B.Sc. (Data Science)
Discipline Specific Course (DSC)
Semester III
BSDB32302T: Data Mining and Visualization

Total Marks: 100
External Marks: 70
Internal Marks: 30
Credits: 4
Pass Percentage: 40%

Objective

This course helps student understand the need for Data Mining and advantages to the business world. They will get a clear idea of various classes of Data Mining techniques, their need, scenarios (situations), scope of their applicability and use of data visualization techniques

Section-A

Unit -1 Data Mining: Introduction, Scope, What is Data Mining; How does Data Mining Works, Predictive Modeling: Data Mining and Data Warehousing: Architecture for Data Mining: Profitable Applications: Data Mining Tools

Unit -II: Data Pre-processing: Overview, Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Generation.

Unit- III: Data Mining Techniques- An Overview, Data Mining Versus Database Management System, Data Mining Techniques- Association rules, Classification, Regression, Clustering, Neural networks.

Unit -IV Clustering: Introduction, Cluster Analysis, Clustering Methods- K means, Hierarchical clustering, Agglomerative clustering, Divisive clustering, evaluating clusters.

Section-B

Unit -V Applications of Data Mining: Introduction, Business Applications Using Data Mining- Risk management and targeted marketing, Customer profiles and feature construction, Medical applications (diabetic screening), Scientific Applications using Data Mining, Other Applications.

Unit -VI Data Visualization: Introduction, Acquiring and Visualizing Data, Simultaneous acquisition and visualization, Applications of Data Visualization, Keys factors of Data Visualization (Control of Presentation, Faster and Better JavaScript processing, Rise of HTML5, Lowering the implementation Bar)

Unit -VII Exploring the Visual Data Spectrum: charting Primitives (Data Points, Line Charts, Bar Charts, Pie Charts, Area Charts), Exploring advanced Visualizations (Candlestick Charts, Bubble Charts, Surface Charts, Map Charts, Infographics).

Unit -VIII Visualizing data Programmatically, Starting with Google charts (Google Charts API Basics, A Basic bar chart, A basic Pie chart, Working with Chart Animations).

Suggested Readings

1. Jiawei Han, Micheline Kamber and Jian Pei, Data Mining Concepts and Techniques, Third Edition, 2000
2. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, Introduction to Data Mining, Pearson 2005,
3. M. Kantardzic, "Data Mining: Concepts, Models, Methods, and Algorithms", 2nd edition, Wiley-IEEE Press, 2011
4. Jon Raasch, Graham Murray, Vadim Ogievetsky, Joseph Lowery, JavaScript and jQuery for Data Analysis and Visualization, 2014
5. Ben Fry, "Visualizing data: Exploring and explaining data with the processing environment", O'Reilly, 2007