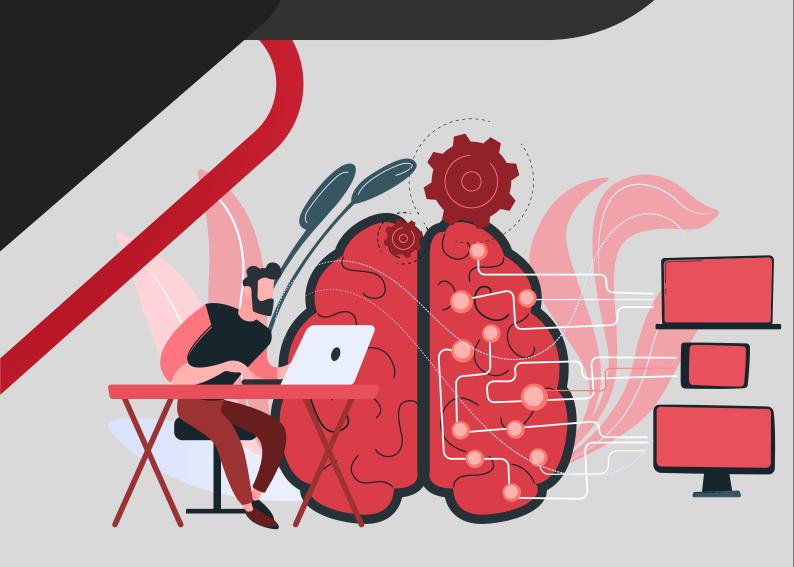


# Machine Learning



### **Introduction to Data Science**

- What is Data Science
- What does data science involves
- ▶ Life cycle of Data Science
- Tools of Data Science
- Introduction to Python

### Python environment Setup and Essentials

- > Introduction to python
- Software installation
- Basic operators and functions
- Data types with python
- Conditional statement

## Mathematical Computing with Python (Numpy)

- Introduction to Numpy
- Introduction to Numpy arrays
- ▶ How to Access Array Elements?
- Indexing, Slicing, Iteration, Indexing with Boolean Arrays
- Dealing with Flat files using Numpy

- Mathematical functions
- Statistical functions (mean, median, average, standard deviation)
- Operations with arrays

## Introduction to Scientific Computing (Scipy)

- Save a search as a report
- Editing reports
- Creating reports with visualizations charts and tables

### Data Manipulation with Pandas

- Introduction to Pandas
- Defining data structures
- Understanding Dataframes
- Importing Data from various sources
- ➤ (Csv, txt, excel etc)
- Missing values
- Data Operations
- File read operations
- Descriptive statistics

## Data Visualization using Matplotlib

- Create plots like scatter plot,
- > Histogram, bar graph, pie chart using Matplotlib
- Grids, axes, plots
- Markers, colour, fonts and styling.

### Data Visualization using Seaborn

- Create plots like scatter plot,
- Histogram, bar graph, pie chart using Seaborn
- Grids, axes, plots
- Markers, colour, fonts and styling.

## Machine learning using scikit-learn

- Machine learning Process Flow
- Machine learning categories
- > Feature selection and extraction in machine learning
- Supervised learning algorithms

#### Regression

#### Simple linear Regression

Applications of linear regression

Building regression models using python

Process to implement linear regression

Coefficient of determination (R- Squared)

Accuracy of model

#### Multiple linear Regression

#### Classification

Logistic Regression
Building Logistic Regression Model
Understanding standard model metrics
(Validation of Logistic Regression Models Standard
Business Outputs

#### Decision Tree

#### Random Forest

- Support Vector Machines
  - K NN
- Naïve Bayes classifier
  - K- nearest neighbor
- Ada Boost
- Gradient Boost
- XG Boost
- Model evaluation techniques concepts of confusion matrix, threshold evaluation with ROCR
- Unsupervised machine learning algorithms K-
- Means Clustering
- Hierarchical Clustering
- Aprior Algorithm

### **Web Scraping in Python**

- Working with Beautiful Soap
- Parsing HTML and XML
- Navigating the document
- Handling CSV files
- Parsing JSON into Python

### Introduction to Deep learning

## Assignment and Live Examples:

- Resumes helping you to create your resume.
- Case study-based approach.
- Placement Assistance.

### **GitHub**

- Creating a Git Account
- ▶ Cloning the repository
- Adding the file
- Committing the file
- Git push
- Removing the file