5-Day Gen AI Intensive Course with Google Learn Guide

Welcome to our 5-Day Gen AI Intensive Course with Google! This was a live event from March 31– April 4, 2025, now made available as a self-paced learning guide for anyone interested in learning more about the fundamental technologies and techniques behind Generative AI.

What's covered:

- **Day 1**: Foundational Models & Prompt Engineering - Explore the evolution of Large Language Models (LLMs), from transformers to techniques like fine-tuning, inference acceleration, and reasoning models. Get trained with the art of prompt engineering for optimal LLM interaction.

- **Day 2**: Embeddings and Vector Stores/Databases - Learn about the conceptual underpinning of embeddings and vector databases, including embedding methods, vector search algorithms, and real-world applications with LLMs, as well as their tradeoffs.

- **Day 3**: Generative AI Agents - Learn to build sophisticated AI agents by understanding their core components and the iterative development process, including recent agentspace development.

- **Day 4**: Domain-Specific LLMs - Delve into the creation and application of specialized LLMs like SecLM and Med-PaLM, with insights from the researchers who built them.

- **Day 5**: MLOps for Generative AI - Discover how to adapt MLOps practices for Generative AI and leverage Vertex AI's tools for foundation models and generative AI applications.

Best of Luck!
- brought to you by Anant Nawalgaria, Mark McDonald, Paige Bailey, and many other contributors from Google.

Set Up

Follow the following steps to get set up before diving into the daily assignments:

- 1. Sign up for a Kaggle account and learn how Notebooks work. Make sure to phone verify your account, it's necessary for the course's code labs.
- 2. We also have a Troubleshooting Guide for the codelabs. Be sure to check there for solutions to common problems.
- 3. Sign up for an AI Studio account and ensure you can generate an API key.
- 4. Sign up for a Discord account and join us on the Kaggle Discord server to connect and collaborate with fellow learners in this course.

Please note that if you would like to post on other channels on the Kaggle discord you will need to link your Kaggle account to discord here: https://kaggle.com/discord/confirmation.

Day 1 (Foundational Large Language Models & Prompt Engineering)

Welcome to Day 1.

Today you'll explore the evolution of LLMs, from transformers to techniques like fine-tuning and inference acceleration. You'll also get trained in the art of prompt engineering for optimal LLM interaction and evaluating LLMs.

The first codelab will walk you through getting started with the Gemini 2.0 API and cover several prompt techniques including how different parameters impact the prompts. In the second codelab, you will also learn how to evaluate the response of LLMs using autoraters and structured output.

Day 1 Assignments:

1. Complete the Intro Unit – "Foundational Large Language Models & Text Generation":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "Foundational Large Language Models & Text Generation" whitepaper.

2. Complete Unit 1 – "Prompt Engineering":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "Prompt Engineering" whitepaper.
- Complete these codelabs on Kaggle:
  - Prompting fundamentals
  - Evaluation and structured data
  - Make sure you phone verify your account before starting, it's necessary for the codelabs
  - Want to have an interactive conversation? Try adding the whitepapers to NotebookLM

[Optional] Read a case study to learn how a leading bank leveraged advanced prompt engineering and other contents discussed in assignments of day 1 to automate their financial advisory workflows, achieving significant productivity gains.

[Optional] Watch the YouTube livestream recording. Paige Bailey will be joined by expert speakers from Google – Warren Barkley, Logan Kilpatrick, Kieran Milan, Anant Nawalgaria, Irina Sigler and Mat Velloso to discuss today's readings and code labs.

Day 2 (Embeddings and Vector Stores/Databases)

Welcome to Day 2.

Today you will learn about the conceptual underpinning of embeddings and vector databases and how they can be used to bring live or specialist data into your LLM application. You'll also explore their geometrical powers for classifying and comparing textual data as well as how to evaluate embeddings.

Day 2 Assignments:

Complete Unit 2: "Embeddings and Vector Stores/Databases":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "Embeddings and Vector Stores/Databases" whitepaper..
- Complete these code labs on Kaggle:
  - Build a RAG question-answering system over custom documents
  - Explore text similarity with embeddings
  - Build a neural classification network with Keras using embeddings.
  - Want to have an interactive conversation? Try adding the whitepapers to NotebookLM

[Optional] Watch the YouTube livestream recording. Paige Bailey will be joined by expert speakers from Google – Andre Araujo, Patricia Florissi, Alan Li, Anant Nawalgaria, Xiaoqi Ren, Chuck Sugnet, and Howard Zhou to discuss embeddings and vector stores/databases.

Day 3 (Generative Agents)

Welcome to Day 3.

Today you'll learn to build sophisticated AI agents by understanding their core components and the iterative development process. You'll also learn more about advanced agentic architectures and approaches such as multi-agent systems, agent evaluation and more.

The codelabs cover how to connect LLMs to existing systems and to the real world. Learn about function calling by giving SQL tools to a chatbot (including an example using Gemini 2.0's Live API), and learn how to build a LangGraph agent that takes orders in a café.

Day 3 Assignments:

Complete Unit 3a - "Generative AI Agents":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "Generative AI Agents" whitepaper.
- Complete these code labs on Kaggle:
  - Talk to a database with function calling.
  - Build an agentic ordering system in LangGraph.

[Optional] Advanced 3b - "Agents Companion":

- Listen to the summary podcast episode for this unit.
- Read the advanced "Agents Companion" whitepaper.
- Want to have an interactive conversation? Try adding the whitepapers to NotebookLM

[Optional] Read a case study which talks about how a leading regulatory reporting solutions provider used an agentic generative AI system to automate ticket-to-code creation in software development, achieving a 2.5x productivity boost.

[Optional] Watch the YouTube livestream recording. Paige Bailey will be joined by expert speakers from Google – Alan Blount, Antonio Gulli, Steven Johnson, Jaclyn Konzelmann, Patrick Marlow, Anant Nawalgaria and Julia Wiesinger to discuss generative AI agents.

Day 4 (Domain-Specific LLMs)

Welcome to Day 4.

In today's reading, you'll delve into the creation and application of specialized LLMs like SecLM and MedLM/Med-PaLM, with insights from the researchers who built them.

In the codelabs, you'll learn how to add real-world data to a model using Google Search and then visualize it with plotting tools via the Live API.You will also learn how to fine-tune a custom Gemini model using your own labeled data to solve custom tasks.

Day 4 Assignments:

Complete Unit 4 – "Domain-Specific LLMs":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "Solving Domain-Specific Problems Using LLMs" whitepaper.
- Complete these codelabs on Kaggle:
  - Tune a Gemini model for a custom task
  - Use Google Search data in generation.
  - Want to have an interactive conversation? Try adding the whitepapers to NotebookLM.

[Optional] Watch the YouTube livestream recording. Paige Bailey will be joined by expert speakers from Google -Donny Cheung, Scott Coull, Ewa Dominowska, Chris Grier, Anant Nawalgaria and Karthik Raman to discuss domain specific models.

Day 5 (MLOps for Generative AI)

Welcome to Day 5.

Discover how to adapt MLOps practices for Generative AI and leverage Vertex AI's tools for foundation models and generative AI applications such as AgentOps for agentic applications.

Day 5 Assignments:

Complete Unit 5 - "MLOps for Generative AI":

- Listen to the summary podcast episode for this unit.
- To complement the podcast, read the "MLOps for Generative AI" whitepaper..
- No codelab for today! During the livestream tomorrow, we will do a code walkthrough and live demo of goo.gle/agent-starter-pack, a resource created for making MLOps for Gen AI easier and accelerating the path to production. Please go through the repository in advance.
- Want to have an interactive conversation? Try adding the whitepapers to NotebookLM

[Optional] Watch the YouTube livestream recording. Paige Bailey will be joined by expert speakers from Google – Sokratis Kartakis, Gabriela Hernandez Larios, Ivan Nardini, Anant Nawalgaria, Elia Secchi, Michael Styer and Saurabh Tiwary to discuss MLOps practices for generative AI.

Bonus Assignment

Check out this bonus assignment: There's more! This bonus notebook walks you through a few more things you can do with the Gemini API that weren't covered during the course. This material doesn't pair with the whitepapers or podcast, but covers some extra capabilities you might find useful when building Gemini API powered apps.