Paper Cade: BCADSC 5.2	**Paper title:** Data Warehousing and Mining	**Teaching Hours** – 5 hrs/week
Total Teaching Hours: 60 Hrs.	**Marks:** Th-80+IA-20	**Credits:** 4

UNIT I
DATA WAREHOUSING, BUSINESS ANALYSIS AND ON-LINE ANALYTICAL PROCESSING (OLAP): Basic Concepts - Data Warehousing Components – Building a Data Warehouse – Database Architectures for Parallel Processing – Parallel DBMS Vendors - Multidimensional Data Model – Data Warehouse Schemas for Decision Support, Concept Hierarchies - Characteristics of OLAP Systems – Typical OLAP Operations, OLAP and OLTP.   **12 Hrs**

UNIT II
DATA MINING: Introduction to Data Mining Systems – Knowledge Discovery Process – Data Mining Techniques – Issues – applications- Data Objects and attribute types, Statistical description of data, Data Preprocessing – Cleaning, Integration, Reduction, Transformation and discretization, Data Visualization, Data similarity and dissimilarity measures.
12 Hrs
UNIT III
DATA MINING - FREQUENT PATTERN ANALYSIS: Mining Frequent Patterns, Associations and Correlations – Mining Methods- Pattern Evaluation Method – Pattern Mining in Multilevel, Multi Dimensional Space – Constraint Based Frequent Pattern Mining, Classification using Frequent Patterns.   **12 Hrs**

UNIT IV
CLASSIFICATION AND CLUSTERING: Decision Tree Induction - Bayesian Classification – Rule Based Classification – Classification by Back Propagation – Support Vector Machines –– Lazy Learners – Model Evaluation and Selection- Techniques to improve Classification Accuracy. Clustering Techniques – Cluster analysis-Partitioning Methods - Hierarchical Methods – Density Based Methods - Grid Based Methods – Evaluation of clustering – Clustering high dimensional data- Clustering with constraints, Outlier analysis-outlier detection methods.   **12Hrs**

UNIT V
WEKA TOOL: Datasets – Introduction, Iris plants database, Breast cancer database, Auto imports database - Introduction to WEKA, The Explorer – Getting started, Exploring the explorer, Learning algorithms, Clustering algorithms, Association–rule learners.   **12Hrs**

References:
1. Jiawei Han and MichelineKamber, ―Data Mining Concepts and Techniques‖, Third Edition, Elsevier, 2012.
2. Alex Berson and Stephen J.Smith, ―Data Warehousing, Data Mining & OLAP‖, Tata McGraw – Hill Edition, 35th Reprint 2016

Additional Reading: .
1. K.P. Soman, ShyamDiwakar and V. Ajay, ―Insight into Data Mining Theory and Practice‖, Eastern Economy Edition, Prentice Hall of India, 2006.
2. Ian H.Witten and Eibe Frank, ―Data Mining: Practical Machine Learning Tools and Techniques‖, Elsevier, Second Edition