INT234:PREDICTIVE ANALYTICS

L:0 T:0 P:4 Credits:2

Course Outcomes: Through this course students should be able to

- review the art and science of predictive analytics to define clear actions that result in improved decisions and business results
- evaluate the use of analytic tools and assist in the selection of industry standard analytics tools
- construct and format data to be most effective to ensure the predictive model meets the business goals
- deduce the characteristics of data sets and compare the trivial data and big data for various applications

List of Practicals / Experiments:

DATA PREPROCESSING

- Managing data with R
- Exploring and understanding data
- Exploring the structure of data
- Exploring numeric variables
- Exploring categorical variables
- Exploring relationships between variables

SUPERVISED LEARNING: CLASSIFICATION

- Lazy learning:Nearest neighbors
- Probabilistic Learning: Using Naive Bayes
- Divide and Conquer: Decision Trees and Rules

SURPERVISED LEARNING : NUMERIC PREDICTION

- Forecasting Numeric Data
- Simple Linear Regression
- Polynomial Regression
- Ordinary least squares estimation
- Correlations
- Multiple Linear Regression

SUPERVISED LEARNING:DUAL USE

- Black Box Methods
- Neural Networks
- Support Vector Machines

UNSUPERVISED LEARNING: CLUSTERING AND PATTERN DETECTION

- K-Means Clustering
- K-means clustering intuition
- K-means random initialization trap
- K-means selecting number of clusters
- Dataset gathering
- Hierarchical Clustering
- Association Rules
- Finding Patterns
- Market Basket Analysis Using Association Rules

MODEL PERFORMANCE

- Evaluation Model Performance
- Improving Model Performance
- Bagging
- Boosting
- Random forests

References:

1. APPLIED PREDICTIVE ANALYTICS: PRINCIPLES AND TECHNIQUES FOR THE PROFESSIONAL DATA ANALYST by DEAN ABBOTT, WILEY, 4th Edition, (2012)

2. DATA ANALYTICS MADE ACCESSIBLE by ANIL MAHESHWARI, AMAZON.COM

3. R IN A NUTSHELL 2E by JOSEPH ADLER, O'REILLY

4. MACHINE LEARNING WITH R by BRETT LANTZ, PACKT PUBLISHING