Training Day 2 Report:

July 1, 2025

Generative Artificial Intelligence (Gen AI) refers to a class of AI systems capable of creating new content such as text, images, audio, video, or code, by learning patterns from existing data. Unlike traditional AI systems that are designed to perform specific tasks (like classification or prediction), Gen AI focuses on generating data that mimics human creativity and intelligence.

Several prominent generative AI models have emerged, including:

- **GPT (Generative Pre-trained Transformer)** by OpenAI – Primarily used for **text generation**, **coding assistance**, summarization, and natural language understanding.

- **Gemini** by Google DeepMind – A **multimodal AI model** capable of handling both **text and image understanding**, useful for complex reasoning, visual tasks, and **code generation**.

- **Claude** by Anthropic – A language model focused on **safe and aligned text-based interactions**, ideal for writing, summarizing, and conversational tasks.

- **LLaMA (Large Language Model Meta AI)** by Meta – An open-source model mainly used in **text processing** and research, with applications in **coding** and experimentation in natural language processing.

- **Grok** by xAI – A **text-based conversational assistant** integrated into the X platform (formerly Twitter), focused on **real-time dialogue** and social media interaction.

- **DALL·E** by OpenAI – Specializes in **image generation** from textual prompts, allowing users to create visuals based on descriptions.

Large Language Models (LLMs): Training and Limitations

Large Language Models (LLMs) are a type of generative AI trained to understand, generate, and interact using human language. These models are based on deep learning

architectures, particularly transformers, and are trained on massive datasets containing text from books, websites, and other digital content. Examples of LLMs include **GPT**, **Claude**, **LLaMA**, **Gemini**, and **Grok**.

Training an LLM involves two main phases:

- **Pretraining:** The model learns general language patterns by predicting the next word in a sentence across billions of tokens. This phase uses unsupervised learning on a large corpus of data.

- **Fine-tuning:** The pretrained model is refined using supervised data or reinforcement learning to specialize it for specific tasks like answering questions, summarizing, or holding a conversation.

Despite their capabilities, LLMs have notable limitations:

- **Hallucination:** They may generate false or misleading information with high confidence.

- **Bias:** Since they learn from internet data, they can inherit societal and cultural biases.

- **Context Limitation:** They have a maximum context window, beyond which they forget earlier parts of the conversation.

- **Lack of Understanding:** LLMs do not truly understand the meaning behind the text—they generate based on patterns.

- **Resource Intensive:** Training and deploying LLMs require significant computational power and energy.

While LLMs have revolutionized natural language processing, ongoing research aims to make them more accurate, ethical, and energy-efficient.