have an opportunity to build an appreciation for the concepts being taught in lectures.
- Experiments shall be performed in the laboratory related to course contents.

### **STUDENTS LEARNING OUTCOMES:**

On successful completion of the course, the student will:

- Students will be able to learn concept of data mining and business intelligence.
- Applying different data mining techniques.
- Students will have good knowledge about business intelligence.
- Student will be able to data mining classification techniques.

### **STUDY MATERIALS:**

#### **Reference Books:**

1. Data Mining concepts and Techniques by Jiawei Han, Micheline Kamber – Elsevier.
2. Data Mining by Arun K. Pujari – University Press.
3. Mordern Data Warehousing, Data Mining and Visualization by George M. Marakas – Pearson.
4. Data Mining by Vikram Puri and P.RadhaKrishana –Oxfrod Press.
5. Data Warehousing by Reema Theraja –Oxford Press

#### **Web Links:**

1. [www.ibm.com/insights/in](http://www.ibm.com/insights/in)
2. [www.sas.com](http://www.sas.com)
3. Oracle Documentation

## LIST OF PRACTICALS:

Sr No.	Name of Practical
1	Study of SQL Server 2008 Databases Engine, Analysis Services and Business Intelligence Development Studio.
2	Design and Create cube by identifying measures and dimensions for Star Schema and Snowflake schema.
3	Process cube and Browse Cube data by replacing a dimension in the grid, filtering and drilldown using cube browser.
4	Create and use Excel Pivot Table report based on data cube using SQL Server 2008.
5	Design and create data mining models using Analysis Service of SQL Server 2008.
6	Perform Different Data Mining Activities using Weka Explorer Tool (Open Source Data Mining Tool).
7	Perform Different Data Mining Activities using Weka Knowledge Flow Tool (Open Source Data Mining Tool).
8	Perform Different Data Mining Activities using Weka Experimental Flow Tool (Open Source Data Mining Tool).
9	Perform Different Data Mining Activities using <b>“Tanagra”</b> Data Mining Tool.
10	Perform Different Data Mining Activities using <b>“Sapina Research Tool”</b> .
11	Perform Different Data Mining Activities using <b>“R-GUI”</b> Tool.
12	Perform Different Data Mining Activities using <b>“Excel Miner”</b> .
13	BI projects: any one of Balanced Scorecard, Fraud detection, Market Segmentation etc. Using any commercial BI tool like SQLServer , Oracle BI, SPSS, Clementine, and XLMiner etc
14	Case Study: To study the research papers on the Advanced topics on data mining and prepare and present the report on it.