ECSEVAC03 DATA ANALYTICS WITH R

UNIT – I

Introduction to R : Introduction - Data Types in R - Few Commands for Data Exploration. Loading and Handling Data in R : Challenges of Analytical Data Processing - Expression, Variables and Functions - Missing Values Treatment - Vectors - Matrices - Factors - List - Common Analytical Tasks - Methods for Reading Data - Comparison of R GUIs for Data Input - Using R with Databases. Exploring Data in R : Data Frames - Load Data Frames - Exploring Data - Data Summary - Finding the Missing Values - Invalid Values and Outliers - Descriptive Statistics - Spotting Problems in Data with Visualization.

UNIT – II

Linear Regression using R : Regression - Model Fitting - Linear Regression - Assumptions of Linear Regression - Validating Linear Assumption. Logistic Regression : Introduction to Generalised Linear Models - Logistic Regression - Binary Logistic Regression - Diagnosing Logistic Regression - Multinomial Logistic Regression Models. Time Series in R : Time Series Data – Reading, Plotting and Decomposing Time Series Data - Forecasts Using Exponential Smoothing - ARIMA Models.

UNIT – III

Decision Tree : Introduction - Decision Tree Representation in R - Appropriate Problems for Decision Tree Learning - Basic Decision Tree Learning Algorithm - Measuring Features - Issues in Decision Tree Learning. Clustering : Basic Concepts in Clustering - Hierarchical Clustering - k-means Algorithm - CURE Algorithm. Association Rules : Frequent Itemset - Mining Algorithm Interfaces - Auxiliary Functions - Sampling from Transaction - Generating Synthetic Transaction Data - Additional Measures of Interestingness.

UNIT – IV

Text Mining : Few Challenges - Text Mining in R - General Architecture of Text Mining Systems - Pre-processing of Documents in R - Core Text Mining Operations - Text Mining Query Languages - Mining Frequent Patterns,

Associations, and Correlations - Frequent Itemsets, Closed Itemsets and Association Rules - Mining Methods - Pattern Evaluation Methods - Sentiment Analysis. Parallel Computing with R : Introduction of R Tool Libraries - Opportunities in HPC to Empower R - Support for Parallelism in R - Comparison of Parallel Packages in R.

UNIT – V

Case Study : Log Analysis - Recommendation Engines - Audience/Customer Insights Analysis - In-store Customer Traffic Prediction - Insurance Fraud Detection - Personalised Product Recommendations - Making User-generated Content Valuable - Credit Card Spending by Customer Groups can be Identified by using Business Needs - Sales Forecasting.

TEXT BOOKS

Bharti Motwani, “Data Analytics with R”, Wiley India Private Limited, 2019.

Seema Acharya, “Data Analytics Using R”, McGraw Hill Education (India) Private Limited, 2018.

REFERENCES

Eric Mayor, “Learning Predictive Analytics with R”, Packt Publishing Limited, 2015.

Simon Walkowiak, “Big Data Analytics with R”, Packt Publishing Limited, 2016.

Umesh R. Hodeghatta and Umesha Nayak, “Business Analytics Using R - A Practical Approach”, Apress, 2017.

Viswa Viswanathan, “Data Analytics with R: A Hands-on Approach”, Infivista Inc., 2nd edition, 2015.

COURSE OUTCOMES

At the end of this course, the students will be able to:

- Develop simple applications and perform data visualisation in R.
- Solve the problems on regression and time series using R.
- Utilize R programming to perform text mining and parallel computing.
- Apply machine learning algorithms on real-time data analytics problems in R.

FACULTY OF EDUCATION

DEPARTMENT OF EDUCATION

UEDUVAC01 - Modern Pedagogical Techniques

Learning Objectives (LO)

The student teacher

- Acquires the knowledge of the concept of Teaching, characteristics of Teaching and Variables in Teaching
- Understands the Need and Importance of Programmed Learning, Computer Assisted Learning Internet and its applications
- Understands the different teaching method
- Acquires the knowledge of Educational E-Resources and Cloud computing

Course Outline

Unit I Concept of Teaching

Meaning and definition of Teaching – Nature and characteristics of Teaching – Variables in Teaching – Relationship between Teaching and Learning.

Unit II Individualised Instruction and ICT

Need and Importance – Characteristics – Principles – Programmed Learning – Computer Assisted Learning (CAL) Advantages and Disadvantages – Information and Communication Technology – Definition – Meaning – Importance and Scope – ICT in Education

Unit III Group Instruction and ICT

Dynamic method – Group discussion, Tutorial, Seminar, Symposium and Brain – storming – Problem solving – Project method – Laboratory – Supervised study – Source method – Dalton Plan – Panel discussion – Workshop – Team Teaching.

Unit IV Educational E-Resources

Meaning – Need and Importance – Classification of E-Resources – offline resources: CD Rom based E-Resources, E-Books, E-dictionaries – MS Office