€TRAINING

Introduction to Data Science

Introduction to Data Science

Introduction

A high-level overview of key Data Science disciplines is provided in this course. Along with an overview of frequent advantages, difficulties, and adoption problems, a fundamental grasp of data science from both a commercial and technological standpoint is given.

You will master the fundamentals of data science in this course, as well as how to use Python, a potent open-source tool. You will learn about fascinating ideas including exploratory data analysis, basic statistics, testing of hypotheses, tools for regression and classification modeling, and an introduction to machine learning.

Course Objectives

At the end of this course the participants will be able to:

- Data Science Tools & Technologies
- Statistics for Data Science
- Python for Data Science
- Exploratory Data Analysis
- Advanced Statistics & Predictive Modeling
- Optimize Model Performance
- Dimensionality Reduction
- Basics of Machine Learning

Targeted Audience

- A beginner who is interested in data science and want to learn basic data science skills
- Those looking for a more robust, structured data science learning program
- Data Analysts, Economists, or Researchers
- Software or Data Engineers

Course Outline

Unit 1: Foundation for Data Science & Probability & Statistics

- Introduction to Data Science
- Analytics Landscape
- Life Cycle of a Data Science Projects
- Data Science Tools & Technologies
- Measures of Central Tendency
- Measures of Dispersion
- Descriptive Statistics
- Probability Basics
- Marginal Probability
- Bayes Theorem
- Probability Distributions
- Hypothesis Testing

Unit 2: Basics of Python & Python Built-in Data Structures

- Install Anacond
- Data Types & Variables
- String & Regular Expressions
- Python list
- Python dictionaries
- Python set
- Python tuple
- Comprehensions

Unit 3: Control & Loop Statements in Python & Functions & Classes in Python

- For Loop
- While Loop
- Break Statement
- Next Statements
- Repeat Statement
- if, ifelse Statements
- Switch Statement
- Writing your own functions UDF
- Calling Python Functions
- Functions with Arguments
- Calling Python Functions by passing Arguments
- Lambda Functions
- Classes & Objects

Unit 4: Working with Data & Analyzing Data using Pandas

- Reading files with Python
- Writing files from Python
- Reading files using Pandas library
- Saving Data using Pandas library
- Clean & Prepare Datasets
- Manipulate DataFrame
- Summarize Data
- Churn Insights from Data

Unit 5: Visualize Data & Advanced Statistics & Predictive Modeling

- Charts using Matplotlib
- Charts using Seaborn
- Charts using ggplot
- ANOVA
- Linear Regression OLS
- Case Study: Linear Regression
- Principal Component Analysis
- Factor Analysis
- Case Study: PCA/FA
- Logistic Regression MLE
- Case Study: Logistic Regression

- K-Nearest Neighbor Algorithm
- Case Study: K-Nearest Neighbor Algorithm
- Decision Tree
- Case Study: Decision Tree

Unit 6: Time Series Forecasting & Introduction to Machine Learning

- Understand Time Series Data
- Visualizing TIme Series Components
- Exponential Smoothing
- Holt's Model
- Holt-Winter's Model
- ARIMA
- Case Study: Time Series Modeling on Stock Price
- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning
- Using Scikit-learn
- Scikit-learn classes
- Case Study: Machine Learning Algorithm