

Generative AI Specialty

Course Duration: 40 Hours (5 Days)

Overview

Generative AI Specialty Course Overview Our Generative AI Specialty course is designed to provide a comprehensive understanding of generative AI technologies over a 5-day (40-hour) period. Participants will begin with an introduction to Generative AI, including architectures and applications like GANs. The course covers Large Language Models (LLMs), focusing on their architecture, types, and varied applications such as text translation and image captioning. Prompt Engineering includes effective techniques for text, image, and code prompting. Advanced labs explore LangChain for creating LLM systems like Retrieval-augmented generation and Question-answering systems. The course also focuses on Fine-tuning techniques and Evaluation using MLflow. Practical labs ensure hands-on experience, empowering attendees to apply learned concepts in real-world scenarios. Pre-requisite: Fundamentals of Python (Machine learning knowledge is an added advantage).

Audience Profile

This comprehensive 5-day course is designed for professionals and enthusiasts aiming to master Generative AI, with a focus on open-source platforms and hands-on labs. It is ideal for individuals with a background in Python and machine learning.

- Data Scientists
- Machine Learning Engineers
- AI Researchers
- Software Developers
- IT Professionals
- Data Analysts
- AI Enthusiasts
- Python Programmers
- Deep Learning Specialists
- Technical Leads and Managers
- UX/UI Designers with an interest in AI
- Automation Engineers
- Research Scholars
- AI Product Developers

Course Syllabus

Module 01: Introduction to Generative AI

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

Module 02: Introduction to Large Language Models (LLMs)

- 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 & Services: Object Detection, Image Segmentation, Image Retrieval, Image Captioning, Visual Q&A, Zero-shot Image Classification, etc.
- Labs

Module 03: Learning Prompt Engineering Using Open-Source Models

- Introduction to Prompt Engineering
- Prompt Engineering Techniques
- Text Prompting with LLaMA (Meta)
- Image Prompting with LLaMA (Meta)
- Code Prompting with LLaMA (Meta)
- Labs

Module 04: Basic LLM Systems (RAG) Using Open-Source Models

- Introduction to Retrieval-Augmented Generation (RAG)
- Overview of LangChain
- Concepts of Embeddings, Retrieval, Chains, and Agents with LangChain
- Labs:
- Build a Simple LLM Application with LangChain
- Create a Chatbot with LangChain
- Develop Vector Stores and Retrievers using LangChain
- Build an Agent with LangChain
- Create a Retrieval-Augmented Generation (RAG) Application with LangChain
- Develop a Conversational RAG Application with LangChain

Module 05: Advanced LLM Systems (Q&A) Using Open-Source Models

- Difference Between RAG and Question-Answering Systems
- Build a Question-Answering System Over Tabular Data with LangChain
- Build a Question-Answering System Over SQL Data with LangChain
- Labs

Module 06: Fine-Tuning Techniques Using Open-Source Models

- Introduction to Quantization
- Optimization of Model Weights (Data Types)
- Modes of Quantization
- Fine-Tuning LLMs (e.g., Meta's LLaMA, Alibaba's Qwen, Google's Gemma)
- Labs

Module 07: Evaluation of Open-Source Models Using MLflow

- Introduction to MLflow
- Build a Machine Learning Model with MLflow
- MLflow Deployment Servers
- LLM Evaluation with MLflow
- Lab: Evaluate a Hugging Face LLM