ALTALUNE TECHNOLOGY

DATA SCIENCE & MACHINE LEARNING WITH PYTHON

Programming with Python

- Day 1
  - All About Python
    - Origin
    - History
    - scope of python
    - popularity of python
    - future of python
    - use of python in different emerging technologies of modern era
    - achievements of python
    - mega projects hosted in python
  - Python Installation and Environment Setup
    - Different Flavours of python like Ipython, Jpython, Cpython, PYPY etc
    - Development Environment and it's Requirements
    - Installation on windows
    - Installation on linux
    - Installation on mac
    - Choosing best IDE for your coding purpose from IDLE, Jupyter, Spyder, PyCharm, vscode, vim
  - Writing first program in python
    - Using vim and command line creating first hello program in python
    - Using vscode or PyCharm for Integrated Development Environments
    - Using Jupyter Notebook and it's features for fast, simple learning process
  - Syntax of python
    - Python coding Style
    - Indentation
    - Comments in python
    - Snake style coding

- Keywords in python
- Built-in functions in python

- Day 2
  - Data type and Data Structures
    - Numbers
    - Strings
    - List
    - Dictionary
    - Tuples
    - Sets
    - Frozen Sets
    - Type Casting
- Day 3
  - Advance Containers from collection module
    - Queues
    - Ordered Dictionary
    - Named Tuple
    - Default Dictionary
    - Counter
    - ChainMap
    - UserDict
  - Control Statements
    - Flow diagrams
    - If – else
    - Nested if else
- Day 4
  - Looping in Python
    - For Loop
    - While Loop
    - Break
    - Continue
    - Else with loop
  - Functions in Python
    - Built-in Functions zip, enumerate, eval, exec, min, max, ord, chr, sorted, reversed, len, sum, power, divmod etc.

- Defining Custom Functions
- Function Calling
- Scope Resolution global, local, nonlocal scope
- Code Reusability
- Recursion

- Day 5
  - Advance Functions in Python
    - Lambda anonymous function
    - Map function
    - Reduce function
    - Filter function
    - Lazy Evaluation
  - Closures and Decorators in Python
    - Memorization using closures
    - Writing extensible functions using Decorators
    - Decorators to implement oops like functionality to functions
- Day 6
  - OOPS
    - Encapsulation
    - Abstraction
    - Data hiding
    - Data Security
    - Access Control
    - Inheritance
    - Polymorphism
    - Overriding
    - Operator Overloading
    - Classes
    - Objects
    - Implementing Data Structures using OOPs
    - Message Passing and Share Space in OOPs
  - Advance OOPS
    - Name Mangling in OOPs
    - MRO (Method Resolution Order)
    - Meta Classes

        - Slots
        - Properties
        - Class methods
        - Static methods
- Day 7
    - Implementation of Data Structures in Python
        - Implementing Stack
        - Implementing Queue
        - Implementing Link-List
    - Generators & Iterators in Python
        - Memory Optimization Techniques
        - Lazy Evaluation
        - Custom Generators & Decorators
        - Zip like generator objects and their working
- Day 8
    - Exception Handling
        - Exceptions and Errors
        - Built in Exceptions
        - Handling Exceptions
        - else keyword with exception
        - finally, keyword to define Clean Up Actions
        - Custom Exception using raise keyword
        - assert keyword for assertions
        - Creating Custom Exception Class
    - File Handling
        - Type of file formats and their significance
        - File Creation and writing data to files
        - Reading data from a file
        - Overwriting files
        - Dealing with excel sheets, csv files
        - Making data persistent to create real life projects
- Day 9
    - Data Serialization
        - Serialization and De-Serialization

- Serializing python objects using pickle, json and shelve modules
- Deserializing Object State to read data from byte file or from network
- Storing Custom Objects to make state machines

- Installing Third Party Modules in Python
  - Pip (Python Package Installer)
  - Installing packages and modules using pip
  - pypi (python package index) repository for package lookup
  - Offline installation of a package using pip
- Database Connectivity
  - Concept of Data Base Management Systems
  - Using sqlite3 to store lite data in database like format
  - CURD (Create, Update, Read, Delete) Operation on sql table
  - Connecting python to a MariaDB or MySQL server using pymysql or mysql-client modules
  - Un-Structure Data Base Management systems like MongoDB

➢ Day 10

- Debugging and Standard Coding in Python
  - Code Analysis using pylint
  - Testing Code using Pytest scripts
  - Debugging in Python using pdb (python debugger)
  - Compiling code to generate .pyc files for faster run
  - Converting python code into object code .exe files to run on windows directly
- Modules and Packages in Python
  - Module name space
  - If __name__ == “__main__” in Python
  - Defining Custom Modules
  - Creating Packages in Python

- Testing Packages
- Importing and using custom packages
- Adding custom modules & packages to PYTHONPATH to use them like standard modules

- Standard Library in Python
  - Os Module for general interaction with Operating System
  - Understaning path submodule of os and it's significance
  - Shutil Module to copy, paste and delete files using python scripts
  - Sys modules for accessing command line arguments, standard input & output streams
  - Subprocess Modules to execute system commands and access their output in python
  - Paramiko module to run commands on remote systems
  - Random module to generate random data
  - Itertool and functool modules for functional programming
  - Zlib for data compression
  - Csv reader
  - Time module to deal with delays and time management in python

- Day 11
  - Graphical User Interface using Tk
    - GUI basics
    - Root window
    - Widgets like button, label, frame, canvas, radio buttons, check buttons, dropdown menus, scroll bars, progress bars, message box, text box, entry widget etc
    - Geometry Managers like place, grid and pack to stich widgets to main frames or root windows
    - Dialog box, top level window implementation
    - Using databases and threading in GUI applications
  - Virtual Environment
    - What is Virtual Environment
    - Difference between Production and Development Environment

- Difference between python and package versions and their importance in projects
- Creating a separate environment for Python Projects
- Activating Virtual Environment
- Installing Different Version of packages in different environments
- Requirement.txt file and it's importance
- README file

- Day 12
  - Text Processing
    - Regular Expressions
    - Compiling re patterns
    - Grouping re patterns
    - Finding all patterns in a text file
    - Parsing data using regular expressions
    - Re flags
  - APIs
    - Application Programming Interfaces and it's use in modern technology
    - Accessing Google place APIs to find any place information and photos
    - Accessing Twitter APIs for data Analysis Purpose
    - Accessing Facebook Graph API to fetch data
  - Web Scraping
    - Requests module in python
    - Fetching a page using requests modules
    - Looking up into headers and content of a page
    - Parsing html data using Beautiful Soup Module
    - Creating a soup and finding information from whole html content

- Day 13
  - GitHub
    - Version Control Systems and their significance

- Installing and configuring Git
- Creating a git repository
- Managing project using git add, commit, log, stash, reset commands
- Distributed systems
- Deploying Central Git repository for team projects
- git fetch, git pull, git push to manage your code remotely and UpToDate with central git repository
- creating branches in git for continuous development
- Discussing CDI (Continuous Development integrations) pipeline using git

- Flask Web Framework of Python
  - Installation and Configuration of Flask
  - Hello World to web using Flask
  - Templates
  - Jinja2 Template Rendering (Include and Extending)
  - Handling GET and POST requests in Flask
  - Integrating HTML, CSS, Flask together
  - Web Forms
  - Cookies and Sessions in Flask
  - Database
  - User Logins
  - Profile Page
  - Ajax in Flask
  - Testing and Debugging
  - Deployment on Linux

- Day 14
  - Data Collection & DBMS (Principle, Tools & Platforms)
    - Database Concepts (File System and DBMS)
    - Database Storage Structures (Tablespace, Control files, Data files)
    - Structured and Unstructured data
    - SQL Commands (DDL, DML & DCL),
    - Data ware Housing concept
    - No-SQL
    - Data Models - XML, working with MongoDB),

- Tools - OLTP and OLAP
- data preparation and cleaning techniques

➢ Day 15 & 16
- Statistical Analysis with Python
  - Introduction to Statistics
  - Descriptive Statistics
  - Summary Statistics
  - Basic probability theory
  - Statistical Concepts (uni-variate and bi-variate sampling, distributions, re-sampling, statistical Inference, prediction error)
  - Probability Distribution (Continuous and discrete- Normal, Bernoulli, Binomial, Negative Binomial, Geometric and Poisson distribution)
  - Bayes' Theorem
  - Central Limit theorem
  - Data Exploration & preparation
  - Concepts of Correlation,
  - Regression,
  - Covariance
  - Outliers etc.

➢ Day 17 & 18
- Data Visualization – Analysis & Reporting
  - Information Visualization
  - Data analytics Life Cycle
  - Analytic Processes and Tools
  - Analysis vs. Reporting
  - Modern Data Analytic Tools
  - Visualization Techniques
  - Visual Encodings
  - Exploratory Data Analysis –Visualization and Exploring Data
  - Interactive visualization
  - Visual Analytics
  - Dashboard Design

➢ Day 19

- Practical Machine Learning
  - Supervised and Unsupervised Learning
  - Uses of Machine learning
  - Future Scope Machine Learning
  - Understanding Representation, Evaluation and Optimization of Models
- Supervise Machine Learning – Regression
  - Implementation of Linear Regression from Scratch
  - Custom Evaluation of Linear Regression
  - Multiple Linear Regression
  - Polynomial Linear Regression
  - Project: Predictive Models using Regression

➢ Day 20

- Supervise Machine Learning – Classification
  - Classification Technique in Machine Learning
  - Implementation of Logistic Regression
  - Evaluation of Logistic Regression Model
  - Pro & Cons of Logistic Regression Algorithms
  - Introduction to Classification Metrices
  - Classification Report and Confusion Matrix
  - Normalization, Feature Selection & Feature Scaling to Optimize Model
  - Image Classification using Logistic Regression
  - Project: Classification Models using Logistic Regression

➢ Day 21

- Supervise Machine Learning – Recommendation
  - Distance Formulas such as Euclidian, Manhattan, Makowski
  - Centroids and Nearest Neighbors
  - K-Nearest Neighbor Model Implementation
  - Elbow method to Determine Best k-value
  - Bag of Words, TF/IDF
  - Text Preprocessing to convert a text to vectors
  - Project: Movie Recommendation Project using K-Nearest Neighbors

➢ Day22

- Supervise Machine Learning – Classification
  - Decision Tree Algorithms and Approaches

- Gini Indexing Algorithms
- ID3 Algorithm for Selecting Root Node
- Implementation of Decision Tree Algorithms
- Evaluation of Logistic Regression Model
- Random-Forest Techniques
- Cross Validation Techniques to Optimize Models
- Project: Fraud Detection Model using Logistic Regression

➢ Day23
  - Supervise Machine Learning – Classification
    - Naïve Bays Classifiers
    - Building a Model using Naïve Bays Classifiers
    - Evaluation and Optimization of Naïve Bays Model
    - Project: Spam Mail Detection using Naïve Bays
    - Understanding Concept of Support Vector Machines
    - Implementation of Support Vector Models for Classification
    - Evaluation and Optimization of SVM models
    - Project: Breast Cancer Classification using SVM Models

➢ Day 24
  - Unsupervised Machine Learning – Clustering
    - What is Clustering and techniques to cluster data
    - Latent Semantic Analysis using Singular Value Decomposition
    - Sentiment Analysis
    - K- means clustering
    - Evaluation & Optimization of K-means Algorithms
    - Project: Topic Modeling using K-means Model
    - Hierarchical Clustering

➢ Day 25 & 26
  - Deep Learning
    - Neural Networks and It's Application
    - Layers and Hyper Parameters
    - Neurons and Activation Functions (SoftMax, sigmoid, relu, tanh)
    - Back Propagation
    - Underfitting & Overfitting
    - K-Fold Cross Validation

- Early Stopping Techniques
- Loss Functions
- Initialization Functions (Xavier Initialization)
- Optimization Techniques
- Gradient Descent
- Stochastic Gradient Descent
- Momentum and Learning Rate
- AdaGrad and RMS Prop Algorithm
- Implementation of Neural Networks for Regression
- Implementation of Neural Networks for Classifiation
- Project: Handwritten Digit Recognition using NN

➢ Day 27 & 28
  - Big Data
    - Distributed Storage and Computing
    - Hadoop Architecture and Implementation
    - Big Data tools - HDFS, SQOOP, FLUME, MapReduce, Hive
  - PySpark
    - Overview
    - Linking with Spark
    - Initializing Spark
    - Resilient Distributed Datasets (RDDs)
    - External Datasets
    - RDD Operations
    - Passing Functions to Spark
    - Working with Key-Value Pairs
    - Shuffle operations
    - RDD Persistence
    - Removing Data
    - Shared Variables
    - Deploying to a Cluster
    - Project on spark using MLib

➢ Day 29 & 30
  - Project Discussion
  - Doubt Solving
  - Future Scopes
  - Case Studies on Data science & Machine learning

- Introducing other tools and techniques
  - R programming
  - Tableau
  - PowerBi
  - ML-Ops