Summer Training TR-103 Prompt Engineering

Day 6 Report

The sixth day of the training focused on exploring Google AI Studio, its advantages, features, and practical applications in prompt engineering and generative AI workflows. Participants gained hands-on exposure to model comparison, media generation, streaming, fine-tuning, multimodal capabilities, and token usage. The session also highlighted differences between free and paid AI models, integration with APIs and SDKs, and advanced concepts such as RAG, diffusion models, multimodal architectures, and Google's video generation model, Veo.

Why use Google AI Studio? (Advantages)

- Offers more model options, enabling users to experiment with different capabilities.
- Provides freely available tokens, making it accessible for learning.
- Includes Comparison Mode, allowing users to evaluate differences between Flash and Pro models:
    - Flash: Faster results
    - Pro: Higher-quality results
- Tokens increase as usage grows, allowing more operations over time.
- Vertex AI is the paid enterprise version that delivers more advanced and refined results; widely used by developers along with OpenAI’s premium offerings.
- Integration with Google Drive:
    - Enables file and dataset storage
    - Tip: Use separate Google accounts because Drive space is limited
- Grounding with Google Search:
    - Provides real-time, updated information
    - Disadvantage: Consumes more tokens

- Thinking Mode:
    - Turned off by default because enabling it increases token consumption
- Streaming Mode:
    - Allows screen-sharing to help the model understand user context (e.g., showing a setup)
    - Occasional issues may occur when linking to a server
- Processes videos, audio, images, and text, showcasing strong multimodal capabilities.

Limitation of OpenAI Free Models

OpenAI's free tier lacks full support for understanding videos, audios, and images, making Google AI Studio more flexible for multimodal tasks.

Tokens in Google AI Studio

Tokens can be accessed through:

- Paper Tokens (free)
- Purchased Tokens (paid)

These tokens support multimodal processing such as audio, video, image, and text inputs.

API Keys and SDKs

When opening Google AI Studio, users can access:

- API Keys — to integrate Google AI into external applications
- SDKs (Software Development Kits) — to build software using Google's generative AI models

Fine-Tuning Options in AI Studio

Google AI Studio provides fine-tuning features that enable customization of pre-trained models.

The session covered:

- RAG (Retrieval-Augmented Generation)
- Notebook LLM

- Diffusion Models (Image generation)
- Multimodal Models (Audio, video, image processing)
- Veo — Google's model for advanced video generation

Multiple API keys can be used for different tasks or projects.

Why Use Pre-Existing Models Instead of Creating Your Own?

Building a model from scratch requires:

- massive datasets
- expensive compute resources
- long training time
- extensive infrastructure

Pre-trained models offer optimized performance, reliability, and instant usability, making them more practical.

MCP – Multimodal Chain Prompting

Participants were introduced to Multimodal Chain Prompting (MCP), a technique that:

- Links multiple prompts in a chain
- Uses templates for consistency
- Enhances model reasoning across different modalities

Core Features Explored in Google AI Studio

- **Comparison Mode:**
  - Allows side-by-side evaluation of outputs from different models.
  - Helps assess quality, speed, tone, and accuracy.
  - Enables informed decision-making when selecting models for projects.
- **Build:**
  - Introduced the process of creating custom AI workflows.

- Participants practiced combining prompts and model settings to build simple AI pipelines.
- Explored how to create custom AI-powered applications using the Build interface.

- **Stream:**
  - Demonstrated real-time output generation for interactive use cases.
  - Participants observed how streaming supports low-latency, dynamic interactions.
  - Useful in scenarios requiring continuous or progressive output.
- **Generate Media:**
  - Used to create AI-generated images, audio, and other media types.
  - Participants generated sample media to explore the creative potential of multimodal models.
  - Demonstrated use cases such as content creation, design support, and multimedia generation.
- **Fine-Tuning:**
  - Explained how pre-trained models can be customized with user datasets.
  - Participants performed sample fine-tuning tasks to see improvements in accuracy and relevance.
  - This helped them understand the role of fine-tuning in specialized applications.

Conclusion

The sixth day of training provided participants with extensive practical experience in using Google AI Studio and its advanced tools. By exploring Comparison Mode, Build, Stream, Generate Media, Fine-Tuning, and multimodal capabilities, attendees gained a strong working understanding of the platform. The session deepened their knowledge of token usage, model selection, and customization techniques while preparing them to engage confidently with more complex AI concepts in the upcoming sessions.