CURRICULUM

AI BOOTCAMP

Module : Python for AI and ML

Installation

- Introduction to the Python programming language and its applications
- Setting up the Python environment: installation of Python and necessary libraries

Configuring the development environment: IDEs, text editors, and Jupyter Notebook

Python Basics I

- Introduction to Python: history, features, and advantages
- Expressions and operators: arithmetic, assignment, comparison, and logical
- Understanding type() function and type inference
- Introduction to data structures: lists, tuples, and dictionaries

Python Basics II

- Recap of Python basics
- Working with arithmetic operators: addition, subtraction, multiplication, division, modulus, and exponentiation
- Using comparison operators: equal to, not equal to, greater than, less than, etc.
- Logical operators: and, or, and not
- Exploring advanced data types: sets and strings manipulation.

Expressions, Conditional Statements & For Loop

- Evaluating expressions: operator precedence and associativity
- Introduction to conditional statements: if, elif, and else
- Executing code based on conditionals.
- Understanding the flow of control in conditional statements
- Iteration using the for loop: range(), iteration over lists, and strings.

While loop, Break and Continue Statements, and Nested Loops

- Working with while loop: syntax, conditions, and examples

- Combining loops and conditionals
- Using the break statement to exit loops prematurely.
- Utilizing the continue statement to skip iterations.
- Implementing nested loops for complex iterations

Functions

- Introduction to functions: purpose, advantages, and best practices
- Defining and calling user-defined functions
- Parameters and arguments: positional, keyword, and default values
- Return statement and function output.
- Variable scope and lifetime
- Function documentation and code readability

Exception Handling and File Handling

- Understanding exceptions: errors, exceptions, and exception hierarchy
- Handling exceptions using try-except blocks: handling specific exceptions, multiple exceptions, and else and finally clauses.
- Raising exceptions and creating custom exception classes
- File handling in Python: opening, reading, writing, and closing files.
- Working with different file modes and file objects

Python Modules: NumPy,Pandas and Matplotlib

- Introduction to the NumPy module: features and applications
- Working with multidimensional arrays: creation, indexing, slicing, and reshaping
- Performing element-wise operations: arithmetic, logical, and statistical
- Overview of the Matplotlib module: data visualization and plotting
- Customizing plots: line properties, markers, colors, labels, and legends

Advanced Topics

- Introduction to Kaggle platform: features and benefits
- Leveraging Kaggle for real-life datasets: data exploration, analysis, and visualization
- Introduction to machine learning modules on Kaggle: scikit-learn, TensorFlow, and PyTorch
- Overview of running machine learning experiments on Kaggle
- Resources for further learning and exploration

Module : Exploratory Data Analysis (EDA) and Machine Learning

Introduction and Missing Value Analysis

- Introduction to Exploratory Data Analysis (EDA)
- Importance of EDA in data analysis
- Steps involved in EDA
- Handling missing values: identification, analysis, and treatment strategies • Imputation techniques for missing values

Data Consistency, Binning, and Outlier Analysis

- Data consistency checks using fuzzy logic
- Binning and discretization techniques for continuous variables
- Outlier detection and analysis methods
- Handling outliers: techniques for treatment or removal

Feature Selection and Data Wrangling

- Importance of feature selection in EDA
- Feature selection techniques: filter methods, wrapper methods, and embedded methods
- Data wrangling: cleaning and transforming data for analysis
- Handling categorical variables: encoding techniques

Inference, Hypothesis Testing, and Visualization

- Inference and hypothesis testing in EDA
- Common statistical tests: t-test, chi-square test, ANOVA, etc.
- Visualization techniques for EDA: histograms, box plots, scatter plots, etc.
- Hands-on practical session for complete EDA using a dataset

Machine Learning Performance Metrics and Naive Bayes

- Evaluation metrics for classification problems: accuracy, precision, recall, F1 score, etc.
- Introduction to Naive Bayes algorithm and its applications
- Implementing Naive Bayes for classification tasks

Logistic Regression, SVM, Decision Trees, and Random Forests

- Logistic Regression: theory, interpretation, and applications
- Support Vector Machines (SVM): concepts, kernels, and use cases
- Decision Trees: construction, pruning, and interpretability
- Random Forests: ensemble learning and feature importance
- Bagging and Boosting: techniques for improving model performance

Hyperparameter Tuning, PCA, and SVD

- Hyperparameter tuning techniques: grid search, random search, and Bayesian optimization
- Principal Component Analysis (PCA): dimensionality reduction and feature extraction
- Singular Value Decomposition (SVD): applications in matrix factorization and data compression

Clustering Introduction, Partitioning Algorithms, and Cluster Evaluation

- Introduction to clustering: unsupervised learning technique
- Partitioning algorithms: K-means, K-medoids
- Hierarchical clustering: agglomerative and divisive approaches
- Density-based clustering: DBSCAN, OPTICS
- Cluster evaluation metrics: silhouette coefficient, Davies-Bouldin index

Regression and Evaluation of Regression Methods

- Introduction to regression analysis
- Linear regression: assumptions, interpretation, and model evaluation • Evaluation metrics for regression: mean squared error, R-squared, etc.
- Other regression methods: polynomial regression, ridge regression, lasso regression

<u>Module : NLP,LLMs, Computer Vision And MLOps</u>

Introduction to Natural Language Processing (NLP) and Large Language Models (LLMs)

- Overview of Natural Language Processing (NLP)
- Evolution of Large Language Models (LLMs)
- Importance and Applications of NLP and LLMs

Fundamentals of NLP

- Linguistic Concepts
- Tokenization and Text Preprocessing
- Part-of-Speech (POS) Tagging

- Named Entity Recognition (NER)
- Sentiment Analysis
- Text Classification
- Word Embeddings and Language Representations

Introduction to Large Language Models

- The Transformer Architecture
- Attention Mechanisms
- GPT, BERT, and Other Key Models
- Pretraining and Fine-Tuning Techniques
- Evaluation Metrics and Benchmarks

Practical Applications of NLP and LLMs

- Chatbots and Conversational AI
- Text Summarization
- Machine Translation
- Content Generation and Creative Writing
- Question Answering Systems
- Semantic Search and Text Mining

Ethical Considerations and Challenges

- Bias and Fairness
- Privacy and Security
- Model Interpretability and Explainability
- Environmental Impact and Computational Requirements

Hands-On Exercises

- Getting Started with NLP Libraries (spaCy, NLTK, Hugging Face Transformers)
- Building a Simple Text Classifier
- Fine-Tuning a Large Language Model for a Specific Task
- Evaluating Model Performance and Error Analysis
-

Future Trends and Opportunities in NLP and LLMs

- Multimodal Models and Human-AI Interaction
- Low-Resource Languages and Transfer Learning
- Knowledge-Enhanced Language Models
- Efficient Training and Deployment Techniques

Computer Vision

- Cascade and HOG classifiers to detect faces
- Face detection using OpenCV and Dlib library
- Detect other objects using OpenCV, such as cars, clocks, eyes, and full body of people

- KCF and CSRT algorithms to perform object tracking
- convolutional neural networks and implement them using Python and TensorFlow
- Detect objects in images in videos using YOLO, one of the most powerful algorithms today
- Recognize gestures and actions in videos using OpenCV
- Create hallucinogenic images with Deep Dream
- Create images that don't exist in the real world with GANs (Generative Adversarial Networks)

Reinforcement Learning

- Fundamentals of Reinforcement Learning
- Sample-based Learning Methods
- Prediction and Control with Function Approximation

Stable Diffusion Models

- Fundamentals of Diffusion Models
- Stable Diffusion in Practice
- Methods, Jobs and Tools of Stable Diffusion

Machine Learning Operations(MLOps)

- **Introduction to MLOps and AI/NLP Fundamentals**
    - Overview of MLOps and its importance in the AI lifecycle
    - Current trends in AI
    - Setting up the development environment

- **Deep Dive into Machine Learning Models for NLP**
    - Understanding NLP models (llama2, GPT, Mistral, etc.)
    - Introduction to Hugging Face Transformers and Datasets
    - Hands-on: Building a simple NLP model with Hugging Face

- **Introduction to FastAPI for ML Model Deployment**
    - Basics of API development with FastAPI
    - Deploying a simple ML model with FastAPI
    - Hands-on: Creating your first ML API with FastAPI

- **Advanced FastAPI Features for Production-Ready APIs**
    - Authentication and authorization in FastAPI
    - Hands-on: Enhancing your ML API with advanced features

- **Introduction to Docker for AI Applications**
    - Basics of Docker and containerization
    - Building Docker images for AI/ML applications
    - Hands-on: Containerizing your FastAPI application

- **Leveraging Lang Chain and LangSmith for Enhanced NLP Applications**
    - Introduction to Lang Chain and its Ecosystem
    - Overview of LangSmith for debugging, testing, evaluating, and monitoring LLM applications
    - Hands-on: Integrating Lang Chain with your NLP models and using LangSmith for enhanced capabilities
- **Advanced Model Deployment with Hugging Face and Lang Chain**
    - Integrating Hugging Face models for advanced NLP capabilities
    - Exploring Lang Chain for building complex NLP applications
    - Hands-on: Deploying a Hugging Face model via FastAPI with LangSmith integration

- **Deploying ML Models on Google Cloud**
    - Overview of Google Cloud Platform (GCP) for ML
    - Introduction to Google Cloud Run
    - Hands-on: Deploying your Dockerized FastAPI application on GCP with LangSmith monitoring