ITShare Company
01122339966-01122339955
Mostafa@iteshare.com

Data Science with Python

Module 1: Introduction to Data Science

- Basics of Data Science
- What and why Data Science
- Terminology

Module 2: Introduction to Machine Learning

- Basics of Machine Learning
- What and why Machine Learning
- Applications of Machine Learning
- Types of Machine Learning
- Main Challenges of Machine Learning

Module 3: Python

- Installation of Anaconda Python Distribution (contd.)
- Decision Control statements
- Data Types with Python
- Operators & Functions

Module 4: Scikit Learn

- Arrays.ipynb
- Array Reshaping
- Maths Functions

Module 5: Numpy

- Arrays.ipynb
- Array Reshaping
- Maths Functions

Module 6: Pandas

- Key Components
- DATA FRAME
- Accessing subset of data - rows, columns, filters

Module 7: Scipy

- Introduction to SciPy
- SciPy Sub Package
- SciPy Sub Package - Statistics and IO

Module 8: Matplotlib

- Plotting Histogram
- Bar Chart
- Pie Chart
- SubPlot

Module 9: Seaborn

- Numerical Data Plotting
- Line Plot
- Categorical Data Plotting
- Regression Plots
-

Module 10: Linear Algebra

- Vectors (2D,3D)
- Dot Product
- Hyperplane

- Square, Rectangle

Module 11: Probability

- Data types and its measures
- Random Variables,its application with variables
- Probability-Application with examples
- Probability distribution with examples
- Sampling Funnel-why And how

Module 12: Statistics

- WHAT IS STATISTICS
- BASIC TERMINOLOGIES IN STATISTICS
- TYPES OF STATISTICS
- DESCRIPTIVE STATISTICS
- MEASURE OF CENTRAL TENDENCY (Mean, median, mode)
- Measures of dispersion (Variance,Standard Deviation,Range-its derivation)
- INFERENTIAL STATISTICS

Module 13: Data pre-processing

- Is your data clean?
- What is Data Pre processing ?

Module 14: Data Cleaning

- Handling Missing data
- Handling Categorical data
- Data cleaning techniques
- Outliers

Module 15: Data Reduction

- Linear Discriminant Analysis LDA
- Principal Component Analysis PCA

Module 16: Exploratory Data Analysis (EDA)

- Introduction
- 2D Scatter-plot

- 3D Scatter-plot
- Univariate, Bivariate and Multivariate
- Histogram
- Box-plot
- Variance, Standard Deviation

Module 17: Feature Engineering

- Introduction
- Need for Feature Engineering in Machine Learning
- Steps in Feature Engineering
- Feature Engineering Techniques

Module 18: Feature Selection

- Filter
- Feature Shuffling
- Feature Importance

Module 19: Performance Metrics & Parameter Tuning

- Confusion Matrix
- ROC Curve
- Cross Validation in Machine Learning
- K fold Cross Validation & Grid search

Module 20: Supervised Learning - Regression I

- Linear Regression - Mathematical Intuition
- Programming of Linear Regression in Python-scikit learn

Module 21: Supervised Learning - Classification

- Multiple Linear Regression
- Multiple Linear Regression - Mathematical Intuition
- Polynomial Regression
- Support Vector Machines

Module 22: UnSupervised Learning - Clustering

- Difference between regression and classification
- Various Algorithms in Classification
- Logistic Regression

- Naive Bayes
- Ensemble Techniques

Module 23: UnSupervised Learning - Association Rule Mining

- Unsupervised Learning
- Types of Unsupervised Learning
- Applications of Unsupervised Learning
- Introduction to Clustering Algorithms
- Types of Clustering Algorithms
- What is K-Means Clustering?
- Implementation of K-Means Clustering
- Improving Models

Module 24: Natural Language Processing

- What is Association Rule Mining?
- Algorithms in Association Rule Mining
- Implementation of Apriori in Python

Module 25: Demo

- Natural Language
- PIPELINE
- Text Tokenization