STAREBS ONLINE PROGRAM

Applied Data Science & Analytics using R

February, 2026

Duration: Total **30 Hours (**20 Sessions) (Including Project work and presentation)

Course Instructor: Debasish Mukherjee

Pre-requisites: Fundamentals of Statistical Inference, Basic Linear Algebra and Calculus

BRIEF OUTLINE OF THE COURSE

EVERY TOPIC IS DEALT WITH Case Study and HANDS ON PRACTICE.

- Introduction to Data Science, Data Science Project Life Cycle, Introduction to Machine Learning, Overview of Supervised and Unsupervised Machine Learning;
- Introduction to R for Data Science; Data Wrangling; Exploratory Data Analysis (EDA), Data Visualization using R;
- Supervised learning for Continuous target, Building Blocks of a ML project, Case Study: Building it all together using R; Gradient Descent, intuition, mathematics, implementation of Gradient Descent for Linear Regression; Overfitting and Under fitting of model, Bias Variance trade-off, Regularization, Ridge, Lasso and Elastic Net Regression;
- Supervised learning for Categorical target, Logistic Regression, Classification Performance Metrics, Decision Tree, Random Forest;
- Introduction to Artificial Neural Network, Activation functions, Forward and Backward Propagation, Implementation using R;
- Unsupervised Machine Learning, Clustering, Dimensionality Reduction;
- Machine Learning for Text data, Natural Language Procession (NLP), Sentiment Analysis;
- PAPER Discussion: Application of Machine Learning on Text Data

Group Project; Presentation and discussion of findings. Group Projects may be evaluated for final grades as decided by the instructor.