Call:- 9574180345

Data Science Syllabus

Statistics Essentials for Analytics

All the topics in the following section will explain the basics of what it is, which scenario you want to use, What math behind it, How to implement with an analytic tool, what inferences you are getting from the final result.

- Understanding the Data
- Probability and its Uses
- Statistical Inference
- Data Clustering
- Testing the Data
- Regression Modelling

Data Science with Python

Module 1: Introduction to Data Science

- What is Data Science?
- What is Machine Learning?
- What is Deep Learning?
- What is AI?
- Data Analytics & it's types

Module 2: Introduction to Python

- What is Python?
- Why Python?
- Installing Python
- Python IDEs
- Jupyter Notebook Overview

Module 3: Python Basics

- Python Basic Data types
- Lists
- Slicing
- IF statements
- Loops

- Dictionaries
- Tuples
- Functions
- Array
- Selection by position & Labels

Module 4: Python Packages

- Pandas
- Numpy
- Sci-kit Learn
- Mat-plot library

Module 5: Importing data

- Reading CSV files
- Saving in Python data
- Loading Python data objects
- Writing data to csv file

Module 6: Manipulating Data

- Selecting rows/observations
- Rounding Number
- Selecting columns/fields
- Merging data
- Data aggregation
- Data munging techniques

Module 7: Statistics Basics

- Central Tendency
- o Mean
- o Median
- o Mode
- o Skewness
- o Normal Distribution
- Probability Basics
- o What does mean by probability?

- o Types of Probability
- o ODDS Ratio?
- • Standard Deviation
- o Data deviation & distribution
- o Variance
- • Bias variance Trade off
- o Underfitting
- o Overfitting
- • Distance metrics
- o Euclidean Distance
- o Manhattan Distance
- • Outlier analysis
- o What is an Outlier?
- o Inter Quartile Range
- o Box & whisker plot
- o Upper Whisker
- o Lower Whisker
- o Scatter plot
- o Cook's Distance
- • Missing Value treatments
- o What is a NA?
- o Central Imputation
- o KNN imputation
- o Dummification
- • Correlation
- o Pearson correlation
- o Positive & Negative correlation

Module 8: Error Metrics

- • Classification
- o Confusion Matrix
- o Precision

- o Recall
- o Specificity
- o F1 Score
- • Regression
- o MSE
- o RMSE
- o MAPE

Machine Learning

Module 9: Supervised Learning

- • Linear Regression
- o Linear Equation
- o Slope
- o Intercept
- o R square value
- • Logistic regression
- o ODDS ratio
- o Probability of success
- o Probability of failure
- o ROC curve
- o Bias Variance Tradeoff

Module 10: Unsupervised Learning

- • K-Means
- • K-Means ++
- • Hierarchical Clustering

Module 11: Other Machine Learning algorithms

- • K – Nearest Neighbour
- • Naïve Bayes Classifier
- • Decision Tree – CART
- • Decision Tree – C50
- • Random Forest