



# INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal - 500 043, Hyderabad, Telangana

## COURSE CONTENT

BUSINESS DATA MINING, WAREHOUSE AND VISULIZATION								
<b>IV Semester:</b> MBA								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P	C	CIA	SEE	Total
CMBD56	Elective	4	-	-	4	40	60	100
		<b>Contact Classes: 45</b>		<b>Tutorial Classes: Nil</b>		<b>Practical Classes: Nil</b>		<b>Total Classes: 45</b>
<b>Prerequisite: Business Analytics</b>								

### I. COURSE OVERVIEW:

The aim of this course is to enhance understanding of data mining concepts, data preprocessing, warehouse efficiency management, classification problems, and data visualization for business. Participants will gain insights into mining various types of data, handling missing values, optimizing warehouse efficiency, building logistic regression models, and creating informative data visualizations for business monitoring.

### II. OBJECTIVES:

**The students will try to learn:**

- I. The concepts of Data Mining and Preprocessing of Data.
- II. The insights of association rule mining and clustering.
- III. The application of Logistic Regression and Sentiment Analytics to solve Business Problems.

### III. COURSE OUTCOMES:

**At the end of the course students should be able to:**

- CO1 Explore the kinds of data that can be mined, major issues in data mining, and the societal impact of data mining.
- CO2 Recognize reasons for data preprocessing, major tasks involved, and techniques for data cleaning, reduction, transformation, and discretization.
- CO3 Analyze the concepts of warehouse operations, order picking methods, and measures of warehouse efficiency.
- CO4 Apply the workforce design, cross-docking, and risk management in warehousing operations.
- CO5 Gain an overview of classification problems, binary logistic regression, and techniques for encoding categorical features.
- CO6 Discuss the data visualization for businesses, covering numerical and non-numerical data visualization techniques.

## IV. COURSE CONTENT:

### **MODULE-I-INTRODUCTION TO DATA MINING, WAREHOUSE AND VISUALIZATION (08)**

Data Mining; Kinds of data that can be mined-Database Data, Data Warehouses, Transactional Data, Other Kinds of Data; Major Issues in Data Mining-Mining Methodology, User Interaction, Efficiency and Scalability, Diversity of Database Types, Data Mining and Society.

### **MODULE -II-DATA PRE-PROCESSING (10)**

Data Preprocessing: An Overview-Reasons to process the data, Major Tasks in Data Preprocessing; Data Cleaning-Missing Values, Noisy Data, Data Cleaning as a Process; Data Reduction-Principal Component Analysis, Histograms, Clustering, Sampling, Data Cube Aggregation; Data Transformation and Data Discretization-Data Transformation by Normalization, Discretization by Binning, Discretization by Histogram Analysis.

Case Study: Handling Missing Values in Melbourne Housing Price Data.

### **MODULE -III-MANAGING WAREHOUSE EFFICIENCY (09)**

Order picking - Picking methods-pick path - Measuring Warehouse Efficiency - Warehouse Workforce design and development - cross docking.

Warehousing Operations: warehousing operations- inbound process, outbound processes, Functions of Warehouse- break-bulk, cross docking, order mixing, Risk management

### **MODULE -IV-CLASSIFICATION PROBLEMS – I-LOGISTICS REGRESSION (10)**

Overview of Classification; Binary Logistic Regression; Classification-Encoding Categorical Features, Building Logistic Regression Model, Printing Model Summary, Predicting on Test Data; Measuring Accuracies-Creating Confusion Matrix, Receiver Operating Characteristic (ROC) and Area Under the Curve; Finding Optimal Classification Cut-off -Youden's Index and Cost-Based Approach.

Case Study: Predicting Employee Attrition on HR Attrition Dataset.

### **MODULE -V-DATA VISUALISATION FOR BUSINESS (08)**

Introduction to Data Visualization for businesses, Visualization of Numerical and Non-Numerical Data Creation of Dashboards using IT Tools, Business Activity Monitoring through Dashboard.

## V. TEXT BOOKS:

1. Ikvinderpal "Data Mining & Warehousing", 1<sup>st</sup> edition, 2020.
2. Max Bramer and Springer "Principles of Data Mining 2020" 8<sup>th</sup> edition, 2020.
3. Szabo, Gungar Polatkan, Oscar Boykin, Chalkiopoulos, "Social Media Data Mining and Analytics", 3<sup>rd</sup> edition. Wiley, 2019.
4. Pang-Ning Tan, "Introduction to Data Mining, Global Edition", 4<sup>th</sup> edition, May 2019.
5. Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal, "Data Mining: Practical Machine Learning Tools and Techniques", 4<sup>th</sup> edition. Elsevier, 2017.
6. Megan Squire, "Mastering Data Mining with Python –Find patterns hidden in your data", 1<sup>st</sup> edition. PACKT Publishing, 2016.
7. Charu C. Aggarwal, "Data Mining and warehousing", 3<sup>rd</sup> edition, 2016.

## VI. REFERENCE BOOKS:

1. Luis Torgo, "Data Mining with R: Learning with Case Studies", 2<sup>nd</sup> edition, CRC Press, 2011.
2. Jiawei Han, Jian Pei, Micheline Kamber, "Data Mining: Concepts and Techniques", 3<sup>rd</sup> edition, Elsevier,2010.
3. Joseph B. Pigus, "Data Mining with Neural Networks", 2<sup>nd</sup> edition, TMH, 2017.
4. Robert Layton, "Learning Data Mining with Python", 2<sup>nd</sup> edition, PACKT Publishing, 2015.
5. Xin-She Yang, "Introduction to Algorithms for Data Mining and ML", 1<sup>st</sup> edition, Academic

- Press, 2019.
6. Boris Kovalerchuk, Evgeni Vityaev, “Data Mining in Finance”, 3<sup>rd</sup> edition, Kluwer Academic Publishers, 2010.
  7. Dinesh K. “Business Analytics: The Science of Data - Driven Decision Making”. 1<sup>st</sup> edition, Wiley, 2009.
  8. Yau, N., “Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics”. John Wiley & Sons.

#### **VII.WEB REFERENCES:**

1. [https://www.amazon.in/dp/B07YG4QSZR/ref=cm\\_sw\\_r\\_apan\\_glt\\_2FGRNQAE4AEV5JQRB7?\\_encoding=UTF8&psc=1](https://www.amazon.in/dp/B07YG4QSZR/ref=cm_sw_r_apan_glt_2FGRNQAE4AEV5JQRB7?_encoding=UTF8&psc=1)

#### **VIII.E-TEXT BOOKS:**

1. [https://www.amazon.in/dp/B00UVBJSAQ/ref=cm\\_sw\\_r\\_apan\\_glt\\_1Y1H0P2MRK1KM0SFG1AJ?\\_encoding=UTF8&psc=1](https://www.amazon.in/dp/B00UVBJSAQ/ref=cm_sw_r_apan_glt_1Y1H0P2MRK1KM0SFG1AJ?_encoding=UTF8&psc=1).
2. [https://www.amazon.in/dp/B075GB7FT5/ref=cm\\_sw\\_r\\_apan\\_glt\\_G8RNA0P6W9YSQZ2N0P16?\\_encoding=UTF8&psc=1](https://www.amazon.in/dp/B075GB7FT5/ref=cm_sw_r_apan_glt_G8RNA0P6W9YSQZ2N0P16?_encoding=UTF8&psc=1).