

Complete Generative AI : From Basics to Expert Level

Are you ready to transform your AI skills and unlock new career opportunities? Our Complete Generative AI Course is designed to take you from beginner to expert in just 10 days. With a focus on practical, hands-on experience, this course will teach you how to leverage Large Language Models (LLMs), master prompt engineering, and build sophisticated AI applications. Whether you're a developer looking to enhance your skillset or an AI enthusiast eager to dive into the world of generative AI, this course has something for you.

Objectives and Outcomes:

By the end of this course, you will be able to:

- Grasp LLM fundamentals and application areas.
- Master prompt engineering for effective LLM interaction.
- Automate and innovate with ChatGPT API for complex systems.
- Leverage vector databases for deep data insights with LLMs.
- Develop business applications using Semantic Kernel and LLMs.
- Utilize LangChain for expansive LLM application development.
- Implement RAG for data-driven, contextual responses.
- Fine-tune LLMs for tailored application performance.
- Rapidly prototype AI apps with Gradio.
- Enhance JavaScript projects with LangChain.js for LLM integration.

Prerequisites:

Basic Python programming

Fundamentals of Machine Learning (optional)

Special Features:

Virtual Machine for Labs: DC Virtual Machine provided for hands-on labs.

Unofficial Courseware: Access to comprehensive unofficial course materials.

OpenAI Access Required: Students need their own OpenAI access. [Get OpenAI Access](#)

Tech Stack To Be Covered



ChatGPT



OpenAI



LangChain.js



Semantic Kernel



Gradio



Hugging Face



Large Language Model

Chapter 01: Prompt Engineering for Developers

In Prompt Engineering for Developers, you will learn how to use a large language model (LLM) to quickly build new and powerful applications. Using the OpenAI API, you'll be able to quickly build capabilities that learn to innovate and create value in ways that were cost-prohibitive, highly technical, or simply impossible before now. This chapter will describe how LLMs work, provide best practices for prompt engineering, and show how LLM APIs can be used in applications for a variety of tasks, including:

- Summarizing (e.g., summarizing user reviews for brevity)
- Inferring (e.g., sentiment classification, topic extraction)
- Transforming text (e.g., translation, spelling & grammar correction)
- Expanding (e.g., automatically writing emails)

In addition, you'll learn two key principles for writing effective prompts, how to systematically engineer good prompts, and also learn to build a custom chatbot. All

concepts are illustrated with numerous examples, which you can play with directly in our Jupyter notebook environment to get hands-on experience with prompt engineering

Chapter 02: Building Systems with the ChatGPT API

In Building Systems with the ChatGPT API, you will learn how to automate complex workflows using chain calls to a large language model. Unlock new development capabilities and improve your efficiency. You'll build:

- Chains of prompts that interact with the completions of prior prompts.
- Systems where Python code interacts with both completions and new prompts.
- A customer service chatbot using all the techniques from this course.

You'll learn how to apply these skills to practical scenarios, including classifying user queries to a chat agent's response, evaluating user queries for safety, and processing tasks for chain-of-thought, multi-step reasoning.

Chapter 03: Vector Databases - Embeddings to Applications

Vector databases play a pivotal role across various fields, such as natural language processing, image recognition, recommender systems and semantic search, and have gained more importance with the growing adoption of LLMs.

These databases are exceptionally valuable as they provide LLMs with access to real-time proprietary data, enabling the development of Retrieval Augmented Generation (RAG) applications.

At their core, vector databases rely on the use of embeddings to capture the meaning of data and gauge the similarity between different pairs of vectors and sift through extensive datasets, identifying the most similar vectors.

This chapter will help you gain the knowledge to make informed decisions about when to apply vector databases to your applications. You'll explore:

- How to use vector databases and LLMs to gain deeper insights into your data.
- Build labs that show how to form embeddings and use several search techniques to find similar embeddings.
- Explore algorithms for fast searches through vast datasets and build applications ranging from RAG to multilingual search.

Chapter 04: Building AI Plugins With Semantic Kernel

Large Language Models (LLMs) are enabling coders and non-coders to build new kinds of applications that harness the power of AI. In this course, you'll learn how to use and create with Microsoft's open source orchestrator, Semantic Kernel. Along the way you'll gain skills in getting the most out of LLMs developing prompts, semantic functions,

vector databases and using an LLM for planning.

After completing this chapter you will be able to:

- Develop sophisticated business applications using LLM's
- Leverage common LLM building blocks such as memories, connectors, chains and planners
- Utilize the open source orchestrator, the Semantic Kernel, in your applications

Using an orchestration SDK such as Semantic Kernel means you can avoid having to learn APIs for each individual AI service, and build integrations that can always stay up to date with the latest advances in AI research rather than solutions that quickly get outdated. Through this course, you'll have all you need to take full advantage of this powerful open source tool.

Chapter 05: LangChain for LLM Application Development

In LangChain for LLM Application Development, you will gain essential skills in expanding the use cases and capabilities of language models in application development using the LangChain framework.

In this course you will learn and get experience with the following topics:

- Models, Prompts and Parsers: calling LLMs, providing prompts and parsing the response
- Memories for LLMs: memories to store conversations and manage limited context space
- Chains: creating sequences of operations
- Question Answering over Documents: apply LLMs to your proprietary data and use case requirements
- Agents: explore the powerful emerging development of LLM as reasoning agents.

At the end of the chapter, you will have a model that can serve as a starting point for your own exploration of diffusion models for your applications. This chapter will vastly expand the possibilities for leveraging powerful language models, where you can now create incredibly robust applications in a matter of hours.

Chapter 06: LangChain - Chat with Your Data

The chapter delves into two main topics: (1) Retrieval Augmented Generation (RAG), a common LLM application that retrieves contextual documents from an external dataset, and a guide to building a chatbot that responds to queries based on the content of your documents, rather than the information it has learned in training.

You'll learn about:

- Document Loading: Learn the fundamentals of data loading and discover over 80 unique loaders LangChain provides to access diverse data sources, including audio and

video.

- Document Splitting: Discover the best practices and considerations for splitting data.
- Vector stores and embeddings: Dive into the concept of embeddings and explore vector store integrations within LangChain.
- Retrieval: Grasp advanced techniques for accessing and indexing data in the vector store, enabling you to retrieve the most relevant information beyond semantic queries.
- Question Answering: Build a one-pass question-answering solution.
- Chat: Learn how to track and select pertinent information from conversations and data sources, as you build your own chatbot using LangChain.

Start building practical applications that allow you to interact with data using LangChain and LLMs.

Chapter 07: Finetuning Large Language Models

When you complete this chapter, you will be able to:

- Understand when to apply finetuning on LLMs
- Prepare your data for finetuning
- Train and evaluate an LLM on your data

With finetuning, you're able to take your own data to train the model on it, and update the weights of the neural nets in the LLM, changing the model compared to other methods like prompt engineering and Retrieval Augmented Generation. Finetuning allows the model to learn style, form, and can update the model with new knowledge to improve results.

Chapter 08: Building Generative AI Applications with Gradio

By the end of the chapter, you'll gain the practical knowledge to rapidly build interactive apps and demos to validate your project and ship faster. What you'll do:

- With a few lines of code, create a user-friendly app (usable for non-coders) to take input text, summarize it with an open-source large language model, and display the summary.
- Create an app that allows the user to upload an image, which uses an image to text (image captioning) to describe the uploaded image, and display both the image and the caption in the app.
- Create an app that takes text and generates an image with a diffusion model, then displays the generated image within the app.
- Combine what you learned in the previous two lessons: Upload an image, caption the image, and use the caption to generate a new image.

- Create an interface to chat with an open source LLM using Falcon, the best-ranking open source LLM on the Open LLM Leaderboard.

Chapter 09: Build LLM Apps with LangChain.js

JavaScript is the world's most popular programming language, and now developers can program in JavaScript to build powerful LLM apps.

This chapter will show webdevs how to expand their toolkits with LangChain.js, a popular JavaScript framework for building with LLMs, and will cover useful concepts for creating powerful, context-aware applications. In this chapter, you will:

- Learn to use LangChain's underlying abstractions to build your own JavaScript apps
- Understand the basics of retrieval augmented generation (RAG)
- Have the structure of a basic conversational retrieval system that you can use for building your own chatbots

Chapter 10: Open-source Models with Hugging Face

In this chapter, you'll select open-source models from Hugging Face Hub to perform NLP, audio, image and multimodal tasks using the Hugging Face transformers library. Easily package your code into a user-friendly app that you can run on the cloud using Gradio and Hugging Face Spaces. You will learn the following objectives:

- Use the transformers library to turn a small language model into a chatbot capable of multi-turn conversations to answer follow-up questions.
- Translate between languages, summarize documents, and measure the similarity between two pieces of text, which can be used for search and retrieval.
- Convert audio to text with Automatic Speech Recognition (ASR), and convert text to audio using Text to Speech (TTS).
- Perform zero-shot audio classification, to classify audio without fine-tuning the model.
- Generate an audio narration describing an image by combining object detection and text-to-speech models.
- Identify objects or regions in an image by prompting a zero-shot image segmentation model with points to identify the object that you want to select.
- Implement visual question answering, image search, image captioning and other multimodal tasks.
- Share your AI app using Gradio and Hugging Face Spaces to run your applications in a user-friendly interface on the cloud or as an API.