
Editorial: Evolution of large language models and their role in shaping general artificial intelligence

Editorial

1

The rise of large language models (LLMs) represents a significant milestone in the development of artificial intelligence. Originally designed to comprehend and generate text that resembles human language, these models have surpassed their original intentions. The foundation for these advancements can be attributed to a distinctive neural network architecture called transformer, which was introduced in the seminal paper “Attention is All You Need.”

Since their inception, LLMs are designed to learn probability distributions to predict the next word in a sequence of words. This capability grants them the versatility to tackle a large spectrum of tasks, such as completing prompts and outperform in domains, including but not limited to answering queries, aiding in writing, translating languages, coding assistance, engaging in conversations, facilitating, brainstorming sessions, offering feedback and managing projects.

The “Attention is All You Need” paper marks a groundbreaking shift in the field of natural language processing with the introduction of the “attention mechanism.” This concept enabled more efficient and powerful handling of textual data, paving the way for the development of models such as ChatGPT. ChatGPT serves as a distinguishing demonstration of the capabilities made possible by the transformer technology. Its effectiveness and functionality rely not only on its architecture but also on extremely large dataset resources and substantial computational capabilities.

A crucial query arises: How do these systems evolve into versatile AI tools that can perform tasks beyond their original design? The emerging capabilities of LLMs such as GPT 4 and Gemini demonstrate their potential to go beyond pre-defined functions. Put simply, their functioning fundamentally relies on making guesses to fill in blanks, often in a manner that is both plausible and stylistically fluent for humans. Users should thus be careful when interacting with LLMs, assuming there is logic and reasoning behind their answers.

Despite their impressive capabilities, LLMs are not without their challenges. Their lack of reasoning capabilities means that incorrect information can be generated in a way that seems convincingly plausible. A notable example is a New York lawyer who potentially faced sanctions for referencing fake cases generated by ChatGPT in a legal brief. This incident emphasizes the need for users to verify the generated content. The legal and ethical implications surrounding LLMs are substantial. The recent lawsuit by *The New York Times* against OpenAI and Microsoft also highlights the complexities surrounding the use of copyrighted and personal data in training these models. Yet another intriguing example is the European Union’s General Data Protection Regulation (GDPR), which poses unique



DTS
3,1

challenges for LLMs. Erasing specific data from these models in the context of the right to be forgotten is virtually impossible.

Looking forward, the potential of LLMs as building blocks in the quest for artificial general intelligence (AGI) is immense. Immediate focus areas include enhancing LLMs with self-fact-checking capabilities and reducing bias and toxicity. The concept of AI alignment is crucial in this regard, ensuring that LLMs adhere to human values and goals. Moreover, the evolution toward multimodal language models promises to expand the capabilities of LLMs beyond mere text generation.

2

The excitement surrounding the breakthroughs in LLM technology in recent years is profound. As we embark on the AI revolution and evolution, the future of these models and their impact on various facets of our society and its digital transformation remains a thrilling and open-ended journey.

YouaKim Badr

Pennsylvania State University, Malvern, Pennsylvania, USA