

Generative AI with Machine Learning Fundamentals

Total Duration: 07 days

Labs: Open-Source platform and Koenig DC will be provided

Section I: Machine Learning Fundamentals

Section Duration: 16 hours

Module 1: Introduction to Python for Machine Learning (0.5 hour)

- Overview of Python
 - Why Python for machine learning?
 - Key features and benefits
- Setting up the Environment
 - Installing Python
 - Setting up Jupyter Notebook/Google Colab
- Basic Syntax
 - Variables and data types
 - Basic operations

Module 2: Python Data Structures (1 hour)

- Lists, Tuples, and Dictionaries
 - Creating and manipulating lists
 - Understanding tuples
 - Using dictionaries
- Operations on Data Structures
 - Indexing and slicing
 - Iterating through data structures
 - Common methods and functions

Module 3: Control Flow and Functions (1 hour)

- Control Flow
 - If-else statements
 - Loops (for, while)
- Functions
 - Defining and calling functions
 - Parameters and return values
 - Lambda functions

Module 4: Introduction to NumPy (1 hour)

- NumPy Basics
 - Importance of NumPy in machine learning
 - Creating arrays
- Array Operations
 - Indexing and slicing
 - Mathematical operations on arrays
- Advanced NumPy
 - Broadcasting
 - Working with multi-dimensional arrays

Module 5: Data Manipulation with pandas (1 hour)

- Introduction to pandas
 - Importance of pandas in data science
 - DataFrames and Series
- DataFrame Operations
 - Loading and inspecting data
 - Indexing and selecting data
 - Handling missing values
- Advanced pandas
 - Grouping and aggregation
 - Merging and joining DataFrames

Module 6: Data Visualization (1 hour)

- Introduction to Matplotlib
 - Plotting basics
 - Creating different types of plots (line, bar, histogram)
- Introduction to Seaborn
 - Statistical data visualization
 - Creating attractive and informative visualizations

Module 7: Introduction to Machine Learning Concepts (0.5 hour)

- What is Machine Learning?
 - Definition and types (supervised, unsupervised, reinforcement)
 - Key concepts and terminology
- Scikit-Learn Basics
 - Overview of the library
 - Key features and benefits
- Building a Simple Model
 - Data preparation
 - Training a model
 - Making predictions
 - Evaluating the model
- Discussion about other frameworks

- TensorFlow
- PyTorch
- Transformers
- GAN
- Diffusers

Module 8: Mini Project (1 hour)

- Mini Project
 - Choose a simple machine learning problem (classification, regression or clustering)
 - Provide the student a dataset based on the problem selected
 - Demonstrate and share the solution based on his choice
 - Guide students for building a ML model (homework if time is constraint)

=====

Section II: Generative AI

Section Duration: 05 days (40 hours)

Module 9: Introduction of GenAI

- Introduction to Generative AI
- Architecture of Generative AI
- Applications of Generative AI using Transformer Library
- Introduction to Generative Adversarial Networks (GANs)
- Labs

Module 10: Introduction of Large language Model

- Architecture of Large Language Models
- Types of Large Language Models (LLMs)
- **Task based Text AI LLMs -**
 - Translation, Summarization, Sentence Similarity, Automatic Speech Recognition, Text to Speech, etc.
- **Major Text AI LLMs -**
 - Llama, Qwen, Cohere, Falcon LLM
- **Image AI Models-**
 - Object Detection, Image Segmentation, Image Retrieval, Image, Image Captioning, Visual QnA, Zero-shot Image Classification, etc.
- Labs

Module 11: Learning Prompt Engineering using Open-Source Models

- Introduction to Prompt Engineering
- Prompt Engineering Techniques
- Text Prompting using Llama (Meta)
- Image Prompting using Llama (Meta)
- Code Prompting using Llama (Meta)

- Labs

Module 12: Basic LLM Systems (RAG) using Open-Source Models

- Introduction to Retrieval Augmented Generation (RAG)
- Introduction to LangChain
- Concept of Embedding, Retrieval, Chain and Agents using LangChain
- Lab: Build a Simple LLM Application using LangChain
- Lab: Build a Chatbot LangChain
- Lab: Build vector stores and retriever using LangChain
- Lab: Build an Agent LangChain
- Lab: Build a Retrieval Augmented Generation (RAG) App using LangChain
- Lab: Build a Conversational RAG Application using LangChain

Module 13: Advanced LLM Systems (QnA) using Open-Source Models

- Difference between RAG & Question Answering system
- Build a Question Answering system over Tabular Data using LangChain
- Build a Question/Answering system over SQL data using LangChain
- Labs

Module 14: Fine-tuning Techniques using Open-Source Models

- Introduction to Quantization
- Optimization of model weights (data types)
- Modes of Quantization
- Fine tuning LLMs (Meta's Llama / Alibaba's Qwen / Google's Gemma)
- Labs

Module 15: Evaluation of Open-Source Models using MLflow

- Introduction to MLflow
- Build a machine learning model using MLflow
- MLflow Deployment Servers
- LLM Evaluation using MLflow
- Lab: Evaluate a Hugging Face LLM

=====