Soft Core	STS 553 : DATA SCIENCE	No. of credits: 3

Course Outcomes:

CO1: Develop in depth understanding of the key technologies in data science.

CO2: Understand the theory of certain Statistical method of data Science.

CO3: Gain practical, hands-on experience with statistics programming languages and big data tools through coursework.

CO4: Demonstrate knowledge of statistical data analysis techniques utilized in company decision making.

CO4: Apply principles of Data Science to the analysis of real-life data. .

CO5: Employ cutting edge tools and technologies to analyze data of Data Science Apply algorithms to build machine intelligence.

Unit I

Introduction to data science, and its product, data science environment/ecosystem. Data visualization - data graphics, elements of visual perception, data manipulation -SQL, merging, aggregating and iterating. (12 hrs)

Unit II

Data preparation - data cleaning, methodologies and its validation. Python/R Data preparation and applications. (8hrs)

Unit III

Machine learning Methods: Bayes model and residual error, linear methods, nearest neighbourhood methods and support vector machines. Neutral network, artificial intelligence. Model selection and Evaluation. (12 hrs)

Unit IV

Random Forests - Bias variance decomposition, regression, classification, randomised models Multi-model data processing- components and applications.

(10 hrs)

A mini project on data science by self study (for internal evaluation only).