Application Development with LLMs on Google Cloud

7h

What you'll learn

- Explore the different options available for using generative AI on Google Cloud.
- Use Vertex AI Studio to test prompts for large language models.
- Develop LLM-powered applications using LangChain and LLM models on Vertex AI.
- Apply prompt engineering techniques to improve the output from LLMs.
- Build a multi-turn chat application using the PaLM API and LangChain.

Target audience

Customers interested in integrating large language models (LLMs) into their application.

Prerequisites

Completion of « Introduction to Developer Efficiency on Google Cloud » or equivalent knowledge.

Course Outline

Module 01: Introduction to Generative AI on Google Cloud

Topics

- Vertex AI on Google Cloud
- Generative AI options on Google Cloud
- Introduction to course use case

Objectives

- Explore the different options available for using generative AI on Google Cloud.

Module 02: Vertex AI Studio

Topics

- Introduction to Vertex AI Studio
- Available models and use cases
- Designing and testing prompts in the Google Cloud console
- Data governance in Vertex AI Studio

Objectives

- Use Vertex AI Studio to test prompts for large language models.
- Understand how Vertex AI Studio keeps your data secure

Activities

- Lab: Exploring Vertex AI Studio

Module 03: LangChain Fundamentals

- Introduction to LangChain
- LangChain concepts and components
- Integrating the Vertex AI PaLM APIs
- Question/Answering Chain using PaLM API

Objectives

- Understand basic concepts and components of LangChain
- Develop LLM-powered applications using LangChain and LLM models on Vertex AI

Activities

- Lab: Getting Started with LangChain + Vertex AI PaLM API

Module 04: Prompt Engineering

Topics

- Review of few-shot prompting
- Chain-of-thought prompting
- Retrieval augmented generation (RAG)
- ReAct

Objectives

- Apply prompt engineering techniques to improve the output from LLMs.
- Implement a RAG architecture to ground LLM models.

Activities

- Lab: Prompt Engineering Techniques

Module 05: Creating Custom Chat Applications with Vertex AI PaLM API

Topics

- LangChain for chatbots
- Memory for multi-turn chat
- Chat retrieval

Objectives

- Understand the concept of memory for multi-turn chat applications.
- Build a multi-turn chat application by using the PaLM API and LangChain

Activities

- Lab: Implementing RAG Using LangChain