

Certification in Data Science with Python

Module 1: Introduction to Python, its Sequences and File Operations

Objectives:

- At the end of this Module, you should be able to:
- Define Python
- Understand why Python is Popular
- Setup Python Environment
- Understand Operands and Expressions
- Write your First Python Program
- Understand Command Line Parameters and Flow Control
- Take Input from the User and Perform Operations on it
- Explain Numbers
- Explain Strings, Tuples, Lists, Dictionaries, and Sets

Topics:

- Overview of Python
- The Companies using Python
- Different Applications where Python is Used
- Discuss Python Scripts on UNIX/Windows
- Values, Types, Variables
- Operands and Expressions
- Conditional Statements
- Loops
- Command Line Arguments

- Writing to the Screen
- Python Files I/O Functions
- Numbers
- Strings and Related Operations
- Tuples and Related Operations
- Lists and Related Operations
- Dictionaries and Related Operations
- Sets and Related Operations

Module2: Deep Dive – Functions, OOPs, Modules, Errors and Exceptions

Objectives:

- At the end of this Module, you should be able to:
- Create and Execute Functions
- Learn Object Oriented Concepts
- Understand Python Standard Libraries
- Define Modules
- Handle the Exceptions

Topics:

- Functions
- Function Parameters
- Global Variables
- Variable Scope and Returning Values
- Lambda Functions
- Object Oriented Concepts
- Standard Libraries
- Modules Used in Python
- The Import Statements

- Module Search Path
- Package Installation Ways
- Errors and Exception Handling
- Handling Multiple Exceptions

Module3: Data Manipulation using NumPy, Pandas & Matplotlib

Objectives:

- At the end of this Module, you should be able to:
- Create Arrays using NumPy
- Perform Various Operations on Arrays and Manipulate them
- Read & Write Data from Text/CSV Files into Arrays and vice-versa
- Create Series and Data Frames in Pandas
- Indexing and Slicing of Data Structures in Pandas
- Reading and Writing Data from Excel/CSV Formats into Pandas
- Data Preparation
- Merging, Concatenation, Combining, Pivoting & Removal
- Data Transformation – Merging, Joining & Concatenation
- Create Simple Plots using Matplotlib
- Learn Different Plot Formats Available in Matplotlib
- Choose the Right Plot Format for a Problem at Hand Judiciously
- Scale and Add Style to your Plots

Topics:

- NumPy – Arrays
- Operations on Arrays
- Indexing, Slicing and Iterating
- Reading and Writing Arrays on Files

- Pandas – Data Structures & Index Operations
- Basic Functionalities of a Data Object
- Merging of Data Objects
- Concatenation of Data Objects
- Types of Joins on Data Objects
- Exploring a Dataset
- Analysing a dataset
- Reading and Writing Data from Excel/CSV Formats with Pandas
- The Matplotlib Library
- Grids, Axes, Plots
- Markers, Colours, Fonts and Styling
- Types of Plots – Bar Graphs, Pie Charts, Histograms
- Contour Plots

Module 4: Introduction to Machine Learning with Python

Objective:

At the end of this module, you should be able to:

- Essential Python Review
- Necessary Machine Learning Python libraries
- Define Machine Learning
- Discuss Machine Learning Use cases
- List the categories of Machine Learning
- Illustrate Supervised Learning Algorithms
- Identify and recognize machine learning algorithms around us
- Understand the various elements of machine learning algorithm like parameters, hyper parameters, loss function and optimization.

Topics:

- Python Revision (numpy, Pandas, scikit learn, matplotlib)
- What is Machine Learning?
- Machine Learning Use-Cases
- Machine Learning Process Flow
- Machine Learning Categories
- Linear regression
- Gradient descent

Module 5: Supervised Learning – I**Objective:**

At the end of this module, you should be able to:

- Understand What is Supervised Learning?
- Illustrate Logistic Regression
- Define Classification
- Explain different Types of Classifiers such as,
- Decision Tree
- Random Forest

Topics:

- What is Classification and its use cases?
- What is Decision Tree?
- Algorithm for Decision Tree Induction
- Creating a Perfect Decision Tree
- Confusion Matrix
- What is Random Forest?

Module 6: Dimensionality Reduction

Objective:

At the end of this module, you should be able to:

- Define the importance of Dimensions
- Explore PCA and its implementation
- Discuss LDA and its implementation

Topics:

- Introduction to Dimensionality
- Why Dimensionality Reduction
- PCA
- Factor Analysis
- Scaling dimensional model
- LDA

Module 7: Supervised Learning – II**Objective:**

At the end of this module, you should be able to:

- Understand What is Naïve Bayes Classifier
- How Naïve Bayes Classifier works?
- Understand Support Vector Machine
- Illustrate How Support Vector Machine works?
- Understand Hyper parameter Optimization

Topics:

- What is Naïve Bayes?
- How Naïve Bayes works?
- Implementing Naïve Bayes Classifier
- What is Support Vector Machine?

- Illustrate how Support Vector Machine works?
- Hyperparameter Optimization
- Grid Search vs Random Search
- Implementation of Support Vector Machine for Classification

Module 8: Unsupervised Learning

Objective:

At the end of this module, you should be able to:

- Define Unsupervised Learning
- Discuss the following Cluster Analysis
- K – means Clustering
- C – means Clustering
- Hierarchical Clustering

Topics:

- What is Clustering & its Use Cases?
- What is K-means Clustering?
- How K-means algorithm works?
- How to do optimal clustering
- What is C-means Clustering?
- What is Hierarchical Clustering?
- How Hierarchical Clustering works?

Module 9: Association Rules Mining and Recommendation Systems

Objective:

At the end of this module, you should be able to:

- Define Association Rules

- Learn the backend of recommendation engines and develop your own using python

Topics:

- What are Association Rules?
- Association Rule Parameters
- Calculating Association Rule Parameters
- Recommendation Engines
- How Recommendation Engines work?
- Collaborative Filtering
- Content Based Filtering

Module 10: Reinforcement Learning**Objective:**

At the end of this module, you should be able to

- Explain the concept of Reinforcement Learning
- Generalize a problem using Reinforcement Learning
- Explain Markov's Decision Process
- Demonstrate Q Learning

Topics:

- What is Reinforcement Learning
- Why Reinforcement Learning
- Elements of Reinforcement Learning
- Exploration vs Exploitation dilemma
- Epsilon Greedy Algorithm
- Markov Decision Process (MDP)
- Q values and V values
- Q – Learning

- values

Module 11: Time Series Analysis

Objective:

At the end of this module, you should be able to:

- Explain Time Series Analysis (TSA)
- Discuss the need of TSA
- Describe ARIMA modelling
- Forecast the time series model

Topics:

- What is Time Series Analysis?
- Importance of TSA
- Components of TSA
- White Noise
- AR model
- MA model
- ARMA model
- ARIMA model
- Stationarity
- ACF & PACF

Module 12: Model Selection and Boosting

Objective:

At the end of this module, you should be able to:

- Discuss Model Selection
- Define Boosting
- Express the need of Boosting

- Explain the working of Boosting algorithm

Topics:

- What is Model Selection?
- Need of Model Selection
- Cross – Validation
- What is Boosting?
- How Boosting Algorithms work?
- Types of Boosting Algorithms
- Adaptive Boosting