UCS654: PREDICTIVE ANALYTICS USING STATISTICS

L	T	P	Cr
2	0	2	3.0

Course Objectives: Advanced analytics requires the use of unstructured data. Uncertainty is a primary characteristic of unstructured data. Statistical methods that relate to correlating information, finding patterns, predictive modeling are essential in dealing comprehensively with data so that it can used as information to make decisions. This course will provide an overview of statistical methods relevant in the world of business analytics. This will be demonstrated through the use of case studies and statistical software.

Probability, conditional probability, random variable, PDF, PMF, joint distribution, statistical independence, variance, co-variance, correlation, differenrent distribution functions, Bayes theorem, central limit theorem.

Sampling-Distributions, Parameter-Estimations, Hypothesis-Testing, Two-population, Tests, Regression and Correlation, UniVariate-Analysis, Multi-Variate, ANOVA.

Mathematical modeling of regression (linear, non-linear, multiple), understanding error in model training (loss, bias, variance, overfitting, underfitting), maximum likelihood estimation to solve regression, transformation of classification to regression.

Basics of Neural Networks, different loss functions, validation and regularization, multilayered, parameter optimization methods

Data generation using modeling and simulation, Association mining, ECLAT, Measuring data similarity and dissimilarity, and TOPSIS.

Course Learning Outcomes (CLOs) / Course Objectives (COs):

Course Objectives:

1. Demonstrate the ability to use basic probability concepts with descriptive statistics.
2. Visualize the patterns in the data.
3. Demonstrate the use of statistical methods to estimate characteristics of the data.
4. Explain and demonstrate the use of predictive analytics in the field of data science.

Text Books:

1. Peter Dalgaard, Introductory Statistics with R, Springer, Second Edition, ISBN: 978-0-387-79053-4
2. Brett Lantz, Machine Learning with R (2nd Edition), www.PacktPub.com.

Reference Books:

1. Online Resources: (e.g., http://r-statistics.co/)
2. Introduction to Machine Learning in R

https://www.kaggle.com/camnugent/introduction-to-machine-learning-in-r-tutorial