Faysal Aziz

Mastering Modern LLMs

ChatGPT, DeepSeek, Gemini & Prompt Engineering in a day

FA-0313

Agentic & Generative AI

Introduction

Unlock the latest in conversational AI, smart prompting, and real-world model applications—taught by an instructor with 30+ years of industry-proven expertise and focused on practical, in-demand skills.

AI is eating the world—and if you're not riding the wave, you're getting left behind. This isn't your typical corporate training session filled with buzzwords and theory. It's a full-day deep dive into the tools that are actually reshaping entire industries right now.

We'll crack open the hood on transformer architecture, decode token economics, and get you hands-dirty with the heavy hitters: ChatGPT, DeepSeek, and Gemini. Whether you're working with free tiers or premium APIs, you'll walk away knowing how to craft prompts that make these models work for you—not against you.

Your guide? Someone who's been in the trenches for 30+ years and has seen tech revolutions come and go. This is battle-tested knowledge from the front lines, not textbook fluff. By the end of the day, you'll have the skills to harness AI like a pro and stay ahead of the curve.

Learning outcomes

By the end of the day, you will be able to:

- Describe transformer architecture, tokenization, and context windows.
- Explain the latest features of ChatGPT (GPT-5, Deep Research, agents).

- Compare DeepSeek (R1, V3, open-source positioning, API pricing).
- Understand Gemini's multimodal capabilities and extended context support.
- Engineer effective prompts using structured prompting techniques.
- Navigate token limitations and manage long-form or multi-step tasks.
- Select the best tool based on task type, resources, and accessibility.

Prerequisites

Participants should have:

- Basic familiarity with programming or AI concepts.
- Eagerness to explore and compare different LLM services.
- A laptop with internet access; optionally, accounts with OpenAI, DeepSeek, and Google.

Training outline

1. · Foundations of AI & Transformers

- How transformer models operate (self-attention, encoder/decoder-only designs).
- Tokenization basics, context window size, and their impact on task performance.

2. · ChatGPT – The State of the Art (GPT-5)

- Overview of GPT-5 release (August 2025): "PhD-level expert" capabilities, reasoning improvements, tool integrations like Gmail/Calendar, and personality modes.
- Deep Research: query allowances by tier (free up to lightweight queries; Pro, Team, Enterprise get more) and use cases for autonomous report generation.

  Agent features:

  - **Operator** for web-based tasks (form-filling, scheduling).
  - **Codex** for coding tasks and local/cloud interaction.
- Multimodal support via GPT (text, image, audio, video) and file/image generation.

3. · DeepSeek – The Efficient Challenger

- Overview of DeepSeek R1, V3: open-source availability, app and web access, API option with usage-based pricing (approx. \$0.55 per million input tokens; \$2.19 per million output tokens).
- R1-0528 update improvements: reasoning, coding accuracy, support for JSON output, function calling.

- Lightweight training efficiency: fewer GPUs required, lower training cost, and open license compared to U.S. counterparts.
- Caution: potential content censorship and geopolitical bias concerns.

4. · Gemini – Google's Multimodal Powerhouse

- Fully multimodal: supports text, image, audio, video in any sequence, with rich cross-modal understanding.
- Giant context window: models like Gemini 1.5 support up to one million tokens (~hours of content).
- Specialized variants: CodeGemma for programming; MedGemma for medical; ShieldGemma for content moderation; DolphinGemma for audio (e.g. dolphins).
- AI Studio: interactive prompt interface with freeform, structured, chat prompts; parameter tuning (temperature, top-K/P); multimodal input; prompt → code export and integration with Vertex AI for production deployment.
- "Guided Learning" mode for educational customization (like ChatGPT's Study Mode).

5. · Prompt Engineering & Token Management

- Use of structured prompts: examples, few-shot prompting, chain-of-thought techniques.
- Handling token limits:
- DeepSeek: efficiency and context varies, managed via API parameters.
- Tips for prompt optimization and effective lengthy tasks.
- Choosing the right tool.

© 2026 Faysal Aziz · Kuala Lumpur