IV B.Tech I Semester

23A31701	GENERATIVE AI & PROMPT ENGINEERING (Professional Core)	L	T	P	C
		3	0	0	3

Course Objectives

- Understand the foundations and working of Generative AI models.
- Explore various generative models like GANs, VAEs, and LLMs.
- Learn prompt engineering techniques to effectively interact with language models.
- Design applications using LLMs with precise control through prompting.
- Understand ethical and societal implications of using Generative AI.

Course Outcomes (COs)
1. Explain the fundamentals of Generative AI, compare model architectures (GANs, VAEs, Transformers), and evaluate their use in generating text, images, and other media.
2. Apply prompt engineering techniques including few-shot learning, output formatting, and debugging to control and guide generative model outputs.
3. Analyze the architecture and capabilities of large language models (LLMs), and build NLP applications using prompt engineering and fine-tuning techniques.
4. Design complex multi-step prompting workflows using tools like LangChain and LlamaIndex, and generate structured or multimodal outputs safely and effectively.
5. Assess the ethical, legal, and societal implications of generative AI, and evaluate its responsible use across fields like healthcare, education, and law.

Unit I: Introduction to Generative AI
Overview of Generative AI and Applications, Generative vs Discriminative Models, Latent Space and Data Generation Concepts, Architectures: GANs, VAEs, Autoregressive Models, Generative AI in Text, Image, Audio, and Video, LLMs: Pretrained Transformers as, Generators, Training Challenges and Evaluation of Generative Models, Case Studies: Image Synthesis, Text Generation.

Unit II: Prompt Engineering Fundamentals
Introduction to Prompt Engineering, Prompt Formats: Zero-shot, One-shot, Few-shot, Prompt Tuning vs Prompt Programming, In-Context Learning & Chain-of-Thought Prompting, Role of Instructions and Examples in Prompts, Controlling Output Style, Tone, and Format, Prompt Failure Cases and Debugging, Prompt Engineering for Coding, Text Completion.

Unit III: Generative Models in NLP
Transformer Architecture Recap (BERT, GPT), GPT-3/4, PaLM, Claude, and LLaMA Architectures, Text Generation Pipelines and APIs (OpenAI, HuggingFace), Prompt Engineering with GPT Models, Fine-tuning vs Instruct Tuning, Retrieval-Augmented Generation (RAG), Evaluation Metrics: BLEU, ROUGE, Perplexity, Building LLM-based Apps with LangChain.

Unit IV: Advanced Prompt Engineering & Tools
Role of Temperature, Top-k, Top-p Sampling, Structured Outputs: Tables, JSON, Function Calls, Agentic Prompting and Multi-step Reasoning, Prompt Chaining and Memory Handling, Prompt Templates for Automation (LangChain, LlamaIndex), Prompt Engineering for Multimodal Models (DALL·E, Gemini, Sora), Safety Layers & Guardrails in Prompting,AutoGPT, BabyAGI, and Agentic Workflow Building.

Unit V: Ethics, Risks, and Applications of Generative AI
Risks: Hallucination, Toxicity, Bias, Deepfakes and Misinformation Challenges, Copyright, IP, and Data Privacy in Generated Content, Evaluation of Responsible AI Outputs, Red Teaming and Safety Testing, Applications in Education, Medicine, Art, and Law, Regulatory Landscape for Generative AI, Future Trends and Research Directions

Textbooks

1. "Deep Learning with Python", François Chollet, Manning, 2nd Edition
2. "Generative Deep Learning", David Foster, O'Reilly, 2nd Edition
3. "Building Systems with ChatGPT", Emmanuel Ameisen (O'Reilly Short Reads)
4. "The Art of Prompt Engineering", Nathan Hunter (Free online eBook)

Reference Books & Papers

1. Vaswani et al., Attention is All You Need
2. OpenAI Technical Reports on GPT Models
3. Papers from NeurIPS, ACL, ICML related to XAI and LLMs
4. LangChain Documentation