**5-Days Gen Ai**

**Day 1: Foundational Large Language Models & Text Generation and Prompt Engineering**

Complete the Intro Unit - “Foundational Large Language Models & Text Generation”, which is:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Founda... ) for this unit (created by NotebookLM, https://notebooklm.google.com/).

➡️ Read the “Foundational Large Language Models & Text Generation” whitepaper - https://www.kaggle.com/whitepaper-fou....

Complete Unit 1 - “Prompt Engineering”, which is:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Prompt... ) for this unit (created by NotebookLM).

➡️ Read the “Prompt Engineering” whitepaper - https://www.kaggle.com/whitepaper-pro....

➡️ Complete this code lab (https://www.kaggle.com/code/markisher...) on Kaggle where you’ll learn prompting fundamentals. Make sure you phone verify (https://www.kaggle.com/settings) your account before starting, it's necessary for the code labs.

**Day 2: Embeddings and Vector Stores/Databases**

Resources mentioned in today's livestream:

Jinhyuk Lee's Google Scholar profile: https://scholar.google.com/citations?...

The original transformer paper: Attention Is All You Need https://arxiv.org/abs/1706.03762

BERT paper explaining bidirectional attention: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding https://arxiv.org/abs/1810.04805

A recent paper from NVidia explaining how to adapt decoder-only language model for embedding generation: NV-Embed: Improved Techniques for Training LLMs as Generalist Embedding Models https://arxiv.org/abs/2405.17428

Whitepaper explaining "Native Integration of the ScaNN Algorithm into AlloyDB Database Internals": https://services.google.com/fh/files/...

Assignments:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Embedd... ) for this unit (created by NotebookLM, https://notebooklm.google.com/).

➡️ Read the “Embeddings and Vector Stores/Databases” whitepaper - https://kaggle.com/whitepaper-embeddi...

➡️ Complete these code labs on Kaggle:

Build a RAG question-answering system over custom documents - https://www.kaggle.com/code/markisher...

Explore text similarity with embeddings - https://www.kaggle.com/code/markisher...

Build a neural classification network with Keras using embeddings - <https://www.kaggle.com/code/markisher>...

**Day 3: Generative AI Agents**

Resources mentioned in today's livestream:

Agent building frameworks:

https://www.langchain.com/langgraph

https://firebase.google.com/docs/genkit

https://github.com/breadboard-ai/brea...

What is RAG? high level page from GCP / Vertex:

https://cloud.google.com/use-cases/re...

Complete Unit 3: “Generative AI Agents”, which is:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Agents... ) for this unit (created by NotebookLM).

➡️ Read the “Generative AI Agents” whitepaper - https://www.kaggle.com/whitepaper-agents

➡️ Complete these code labs on Kaggle:

Talk to a database with function calling - https://www.kaggle.com/code/markisher...

Build an agentic ordering system in LangGraph - https://www.kaggle.com/code/markisher...

**Day 4: Domain-Specific LLMs**

Complete Unit 4: “Domain-Specific LLMs”, which is:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Solvin... ) for this unit (created by NotebookLM).

➡️ Read the “Solving Domain-Specific Problems Using LLMs” whitepaper - https://www.kaggle.com/whitepaper-sol...

➡️ Complete these code labs on Kaggle:

[Optional] Use Google Search data in generation. (Note: Grounding with Google Search has been released as a limited launch and is not available in all locations. The EEA, UK, and CH regions will be supported at a later date) - https://www.kaggle.com/code/markisher...

Tune a Gemini model for a custom task - <https://www.kaggle.com/code/markisher>...

**Day 5: MLOps for Generative AI**

Neptune:https://github.com/google-deepmind/ne...

Gecko: Text-to-image evaluation: https://arxiv.org/abs/2404.16820

Google Cloud Vertex AI: vertex ai https://cloud.google.com/vertex-ai/

Bigquery https://cloud.google.com/bigquery/

Vertex AI GenAI evaluation Service: https://cloud.google.com/vertex-ai/ge...

Complete Unit 5: “MLOps for Generative AI”, which is:

➡️ [Optional] Listen to the summary podcast episode ( • Whitepaper Companion Podcast - Operat... ) for this unit (created by NotebookLM, https://notebooklm.google/).

➡️ Read the “MLOps for Generative AI” whitepaper - https://www.kaggle.com/whitepaper-ope...

➡️ No code lab for today! We will do a code walkthrough and live demo of goo.gle/e2e-gen-ai-app-starter-pack (https://goo.gle/e2e-gen-ai-app-starte..., a resource created for making MLOps for Gen AI easier and accelerating the path to production. Please go through the repository in advance.

SUBSCRIBE: https://www.youtube.com/c/kaggle?sub\_...

**About Kaggle:**

Kaggle is the world's largest community of data scientists. Join us to compete, collaborate, learn, and do your data science work. Kaggle's platform is the fastest way to get started on a new data science project. Spin up a Jupyter notebook with a single click. Build with our huge repository of free code and data. Stumped? Ask the friendly Kaggle community for help.

**Follow Kaggle online:**

Visit the WEBSITE: http://www.kaggle.com/?utm\_medium=you...

Like Kaggle on FACEBOOK: http://www.facebook.com/kaggle?utm\_me...

Follow Kaggle on TWITTER: http://twitter.com/kaggle?utm\_medium=...

Check out our BLOG: http://blog.kaggle.com/?utm\_medium=yo...

Connect with us on LINKEDIN: http://www.linkedin.com/company/kaggl...

**Advance your data science skills:**

Take our free online courses: http://www.kaggle.com/learn/overview?...

Get started with Kaggle Kernels: http://www.kaggle.com/docs/kernels?ut...

Download clean datasets from Kaggle: http://www.kaggle.com/docs/datasets?u...

Sign up for a Kaggle Competition: http://www.kaggle.com/docs/competitio...

Explore the Kaggle Public API: http://www.kaggle.com/docs/api?utm\_me...