

Introduction to LLMs

Lara Wehbe - TheAIEngineers - October 2024

Outline

- Introduction to LLMs
- Understanding Transformers Architecture
- Pre-trained models and Fine-tuning
- Vector DBs and RAG Architecture
- Training LLMs
- Practical Applications of LLMs
- Ethical Considerations and Limitations

Draft

Break down the course into modules. Here's a suggested outline:

Module 1: Introduction to LLMs

- What are LLMs?
- Historical context (from GPT-1 to GPT-4, BERT, etc.)
- Applications in real-world scenarios (chatbots, content generation, etc.)

Module 2: Understanding Transformer Architecture

- Overview of transformers
- Attention mechanisms
- Self-attention and cross-attention

Module 3: Pre-Trained Models and Fine-Tuning

- Overview of pre-trained models (GPT, BERT, T5, etc.)
- Fine-tuning vs. training from scratch
- Working with popular libraries (Hugging Face, OpenAI, etc.)

Module 4: Training LLMs

- Datasets and tokenization
- Training from scratch
- Evaluation metrics

-

Module 5: Practical Applications of LLMs

- Building chatbots and virtual assistants
- Content creation (text generation, summarization)
- Translation and Q&A systems

Module 6: Ethical Considerations and Limitations

- Bias in LLMs
- Ethical concerns (misuse, misinformation)
- Addressing fairness and transparency

3. Hands-On Projects

Include projects to reinforce learning:

- Fine-tune GPT-3 for a custom task.
- Build a chatbot using Hugging Face Transformers.
- Train a language model on a small custom dataset.

4. Tools and Frameworks

- **Hugging Face Transformers:** For model loading, fine-tuning, and deployment.
- **OpenAI GPT models:** If you want to focus on specific GPT architectures.
- **Google Colab** or **Jupyter Notebooks** for coding exercises.
- **PyTorch** or **TensorFlow** for advanced students interested in model training.

Introduction to LLMs

LLMs

Definition

What are LLMs?

LLMs are advanced AI models designed to understand and generate human language.

At its core, a large language model is a type of machine learning model that can understand and generate human language via deep neural networks.

Why are they Important?

They revolutionize tasks like content creation, virtual assistants, translations, and more.

LLMs

Definition

Why is it “Large”?

What makes a language model “large” is the size of its architecture. This, in turn, is based on the artificial intelligence of neural networks, much like the human brain where neurons work together to learn from and process information.

LLMs

Definition

Why is it “Large”?

What makes a language model “large” is the size of its architecture. This, in turn, is based on the artificial intelligence of neural networks, much like the human brain where neurons work together to learn from and process information.

Transformers Architecture

Pre-trained models and Fine-tuning

RAG Architecture

Training LLMs

Practical Applications

Ethical Considerations