
Machine Learning with R / Python

Duration: 7-8 weekends

Prerequisites

- Basic knowledge of any Programming Language.
- Basic knowledge of Database (SQL) and files (MS Excel, CSV etc.)
- Basic high school Algebra and Geometry.

Mathematics Fundamentals

1. Fundamentals of Statistics

2. Introduction to Statistics

- Types of data
- Measures of central tendency and dispersion
- Statistical Graphics

3. Probability and Probability Distributions

- Binomial Distribution
- Poisson Distribution
- Normal Distribution

R Programming

1. R Programming Basics

- Introduction to R
- Data Types
- Reading data, Subsetting Data
- Visualizing the Data
- Input Output Sub setting
- Control structure
- Functions
- Data Exploration
- Data Harmonization

2. Descriptive & Inferential Statistics

- Estimation Theory
- Sampling Distribution
- Point Estimation
- Interval Estimation
- Sampling Distribution
- Test of Hypothesis
- Inference about one population means
- Inference about two populations means
- Analysis of Variance Concept
- Inference about one & two population (Means & Proportion)
- Analysis of Variance (1 Way & 2 Way)

Python Programming

1. An Introduction to Python

2. Python Core Objects and Functions

- Built in modules (Library Functions)
- Numeric and Math's Module
- String/List/Dictionaries/Tuple
- Complex Data structures in Python

- Arbitrary data types and their Data Structure
- Python built in function
- Python user defined functions
- Python packages and functions

Machine Learning with Python

1. Concepts majorly used to explain are:

A. NUMPY

B. SCIKIT LEARN

C. PANDAS

- Supervised Algorithms
- Introduction to Machine Learning
- Naïve Bays Algorithm
- K-Nearest Neighbor Algorithm
- Decision Tress (SingleTree)
- Regression
- Correlation coefficient
- Simple Linear Regression
- Multiple Linear Regression
- Logistic Regression
- Time Series Analysis
- Moving Average
- Simple Exponential Smoothing
- Holt-Winters Method
- ARIMA Models
- Support Vector Machines
- Random Forest
- Support Vector Machines
- Model Ensembling
- Bagging
- Boosting
- Stacking

2. Unsupervised Learning Algorithms

- Cluster Analysis
- Hierarchical Clustering
- K-means Clustering
- Association Rules Mining
- Principal Components Analysis

3. Natural Language Processing

- Term Document Matrix
- TF-IDF
- Word Cloud
- Recommendations Systems

4. Neural Network