



Python Programming with Data Analytics, Machine Learning and AI - Detailed Course Content

Python Environment Setup and Essentials

- Introduction to Python Language, features, the advantages of Python over other programming languages
- Python installation, Windows, Mac & Linux distribution for Anaconda Python, deploying Python IDE, basic Python commands, data types, variables, keywords and more.

Python Language Basic Constructs

- Built-in data types in Python, tabs and spaces indentation, code comment, character, variables and names
- Python built-in data types, Numeric, int, float, complex, list tuple, set dict, containers, text sequence, exceptions, instances, classes, modules, Str(String)
- Ellipsis Object, Null Object, Ellipsis, Debug, basic operators, comparison, arithmetic, slicing and slice operator, logical, bitwise, loop and control statements, while, for, if, break, else, continue.

NumPy for Mathematical Computing

- Introduction to NumPy arrays and matrices
- Indexing of Numpy array, datatypes, broadcasting of array math, standard deviation, conditional probability, correlation, and covariance.
- **Hands-on Exercise** – How to import NumPy module, creating an array using ND-array, calculating standard deviation on an array of numbers, calculating the correlation between two variables.

SciPy for Scientific Computing

- Introduction to SciPy and its functions, building on top of NumPy, cluster, linalg, signal, optimize, integrate, subpackages, SciPy with Bayes Theorem.
- **Hands-on Exercise** – Importing of SciPy, applying the Bayes theorem on the given dataset.

Pandas for Data Analysis

- Introduction to Python dataframes
- Importing data from JSON, CSV, Excel, SQL database
- NumPy array to dataframe, various data operations like selecting, filtering, sorting, viewing, joining, combining, how to handle missing values, time series analysis, linear regression.
- **Hands-on Exercise** – working on importing data from JSON files, selecting record by a group, applying filter on top, viewing records, analyzing with linear regression, and creation of time series.



Introduction to Machine Learning

- Need of Machine Learning
- Introduction to Machine Learning
- Types of Machine Learning – Supervised, Unsupervised and Reinforcement Learning.
- Why Machine Learning with Python.
- Applications of Machine Learning

Supervised Learning and Linear Regression

- Introduction to supervised learning
- Types of Supervised Learning – Regression & Classification
- Introduction to Regression, Simple Linear Regression, Multiple Linear Regression, Assumptions in Linear Regression, Math behind Linear Regression
- **Hands-on Exercise** – Implementing linear regression from scratch with python. Using Python library Scikit-Learn to perform simple linear regression and multiple linear regression. Implement train-test split and predict the values on the test set.

Classification and Logistic Regression

- Introduction to Classification
- Linear regression vs Logistic Regression
- Math behind Logistic Regression, detailed formulas, logit function and odds, confusion matrix and Accuracy, true positive rate, false positive rate, Threshold evaluation with ROC.
- **Hands-on Exercise** – Implementing logistic regression from scratch with python. Using Python library Scikit-Learn to perform simple logistic regression and multiple logistic regression. Building a confusion matrix, to find out the accuracy, true positive rate, and false positive rate.

Decision Tree and Random Forest

- Introduction to tree-based classification
- Understanding Decision Tree,
- Impurity Function – Entropy, understand the concept of information gain for right split of node, Impurity Function – Information gain
- Understand the concept of information gain for right split of node
- Impurity Function – Gini index, understand the concept of Gini Index for right split of node, overfitting & pruning, pre-pruning, post-pruning, cost-complexity pruning, Introduction to ensemble techniques, Understanding Bagging, Introduction to Random Forest, Finding the right number of trees in Random Forest.
- **Hands-on Exercise** – Implementing decision tree from scratch in Python. Using Python library Scikit-Learn to build a decision tree and random forest. Visualizing the tree and changing the hyperparameters in the random forest.



Naïve Bayes and Support Vector Machine

- Introduction to probabilistic classifiers
- Understanding Naïve Bayes
- Math behind Bayes theorem
- Understanding Support Vector Machine
- Kernel Functions in Support Vector Machine, Math behind svm.
- **Hands-on Exercise** – Using Python library Scikit-Learn to build Naïve Bayes Classifier and Support Vector Classifier.

Unsupervised Learning

- Types of Unsupervised Learning- Clustering and Dimensionality Reduction.
- Types of clustering, introduction to k-means clustering, the math behind k-means,
- Dimensionality reduction with PCA.
- **Hands-on Exercise** – Using Python library Scikit-Learn to implement K-means clustering. Implementing PCA on top of a dataset.

Introduction to Deep Learning

- Introduction to deep learning with neural networks
- Biological neural network vs Artificial neural network
- Understanding perceptron learning algorithm, introduction to deep learning frameworks
- TensorFlow-Constants, Variables and Place-holders

Machine Learning Projects

Project 1: Customer Churn Classification

Topics: This is a real-world project that gives you hands-on experience in working with most of the machine learning algorithms. The main components of the project include the following:

- Manipulating data to extract meaningful insights
- Visualizing data to find patterns among different factors
- Implementing these algorithms – linear regression, decision tree, naïve Bayes

Project 2: Recommendation for Movie, Summary

Topics: This is a real-world project that gives you hands-on experience in working with a movie recommender system. Depending on what movies are liked by a particular user, you will be in a position to provide data-driven recommendations. This project involves understanding recommender systems, information filtering, predicting ‘rating’, learning about user ‘preference’ and so on. You will exclusively work on data related to user details, movie details, and others. The main components of the project include the following:



- Recommendation for movie
- Two Types of Predictions – Rating Prediction, Item Prediction
- Important Approaches: Memory-Based and Model-Based
- Knowing User Based Methods in K-Nearest Neighbor
- Understanding Item Based Method
- Matrix Factorization
- Decomposition of Singular Value
- Data Science Project discussion
- Collaboration Filtering
- Business Variables Overview