Training Day 2 Report

Date: 2 July 2025
By: Sahil Stathia Sharma
URN: 2302658
CRN: 2315255

Overview

On the second day of the training session, we focused on Large Language Models (LLMs), a core area of artificial intelligence. These models are capable of understanding and generating human-like text based on vast amounts of training data. The session gave us insight into how LLMs work, their components, the training process, and their impact across different sectors like education, healthcare, customer support, and creative industries. Real-world examples such as ChatGPT, Claude, and Gemini were discussed to help us connect theory to practice. The session also introduced ethical concerns and limitations associated with AI technologies.

We Explored

Definition and real-world analogy of LLMs

Key terms like Token, Parameter, Prompt, Fine-tuning, Inference

Working process of LLMs through tokenization and transformer models

Evolution of LLMs from GPT-1 to GPT-4o and Gemini

Three main stages of training: Pretraining, Fine-tuning, and RLHF

Prompt Engineering types: Zero-shot, Few-shot, Chain-of-thought

Real-world applications in education, healthcare, customer support, law, and content creation

Limitations like hallucinations, biases, and lack of actual understanding

Ethical concerns and the future of LLMs in multimodal AI, autonomous agents, and mobile deployment

1. Definition and Real-World Analogy of LLMs

Large Language Models (LLMs) are advanced artificial intelligence models trained on massive datasets of text to understand and generate human-like language. They can answer questions, write essays, translate languages, generate code, and more.

Analogy: Think of an LLM as a super-smart librarian who has read almost every book in the world. When you ask a question (prompt), the librarian uses knowledge from all those books to give you the best possible answer even though it doesn't understand the world like a human**.**

2. Key Terms

 - Token: The smallest unit of text (e.g., a word, part of a word, or symbol) that the model processes
.- Parameter: The adjustable weights (like neurons) in the model. GPT-3 has 175 billion of these!
 - Prompt: The input or instruction given to the model (e.g., 'Explain photosynthesis').
 - Fine-tuning: The process of training a pretrained model on a smaller, specific dataset.
 - Inference: When a trained model is used to generate predictions or responses to new prompts.

3. How LLMs Work: Tokenization & Transformer Architecture

Tokenization breaks input text into tokens (e.g., 'ChatGPT is cool' ['Chat', 'G', 'PT', 'is', 'cool']).
Transformer models use attention mechanisms to understand relationships between tokens and generate meaningful responses.

4. Evolution of LLMs

GPT-1 (2018): 117M parameters, basic understanding
GPT-2 (2019): 1.5B parameters, better generation
GPT-3 (2020): 175B parameters, strong few-shot capabilities
GPT-3.5 (2022): Improved reasoning and speed
GPT-4 (2023): Multimodal input (text + image)
GPT-4o (2024): Real-time vision, audio, and language
Gemini (202325): Advanced reasoning and multimodal capabilities

5. Three Main Stages of Training

 1. Pretraining: Model learns patterns from massive text data.
 2. Fine-tuning: Model is refined on domain-specific data.
 3. RLHF (Reinforcement Learning from Human Feedback): Aligns model with human preferences.

6. Prompt Engineering Types

- Zero-shot: No example; general training is used.

Prompt: 'Translate to French: Good morning.'

- Few-shot: Few examples are given.

Prompt: 'Hello Bonjour, Thank you Merci, Good morning ?'

- Chain-of-thought: Step-by-step reasoning.

Prompt: 'John > Mary > Sam. Who's the oldest? Let's think step by step...'

7. Real-World Applications

Education: Personalized tutoring, content generation.
Healthcare: Record summarization, chatbots.
Customer Support: 24/7 AI assistants.
Law: Document summarization, legal search.
Content Creation: Blogs, scripts, posts.

8. Limitations of LLMs

- Hallucinations: Generates false or made-up information.
- Biases: May reflect harmful stereotypes from training data.
- No true understanding: Predicts patterns but doesn't understand context like humans.

9. Ethical Concerns and Future of LLMs

Ethics:
- Misinformation, plagiarism, privacy concerns.
- Copyright and plagiarism issues.
- Privacy concerns with sensitive data input.

Future:
- Multimodal AI: Text, image, video, audio.
- Autonomous agents: AI acting independently.
- Mobile deployment: Efficient models on smartphones.

Conclusion:

LLMs are transforming human-machine interaction. With responsible development, they offer powerful tools for education, business, healthcare, and creativity shaping the future of AI in everyday life.