

Advanced LLMs Programme for Software Engineers

This comprehensive training on Large Language Models equips participants with the skills to train, fine-tune, and deploy advanced LLMs effectively. Attendees will gain proficiency in areas like model architecture, data handling, and ethical alignment, enabling them to create foundational models from scratch and adapt them to the goals or domains. This training requires a background in computer science.

14 MODULES OF 6 HOURS EACH – 2 INTENSIVE WEEKS

START DATE: APRIL 1st, 2024 – PRICE: 5000€.

History of Language Models

4 hours interactive lecture

The history of language models traces back to the early days of computational linguistics, evolving from simple rule-based systems to complex neural network-based models. Initially, language models were based on statistical methods, like n-gram models, which predicted words based on the preceding 'n' words. The advent of machine learning led to the development of more sophisticated models, including Hidden Markov Models and, later, neural network approaches like RNNs and LSTMs. The breakthrough came with the introduction of transformer models, such as GPT and BERT, which significantly improved performance in natural language processing tasks. This session will explore this historical progression, emphasising key milestones and technological advancements.

State-of-the-art of LLMs

2 hours lecture + 6 hours workshop



PI SCHOOL

MACHINE INTELLIGENCE
MEETS HUMAN CREATIVITY

Current state-of-the-art Large Language Models (LLMs) like OpenAI GPT, Anthropic Claude, Meta LLaMA and others represent significant advancements in natural language understanding and generation. These models, built on transformer architectures, have demonstrated remarkable abilities in generating coherent and contextually relevant text. This session will delve into the latest achievements in LLMs, including their capabilities in various NLP tasks such as translation, summarisation, and question-answering. We'll also discuss the ongoing research challenges, such as improving model efficiency and handling nuanced language understanding. The session aims to provide a comprehensive overview of where LLM technology stands today and its potential future trajectory.

Advanced Prompting

2 hours lecture + 4 hours workshop

Advanced prompting techniques enable more effective interactions with LLMs, enhancing their ability to generate desired outputs. This training segment focuses on crafting prompts that leverage the model's strengths and mitigate limitations. We'll explore strategies like prompt engineering, which involves structuring input text to guide the model towards specific types of responses. Additionally, we'll cover techniques for iterative refinement of prompts, using model feedback to hone in on more accurate and relevant outputs progressively. By mastering advanced prompting, users can significantly improve the utility and precision of LLMs in various applications.

The Architecture of an LLM

4 hours lecture + 2 hours workshop

The architecture of a Large Language Model is pivotal to its functioning. This training will provide a deep dive into the transformer architecture, the backbone of modern LLMs. We will dissect the key components like attention mechanisms, which allow the model to weigh different parts of the input text





differently, and positional encodings, which give the model a sense of word order. Understanding these architectural elements is crucial for appreciating how LLMs process and generate language and making informed model customisation and application decisions.

Data Preparation

2 hours lecture + 4 hours workshop

Data preparation is critical in training LLMs, impacting their performance and capabilities. This session covers the processes of data collection, cleaning, and preprocessing. We'll discuss strategies for building a diverse and comprehensive dataset, techniques for cleaning and normalising text data, and considerations for ethical and unbiased data curation. Additionally, the importance of data variety to prevent overfitting and ensure the model's generalizability will be emphasised. This foundational knowledge is essential for anyone training or fine-tuning language models.

Tokenisation & Embeddings

2 hours lecture + 4 hours workshop

Tokenization and embeddings are fundamental concepts in the operation of LLMs. Tokenisation involves breaking down the text into smaller units (tokens), while embeddings translate these tokens into numerical vectors that the model can process. This training will cover various tokenisation methods, like subword tokenisation, and their impact on model performance. We will also delve into the concept of embeddings, exploring how they capture semantic and syntactic information and the role of contextual embeddings in modern LLMs. Understanding these concepts is crucial for effectively working with and modifying language models.





PI SCHOOL

MACHINE INTELLIGENCE
MEETS HUMAN CREATIVITY

Training

2 hours lecture + 4 hours workshop

Training Large Language Models is a complex process requiring a deep understanding of various aspects like model architecture, data, and computational resources. This session will cover the training process, including setting up the training environment, selecting appropriate loss functions, and understanding backpropagation in LLMs. We'll also discuss strategies for efficient training, such as distributed training and mixed-precision computing, which are crucial for handling large-scale models. This training aims to provide the necessary knowledge to train or fine-tune LLMs effectively for specific tasks or datasets.

Deployment & Inference

2 hours lecture + 4 hours workshop

Deploying and running inference with LLMs presents unique challenges, especially considering their size and computational requirements. This training covers the end-to-end process of deploying LLMs, from model compression techniques to reducing size and computational demands and choosing the right deployment platforms. We'll discuss inference techniques, batch processing, and real-time processing considerations.

Additionally, we'll explore monitoring and maintaining model performance in production, ensuring reliability and efficiency in real-world applications.

Context Injection

6 hours workshop

Context injection in LLMs is a technique to enhance the model's understanding and generation of text by providing relevant background information. This training will discuss methods to effectively incorporate





context into prompts or model inputs, enhancing the relevance and accuracy of the output. We'll explore different approaches to context injection, including explicit context addition and designing prompts that implicitly guide the model to consider the desired context. This session aims to equip participants with skills to utilise context to improve LLM performance in various applications effectively.

Fine-tuning

6 hours workshop

Fine-tuning is a process of adjusting a pre-trained LLM to perform better on specific tasks or datasets. This session will cover the essentials of fine-tuning, including selecting an appropriate base model, preparing task-specific datasets, and adjusting training parameters to adapt the model to new tasks. We'll discuss strategies to avoid common pitfalls like overfitting and how to evaluate the performance of the fine-tuned model. This knowledge is vital for anyone looking to customise LLMs for specialised applications.

Alignment

6 hours workshop

Alignment in the context of LLMs refers to aligning the model's outputs with desired behaviours or ethical guidelines. This training will focus on methods to ensure that model outputs are accurate and align with ethical, cultural, and social norms. We'll discuss techniques like reward modelling and reinforcement learning from human feedback to guide models towards generating safe, fair, and unbiased content. Understanding and implementing alignment is crucial for responsible AI development and deployment.

Instruct Fine Tuning

6 hours workshop





Instruct fine-tuning involves training LLMs to follow instructions more effectively. This segment will cover techniques to improve the model's ability to understand and execute task-specific instructions. We'll explore the training process with instruction-based datasets, where models learn to generate outputs that closely follow the given directives. This approach is particularly useful in applications requiring precise control over model outputs, such as content generation or automated responses.

Chain of Thought Fine Tuning

6 hours workshop

Chain of Thought fine-tuning focuses on improving the LLM's ability to solve complex problems by mimicking human-like reasoning processes. This training will delve into methods to train models to generate intermediate steps or explanations along with their final output, enhancing transparency and understandability. We'll explore dataset preparation and training techniques to encourage the model to articulate its reasoning process. This approach is particularly beneficial for problem-solving, reasoning, and explanation generation tasks.

Multimodal Training

6 hours workshop

Multimodal training involves training LLMs to process and integrate multiple data types, such as text, images, and audio. This session will cover the basics of multimodal AI, discussing how to integrate different data modalities into a unified model. We'll explore the challenges and techniques in multimodal data processing, model architecture adaptations for multimodal inputs, and the applications of multimodal LLMs in various fields. This training is essential for understanding the expanding capabilities of AI beyond text-based processing.

