PROGRAM	**Master of Business Administration – Business Analytics**
SEMESTER	**II**
COURSE TITLE	**Data Mining for Business Analytics using R**
COURSE CODE	**04MB0211**
COURSE CREDITS	**03**
COURSE DURATION	**42 Hrs (42 sessions of 60 minutes each)**

COURSE OUTCOMES:

- Appraise the importance of various data mining techniques and define key terms associated with Data Mining.
- Explain various aspects of data and Explore data using different visualization techniques.
- Apply classification model for prediction.
- Analyze underneath patterns in the data such as natural groupings using cluster analysis.
- Evaluate Association Mining rule to determine pattern in transactions.

COURSE CONTENTS:

Unit No	Unit / Sub Unit	Sessions
I	**Introduction, Data and Pre-processing:** What is data mining? Why data mining? Data Mining Process, Relation to Business Intelligence Techniques, Introduction to Data Mining Task (Classification, Clustering, Association Analysis, Anomaly Detection). What is a model? Basic terminologies, Predictive modelling, Real-world Data mining Applications. Data and Pre-processing: Understanding of Data, what is Data? Types of Attributes, Properties of attribute values, Sampling, Data Normalization, Data Cleaning, Similarity Measures, and Feature Selection.	10
II	**Classification:** Decision-Tree Based Approach, Rule-based Approach, Instance-based classifiers (Nearest Neighbour), Classification Model Selection and Evaluation**.**	6
III	**Association Rules:** Discovering and Generating Association Rules in Transaction Databases, Data Format, Frequent Itemset Generation, Frequency-plot and The Apriori Algorithm, Rule Generation, The Process of Rule Selection and Interpreting the Results, FP-Growth Algorithm and Evaluation of Association Patterns.	8
IV	**Cluster Analysis:** Introduction to Cluster Analysis, Methods of Measuring Distance Between Two Records and Between Two Clusters, Hierarchical Agglomerative Clustering, Non-hierarchical Clustering: The k-means Algorithm, The DBSCAN Algorithm, Cluster Evaluation.	8
V	**Data Visualization:** Use of Data Visualization in Descriptive Analytics, Tables and Graphs, Basic Charts: Bar Charts, Line Graphs and Scatter Plots, Editing Charts and Graphs using Colors, Size, Legend, Rescaling and Zooming the Plots, Parallel Coordinates Plots, Visualization of Geographical Data, Applications.	10

EVALUATION:

The students will be evaluated on a continuous basis and broadly follow the scheme given below:

	Component	Weightage
A	Continuous Evaluation Component (Assignments / Quizzes /Class Participation etc.)	20% (C.E.C.)
B	Internal Assessment (Lab based Practical Examination using R-software)	30% (I.A.)
C	End-Semester Practical Examination	50% (Practical/VIVA)

SUGGESTED READINGS:

Text Books:

Sr. No	Author/s	Name of the Book	Publisher	Edition and Year
T-01	Galit Shmueli, Peter Bruce, Inbal Yahav, Nitin R Patel	Data Mining for Business Analytics: Concepts Techniques and Applications in R	Wiley	2017
T-02	Jiawei Han	Data Mining: Concepts and Techniques	Morgan Kaufmann Publishers	3rd Edition, 2012
T-03	Pang-Nin Tan, Vipin Kumar	Introduction to data Mining	Pearson India	2016

Reference Books:

Sr. No	Author/s	Name of the Book	Publisher	Edition and Year
R-01	Charu C. Aggarwal	Data Mining	Springer International Publishing	2015
R-02	Johannes Ledolter	Data Mining and Business Analytics with R	John Wiley & Sons	2013
R-03	Yanchang Zhao	Data Mining Applications with R	Academic Press	2014