Xerxez Solutions
Cooperate Training Road Map
On
Python Programming for Artificial Intelligence and Data Science – Advance Level
(Corporate Training | Education Industry)

This document provides the curriculum outline of the Knowledge, Skills and Abilities that a **Machine Learning Developer** and **Research/Data Analyst** can be expected to demonstrate.

Prerequisite:

- Fundamentals of Python Programming, Data Structure and Object-Oriented Technique.
- Basic Knowledge of Git/GitHub, Jupyter Notebook and PyDoc.
- Fundamental Knowledge of Database and PgAdmin.
- Understanding of Visual Studio Framework.

Out Come:

After attending this training, the trainees will gain the below skills on Full Stack AI/ML Model Design, Development, Deploy & DevOps/MLOps Orchestration with end-to-end security using GitHub Enterprises.

- Advanced introduction to concurrent programming: subprocesses, multiprocessing, threads, and concurrent futures.
- Build 10 real-world Python programs using the professional object-oriented programming approach and use Git to track your code changes.
- Complete independent Python projects and compare the solutions and write highly organized modular code.
- Learn advanced core Python concepts such as class methods, static methods, abstract classes, database management and pyspark etc.
- Learn the PEP8 Style Guide for Python code, code refactoring and python packages.

Local setup (Physical Mode)	Remote Lab Setup (Optional)	GitHub Account
Laptop/Desktop with high-speed internet connection, Windows 10 and above	OS: Windows 10 and above	One Account
Memory: 4 GB RAM	Memory: 32 GB RAM	**Python Interpretor**
CPU: 1 CPU Cores	CPU: 8 CPU Cores	3.8 and Above
Storage: 20 GB	Storage: 500 GB SSD	VSCode

Topics Covered:

Module 1: Advanced Techniques for Accessing, Cleaning, Aggregating, and Visualizing Your Data

- Data Access: CSV, Table and Excel Reader, Accessing the database.
- Data Cleaning:
  - Row Filter,
  - Column Filter,
  - Data Manipulation (Number, String and Rules Engine, Math Formula, String manipulation and Rule Engine),
  - Column Expression Node
- Aggregation and Data Blending:
  - What is Data Aggregation.
  - Basic data aggregation using GroupBy node.
  - The pivoting node
  - The join operation and Joiner node
  - The concatenate node
- Data Visualization:
  - Data Explorer.
  - Scatter Plot.
  - 3 steps to build an interactive scatter plot.
- Data Export and Reporting:
  - CSV writer node
  - Practice reporting with BIRT

Hands-On Session: Data Access, Data Cleaning, Aggregation and Data Blending, Data Visualization, Export and Import Methods

Module 2: From Data Manipulation to Advanced Algorithms and Recommendation Systems
- **Date & Time & Databases**
  - String to Date & Time Node.
  - Extract Date & Time Fields Node.

- o Modify Date/Time/ Time zone nodes.
- o Moving average and moving aggregation Node.
- o Manipulating Data on databases.
- **Flow Variable and Components**
  - o Flow variables.
  - o From Data to Variables.
  - o Path Data Type and Flow Variables.
  - o What is a component?
  - o Component Configuration
  - o Sharing and Linking Components
- **Work Control**
  - o What is a loop?
  - o How to build a counting loop
  - o How to build generic loop
  - o Loop End Nodes
  - o Loop Commands
  - o Group Loop Start Node and Chunk Loop Start Node
  - o Deciding between Parallel Branches
- **PostgreSQL Database and PQAdmin**
  - o Database Connection
  - o Create Database and Table
  - o Insert Data, Select Data and Where Clause
  - o Order By, Update Table, Delete Data, Drop Table
  - o Limit, Join and Cursor Objects.
  - o Rest API

<mark>Hands On: Data and Time, Flow Variable and Components, Work Control and PostgreSQL Database</mark>

Module 3: Exploring the World of Machine Learning: From Supervised to Unsupervised Learning and Recommendation Systems

- **Introduction to Machine Learning**
  - o Supervised and Unsupervised Learning
- **Supervised Learning**
  - o Regression - Linear Regression, Multivariate Linear Regression, Root mean square error and mean absolute error, Pros and Cons
  - o Classification - Logistic regression, Setting up Threshold, Performance measures – Precision and Recall, Naïve Bayes, Evaluation of Model.
  - o K-nearest neighbour algorithm.
  - o Support Vector Machine.
- **Unsupervised Learning – Clustering**
  - o Types of Clustering - K – means Clustering, Importance of Scaling, Applications of Clustering, Advantages and Disadvantages of K-means clustering.
  - o Visual Analysis of Clustering.

- o Dynamic Clustering.
  - o Dimensionality reduction – PCA
- **Recommendation System**
  - o Content based filtering.
  - o Collaborative filtering.
  - o Similarity measures.
  - o The case of recommending songs.
  - o Hybrid Systems and other methods.

Module 4: From Data to Intelligence: A Comprehensive Guide to Machine Learning, NLP, and Computer Vision Techniques

- **Neural Network**
  - o Activation Functions
  - o Feed Forward Neural Network
  - o Topology of a Neural Network
  - o Error and Loss Function
  - o Training a Neural Network
  - o Optimization
  - o Gradient Descent and Backpropagation
  - o Learning Rate and Momentum
- **Decision Tree**
  - o Decision tree - CART
  - o CART - Example
  - o Pruning
  - o Ensemble Techniques
- **Random Forest**
  - o Bagging
  - o Boosting
  - o Boosting - A Visual Example
  - o Stacking
- **Natural Language Processing**
  - o Different tasks in NLP
  - o How are NLP problems solved?
  - o Text Extraction/Web scrapping
  - o Building a model
- **Computer Vision**
  - o Types of Computer Vision problems
  - o Pixel
  - o How does a computer see an image?
  - o 3D Images
  - o Resolution
  - o Image Transformations
  - o Convolution

- o Pooling
- o Significance of Convolution & Pooling
- o Convolutional Neural Networks

<mark>**Hands On:** Neural Network, Decision Tree, NLP and Computer Vision</mark>

Module 5: Introduction to PySpark and Big Data
- Introduction and Environment Setup
- SparkContext
- RDD
- Broadcast and Accumulator
- SparkConf and SparkFile
- Storage Level
- MLLib
- Serializer

<mark>**Hands On:** Environment Setup and Fundamentals of PySpark</mark>

CONTACT US

Email: info@xerxez.in
WhatsApp: 9164315460
Website: www.xerxez.in