Faysal Aziz

Generative AI & Prompt Engineering

Mastering the Future From transformers to agentic systems – real skills, modern models, instant deployment - 2 days

FA-0225 · 2 days

Agentic & Generative AI

Introduction

Generative AI has moved beyond the experimental phase and is now reshaping industries from business operations to software development, creative fields, and automation systems. This course starts with core AI fundamentals—covering machine learning principles, deep learning frameworks, neural network architectures, and transformer models—before advancing to current large language models, prompt optimization strategies, API implementation, and modern autonomous AI systems.

Instead of focusing on theoretical concepts, the course emphasizes hands-on, enterprise-ready applications taught by an instructor with extensive real-world experience using the most current models and deployment methodologies.

Learning outcomes

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

- Explain core AI paradigms: machine learning, deep learning, neural networks, transformers
- Compare the strengths and limitations of major LLMs (proprietary vs open-source) as of mid-2025
- Integrate with LLM APIs under real-world constraints (rate limits, cost, security)
- Design and refine high-quality prompts using modern techniques (zero-shot, few-shot, CoT, role-prompting)
- Build prompt engineering systems robust to hallucinations and bias

- Understand agentic AI, Model Context Protocol systems, and multi-agent orchestration
- Apply knowledge via real-world labs, scenarios, and deployment strategies

---

Prerequisites

- Technical background: programming or software experience preferred
- Basic understanding of probability/statistics and neural network concepts
- Familiarity with REST APIs or SDKs

Training outline

1. AI Foundations

- **Machine Learning basics**: supervised, unsupervised, reinforcement learning
- **Deep Learning & Neural Networks**: perceptrons, feed-forward nets, backpropagation
- **Transformer architecture**: attention mechanism, encoder-decoder vs decoder-only designs
- Pretraining vs fine-tuning vs instruction tuning

2. Modern LLM Landscape

- **OpenAI GPT-4 family**: GPT-4.1, GPT-4o multimodal, reasoning and long-context strengths
- **Anthropic Claude 4 (Sonnet & Opus)**: deep reasoning, long tasks, safety features, tool-enabled workflows
- **Google Gemini (2.0 / 2.5 Pro & Flash)**: multimodal, agentic capabilities, million-token context windows
- **Meta Llama 3.x**: open-source, multilingual, competitive benchmarks with cost efficiency
- **Mistral models (Mixtral 8×7B, Medium 3, Devstral)**: open-source mixture-of-experts, strong coding performance
- **Emerging models**: DeepSeek-LLM, Nemotron, Qwen, Gemma 3 open-models, Domain-specific SLMs

3. Comparative Dimensions

- Performance by task: reasoning, coding, creativity, multilingual
- Context window size, token limits, latency, cost per million tokens
- Licensing and deployment: proprietary vs open-source, on-prem vs cloud
- Ecosystem support: fine-tuning, tools, platforms (e.g. Azure, Vertex AI, Hugging Face)

4. Deployment & API Integration

- API concepts: endpoints, rate limits, batching, streaming
- Security, privacy, and compliance considerations in enterprise deployments
- Monitoring, prompt versioning, feedback loops within MLOps pipelines
- Hybrid architectures: combining LLMs with small language models or knowledge graphs (GraphRAG)

5. Prompt Engineering Fundamentals

- Definition and business value in 2025 context (market size, error-rate risks)

  Key techniques:

  - Zero-shot, few-shot, chain-of-thought prompting, role prompting
  - Adaptive prompting and vibe-coding trends (Andrew Ng's lazy prompting)
- Reducing hallucinations and bias: permission to say "I don't know," ask for citations, structured fallback logic
- Frameworks and best practices for robust prompt pipelines

6. Prompt Engineering in Practice

- Prompt design labs: refining prompts with debugging, evaluation metrics
- Performance testing: quality assessments, human-in-the-loop validation
- Building reusable prompt libraries and templates
- Tools and orchestration: LangChain, prompt chaining, memory, tool use integration

7. Agentic AI and Multi-Agent Systems

- Definition of agentic workflows, LAMs (large action models), multi-agent cooperatives
- Model Context Protocol (MCP): managing context across agents and tool integrations
- Designing autonomous agents: use cases in customer service, research assistants, code-based agents
- Safety, governance, and orchestration in agentic systems

8. Ethics, Governance & Risk Mitigation

- **MCP Vulnerabilities**: Secure design, hardened SDKs, least privilege model
- **Operational Guardrails**: Human-in-the-loop checkpoints, audit logs, traceability
- **Responsible Agent Use**: Privacy compliance, bias monitoring, transparency

9. Real-World Scenarios & Deployment Use Cases

- Designing pipelines for enterprise content generation with guardrails and determinism
- Prompt-driven automation in software engineering, customer support, and R & D
- End-to-end deployment walkthroughs: model selection, integration, monitoring, prompt tuning

---

© 2026 Faysal Aziz · Kuala Lumpur