



Training 4 Programmers

To discuss this course and customizations:

Call: +1 434-509-5680 or Email: sales@training4programmers.com

Hands-On Generative AI with Azure

Duration

3 days

Description

This course provides a comprehensive exploration of Generative AI and Large Language Models (LLMs), focusing on their architectures, applications, and ethical considerations. Participants will learn to harness the power of transformers, parameter-efficient fine-tuning, and advanced prompt engineering techniques, with a special emphasis on the LangChain framework. The course also delves into the practical aspects of developing AI-driven applications, such as chatbots, using Microsoft Azure and integrating vector databases for efficient data retrieval. Through hands-on projects and case studies, participants will gain the skills necessary to build, deploy, and optimize AI models and applications, while adhering to responsible AI practices.

Objectives

- Understand the fundamentals of Generative AI, Large Language Models (LLMs), and Azure OpenAI services.
- Explore different generative AI models, including Google's Gemini, ChatGPT, and Bing GPT-4, with a focus on text-based engines.
- Analyze ethical considerations in AI, such as bias, fairness, transparency, and the importance of responsible AI practices.
- Dive deep into LLM architectures, including transformers and techniques like parameter-efficient fine-tuning and prompt engineering.
- Develop skills in prompt engineering, including the use of the LangChain framework to enhance LLM performance.
- Learn how to develop AI-powered applications, particularly chatbots, using Microsoft Azure and other tools.
- Gain proficiency in working with vector databases and Retrieval-Augmented Generation (RAG) for optimized data retrieval.
- Apply learned concepts through practical implementation and case studies, including building and deploying AI models and applications.

Prerequisites

No Azure or generative AI experience is required, but experience with both is recommended.



Training 4 Programmers

To discuss this course and customizations:

Call: +1 434-509-5680 or Email: sales@training4programmers.com

Training Materials

All students receive comprehensive courseware covering all topics in the course. Courseware is distributed via GitHub in the form of documentation and extensive code samples. Students practice the topics covered through challenging hands-on lab exercises.

Software Requirements

Students will need a free, personal GitHub account to access the courseware. Student will need permission to install Visual Studio Code and Visual Studio Code Extensions on their computers.

Outline

- Overview of Generative AI and LLMs
 - An Overview of Generative AI, Azure OpenAI, and Large Language Models (LLMs)
 - Generative AI & LLMs
 - Generative AI Models and Azure OpenAI Base Models
 - GenAI Text based engines: Google's Gemini, ChatGPT, Bing GPT4, Poe and Lambda
 - Ensuring Responsible AI with Azure OpenAI's Policies
 - LLM use cases and tasks
 - LLM safety concerns- bias, fairness, transparency, usage restriction, content filtering
 - Understanding jailbreaks and LLM exploits
- LLM Architectures and Components
 - Transformers architecture
 - Generating text with transformers
 - Parameter efficient fine-tuning (PEFT)
 - Pre-training large language models
 - Fine-tuning with prompt engineering
 - Prompts and prompt engineering, Low-Rank Adaption, Prefix tuning
 - Adjusting parameters and understanding model performance
- Prompt Engineering and LangChain Framework
 - The Art of Crafting Prompts: Principles and Techniques
 - Introduction to LangChain framework, modules and components
 - Prompts with LangChain
 - Langchain Experiments
 - Prompt templates and dynamic prompts
 - Helping LLMs reason and plan with chain-of-thought
- LLM Application Development



Training 4 Programmers

To discuss this course and customizations:

Call: +1 434-509-5680 or Email: sales@training4programmers.com

- Building a chatbot on Microsoft Azure
- Questions and Answers chatbot on Azure
- Using the LLM in applications (OpenAI-Azure Hosted)
- Interacting with external applications
- LLM application architectures
- Vector Databases and RAGs (Retrieval-Augmented Generation)
 - Introduction to Vector DB and its variants
 - Indexing and query optimization strategies for Vector Databases
 - Building and searching Vector Indexes with Azure AI Search and blobs
 - Introduction to RAGs, concepts and architecture
 - RAG Variants and implementation
 - RAG components
 - RAG Assessment and Optimization
- Practical Implementation and Case Studies
 - Putting it all together: Building a SQL Assistant chatbot with information retrieval and bots for custom data
 - Generative AI Project Lifecycle Cheat Sheet
 - Preparing models for deployment with Flask API
 - Use Case: Going open source- comparing open-source LLMs vs Azure and OpenAI
 - Processing and analyzing unstructured data with LLMs