Summer Training TR-103 Prompt Engineering

Day 3 Report

The third day of the training focused on understanding Large Language Models (LLMs)—their working principles, core components, evolution, limitations, and real-world applications. Participants learned how LLMs generate human-like text using vast datasets and transformer architectures. The session introduced essential terminology such as tokens, parameters, prompts, inference, and fine-tuning, supported by visual slides and interactive quiz activities.

Large Language Model (LLM)

- A Large Language Model (LLM) is a type of AI trained to understand and generate human-like text.
- It uses vast datasets (books, websites, conversations) to learn patterns in language.
- It is like a super-charged autocomplete system that understands context deeply.

eg. ChatGPT, Google Gemini, Claude, Llama

Analogy: Like a human who has read billions of books and can respond to any topic.

Key Terms

1. **Token:** The smallest unit of text (a word or part of a word).
2. **Parameter:** An adjustable component of the model—like a brain cell that stores learned patterns.
3. **Prompt:** The input or question given to the model.
4. **Fine-Tuning:** Customizing a model using specific datasets to improve performance.
5. **Inference:** The model's generated response or output.

Analogy: Tokens are ingredients, parameters are the recipe rules, and inference is the final dish.

Evolution of Language Models

Year	Model	Parameters	Creator
2018	GPT-1	117M	OpenAI
2019	BERT	110M	Google
2020	GPT-3	175B	OpenAI
2023	LLaMA, Claude	~70B–100B	Meta, Anthropic
2024	GPT-4o, Gemini 1.5	~200B	OpenAI, Google

How does an LLM work?

1. **Prompt typed:** "What is the capital of France?"
2. **Model tokenizes the prompt:** ["What". "is", "the", "capital", "of", "France", "?"]
3. The transformer processes the tokens through multiple layers using attention mechanisms.
4. **The model predicts the next best token:** "Paris".

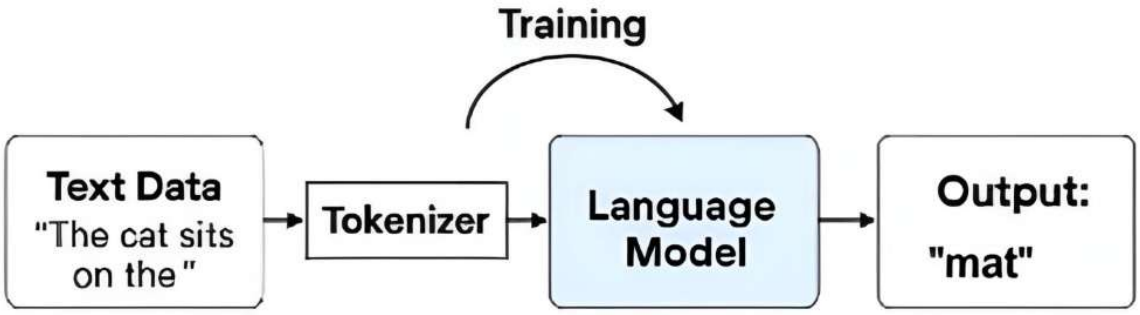
Output: A smart, context-aware response.

Key Insight: LLMs do not truly understand like humans, they predict the most likely next tokens based on patterns learned from training data.

Inside the LLM – Transformers

- **Self-attention:** Helps the model identify which words in a sentence are most related to each other.
- **Feed-forward networks:** Learn deeper patterns and refine the meaning extracted from attention.
- **Positional encoding:** Allows the model to remember word order, since transformers do not process text sequentially.

Analogy: Reading a sentence by emphasizing important words depending on the context.



- **Flowchart Explanation:**

1. Text Data (Input):"

- This is raw text provided to the model during training or inference.
- The model cannot work directly with text, so it must be converted.

eg. "The cat sits on the"

2. Tokenizer:

- Breaks the sentence into tokens (small units such as words or sub-words).
- Converts these tokens into numbers (IDs) the model can understand.

eg. tokenization:

["The", "cat", "sits", "on", "the"]

3. Language Model (Transformer):

- This is where all the learning and prediction happens.
- The model looks at the input tokens and tries to predict the next token.
- Uses self-attention, positional encoding, and feed-forward layers during processing.

4. Output (Prediction):

- Based on patterns learned from training, the model predicts the next likely word.
- The model generates this because it has seen similar phrase patterns ("The cat sits on the mat") millions of times during training.

eg. prediction: "mat"

- The flow is:

Text → Tokens → Model learns patterns → Predicts next word

- LLMs don't "understand" like humans — they predict the most likely next token using everything they've learned.

Training LLMs

1. Pretraining: The model reads massive amounts of text and learns to predict the next word.

2. Fine-tuning: The model is refined on specialized datasets (e.g., medical or legal text) to perform better in specific domains.

3. RLHF (Reinforcement Learning with Human Feedback): Humans rate or correct model responses, helping it learn which answers are better.

eg.: Like teaching a parrot basic words first (pretraining), then training it on specific phrases (fine-tuning), and finally correcting it whenever it makes mistakes (RLHF).

Applications of LLMs

LLMs are widely used in:

- Chatbots & virtual assistants
- Education & tutoring
- Healthcare support
- Legal document drafting

- Creative writing
- Automation tools

A demo prompt was shown demonstrating real-time AI responses.

Prompting: How to Talk to LLMs

The trainer explained best practices:

- Clear instructions
- Role assignment
- Context + constraints
- Examples to guide the model

This helps improve accuracy and reliability of AI outputs.

Limitations of LLMs *(from screenshots)*

1. **Hallucinations** – Can produce false or misleading information.
2. **Bias** – Reflects bias present in training data.
3. **No Real Understanding** – Does not "think"; only predicts patterns.
4. **Context Length Limitations** – Restricted memory span.

Tip: Always verify important outputs.

Ethical Concerns

The session addressed responsible AI use, highlighting issues such as:

- **Misinformation generation**
- **Data privacy risks**
- **Deepfakes & impersonation**
- **AI responsibility & fairness**

A case study from the slides was discussed:

- **Microsoft's Tay chatbot**, which malfunctioned due to biased inputs within 24 hours.

Quiz Session (Based on Screenshots)

An interactive quiz reinforced learning. Sample questions included:

- **What does LLM stand for?** → Large Language Model
- **Which architecture powers LLMs?** → Transformer
- **What is a token?** → A language unit
- **Which field is least associated with LLMs?** → Agriculture
- **Limitations of LLMs?** → Hallucinations
- **Prompt engineering helps with…?** → Better outputs

Screenshots showed real-time responses and checkmark-marked answers.

The Future of LLMs

According to the closing slides, the future will feature:

- **Multimodal models** (text + image + audio)
- **Autonomous agents** (e.g., Devin-like systems)
- **Faster open-source models**
- **On-device LLMs** (Apple, Android)

Participants discussed where they see LLMs being used next.

Further Learning & Resources

The trainer recommended useful sources (from screenshots):

- OpenAI Cookbook
- Hugging Face courses
- YouTube: CodeEmporium – GPT Explained Simply
- **Books:**
    - *You Look Like a Thing and I Love You* – Janelle Shane

- *Artificial Intelligence: A Guide for Thinking Humans* – Melanie Mitchell

Conclusion

Day 3 offered a comprehensive understanding of how LLMs work, their inner architecture, key concepts, ethical challenges, and future developments. Through explanations, diagrams, analogies, and quizzes from the presentation slides, participants gained a strong conceptual foundation for working with modern AI systems. The session strengthened their knowledge of LLM operations, prompting strategies, and the responsibilities associated with deploying AI responsibly.