R COURSES

1. Introduction to programming in R
2. Basic statistics using R
3. Credit risk management using R
4. Portfolio optimization in R
5. Artificial neural networks using R

INTRODUCTION TO PROGRAMMING IN R

Learning Outcome Statements

- Learn how to program in R
- Learn to use R for data analysis
- Learn to install and configure a statistical programming environment
- Understand generic programming concepts as they apply to R

Key Contents

- Learning R
  - Objects,
  - Data types, Data Structures
  - Assignments
- Working with R
  - Installing Packages
  - Working with Scripts
  - Setting up working Directories
  - Sub-setting data with square bracket
  - Importing Data
  - Logical Statements
- Doing Math in R
  - Math Function
  - Functions for Statistical Distributions
  - Sorting and Operations on Vectors and Matrices
- Plots in R
  - Histograms
  - Bar plot
  - Box Plots
  - Co-Plot
  - Pairs
  - QQ Norm
  - QQ Plot
  - Scatterplots
- Control Statements
  - Conditionals statements
  - Loops
  - Functions

  Writing your own Function

BASIC STATISTICS USING R

Learning Outcome Statements

- Learn to sample and explore data
- Introduction to basic probability theory and Bayes' rule
- Learn about various continuous and discrete probability distributions
- Discuss hypothesis testing and relate testing procedures back to estimation via confidence intervals.

Key Contents

- Data Description
  - Descriptive Statistics
  - Exploratory Data Analysis
  - Multivariate Data and Data Frames
- Probability
  - Sample Spaces
  - Events
  - Properties of Probability
  - Counting Methods
  - Conditional Probability
  - Independent Events
  - Bayes' Rule
  - Random Variables
- Discrete Distributions
  - Discrete Random Variables
  - Discrete Uniform Distribution
  - Binomial Distribution
  - Other Discrete Distributions
- Continuous Distributions
  - Continuous Random Variables
  - Continuous Uniform Distribution
  - Normal Distribution
  - Other Continuous Distributions
- Multivariate Distributions
  - Joint and Marginal Probability Distributions
  - Joint and Marginal Expectation
  - Conditional Distributions
  - Independent Random Variables
  - Bivariate Normal Distribution
  - Multinomial Distribution
- Sampling Distribution
  - Simple Random Samples
  - Central Limit Theorem
  - Sampling Distributions of Two-sample Statistics
- Estimation
  - Point Estimation
  - Confidence Interval for Means
  - Confidence Intervals for Difference of Means
  - Confidence Intervals for Proportions
  - Confidence Intervals for Variances
  - Fitting Distributions
  - Sample Size and Margin of Error
- Hypothesis Testing
  - Tests for Proportion
  - One Sample Tests for Means and Variances
  - Two-Sample Tests for Means and Variances
  - Analysis of Variance
- Simple Linear Regression
  - Estimation
  - Residual Analysis
- Multiple Linear Regression
  - Multiple Linear Regression Model
  - Estimation and Prediction
  - Polynomial Regression
  - Residual Analysis

CREDIT RISK MODELING USING R

Learning Outcome Statements

- Gain an introduction to credit risk modeling
- Understand how logistic regression is used in credit risk modeling
- Understand how classification trees are used in credit risk modeling
- Learn to evaluate and compare credit risk models

Key Contents

Basic R
- Learning R
- Doing Statistics in R

Machine Learning Algorithms
- Linear Regression
  - Regression Hypotheses and Assumptions
  - Ordinary Least Squares (OLS)
  - Computing Regression Coefficients
  - Coefficient of Determination
  - Interpretation of the Regression Output
- Multiple Linear Regression
  - Types of Multiple Linear Regression
  - Building the Multiple Linear Regression Model
  - Two-Variable Model
  - Computing Regression Coefficients
  - Coefficient of Determination and Adjusted R2
  - Interpretation of the Regression Output
- Logistic Regression
  - Probability
  - Odds and Odds Ratio
  - Logit and Inverse-logit functions
  - Computing Regression Coefficients
  - Coefficient of Determination and Adjusted R2
  - Interpretation of the Regression Output
- Decision Tree & Random Forest
  - Information theory
  - Structure of the tree
  - Building the tree
  - Evaluation of tree (Confusion Matrix, ROC Curves, Pearson Correlation Coefficient, Area Under Curve)
  - Prediction using tree
  - Overfitting and Underfitting
  - Cross Validation
  - Pruning Trees
  - Bagging
  - Random Forest

Introduction to Credit Risk Modeling
- Approach to Model Building
- Architectural Suggestions
- Reading Data In
- Binning Example
- Example of Binning or Coarse Classifying in R
- Breaking Data into Training and Test Sample
- Traditional Credit Scoring
- Traditional Credit Scoring Using Logistic Regression in R
- Calculating ROC Curve for model
- Calculating KS Statistic
- Calculating top 3 variables affecting Credit Score Function in R

Credit Risk Modelling in R
- Using Bayesian N Using Traditional recursive Partitioning
- Comparing Complexity and out of Sample Error
- Compare ROC Performance of Trees
- Converting Trees to Rules
- Bayesian Networks in Credit Scoring
- Using Traditional recursive Partitioning
- Comparing Complexity and out of Sample Error
- Compare ROC Performance of Trees
- Converting Trees to Rules
- Conditional inference Trees
- Using Random Forests
- Calculating Area under the Curve
- Cross Validation
- Cutting Edge techniques: Party Package (Unbiased Non parametric methods-Model Based Trees)

PORTFOLIO OPTIMIZATION IN R

Learning Outcome Statements

- Gain an introduction to robust portfolio optimization techniques as a remedy to the outlier sensitivity encountered by plain Markowitz optimization
- Understand the concept of portfolio diversification and concepts of the most diversified, equal risk contributed and minimum tail-dependent portfolios
- Understand the concept of conditional value at risk and draw-down of a portfolio
- Understand the concept of copula opinion pooling and the construction of a wealth protection strategy

Key Contents

Brief Course in R
- Working with R
- Classes, methods and functions

Modern Portfolio Theory
- Markowitz portfolios
- Empirical mean-variance portfolios

Portfolio Optimization Approaches
- Robust portfolio optimization
- Robust statistics
- Selected robust estimators
- Robust optimization
- Synopsis of R packages
  - *covRobust*
  - *fPortfolio*
  - *MASS*
  - *robustbase*
  - *robust*
  - *rrcov*

Diversification
- Most-diversifed portfolio
- Risk contribution constrained portfolios
- Optimal tail-dependent portfolios
- Synopsis of R packages
  - *cccp*
  - *DEoptim, DEoptimR, and RcppDE*
  - *FRAPO*
  - *PortfolioAnalytics*

Risk-optimal portfolios
- Mean-VaR portfolios
- Optimal CVaR portfolios
- Synopsis of R packages
  - *fPortfolio*
  - *FRAPO*
  - Packages for linear programming
  - *PerformanceAnalytics*

Tactical Asset Allocation Overview 274
- Univariate time series models
- Multivariate time series models
- The Black–Litterman approach
- Copula opinion and entropy pooling
- Synopsis of R packages
  - *BLCOP*
  - *dse*
  - *fArma*
  - *forecast*
  - *MSBVAR*
  - *PortfolioAnalytics*
  - *urca* and *vars*
- Black–Litterman portfolio optimization
- Copula opinion pooling
- Entropy pooling
- Protection strategies

Probabilistic Utility
- Overview 339
- Concept of probabilistic utility
- Markov Chain Monte Carlo
- Monte Carlo approaches
- Markov chains
- Metropolis–Hastings algorithm
- Synopsis of R packages
  - *Packages for conducting MCMC*
  - *Packages for analyzing MCMC*

NEURAL NETWORKS INR

Learning Outcome Statements

- Get familiar with simple mathematical concepts relevant to neural networks
- Understand data representation of neural networks
- Learn to use neural networks to solve real problems
- Examine the core components of neural networks: layers, networks, objective functions, and optimizers.
- Learn about the three most common use cases of neural networks: binary classification, multiclass classification, and scalar regression
- Gain a brief introduction to Keras

Key Contents

Data representation of neural networks
- Scalars
- Vectors
- Matrices
- 3D tensors and higher-dimensional tensors
- Manipulating tensors in R
- Vector data
- Time series data or sequence data
- Image data
- Video data

Tensor Operations
- Element-wise operations
- Operations involving tensors of different dimensions
- Tensor Reshaping
- Geometric interpretation of tensor operations
- Geometric interpretation of deep learning

Gradient-based Optimization
- Derivative of a tensor operation
- Stochastic gradient descent
- Backpropagation algorithm
- Training neural networks using gradient descent

Anatomy of a Neural Network
- Layers
- Models
- Loss functions and optimizers

Introduction to Keras
- Keras
- Tensorflow
- Theano
- CNTK

Setting Up a deep-learning workstation

Case Studies
- Binary classification example
- Multiclass classification example
- Regression example