

Large Language Models: What They Are, Why They Are Important, and What They Fail At

Roberto Navigli¹

¹*Sapienza University of Rome, Italy*

Abstract

The advent of Large Language Models (LLMs) like GPT-4 represents a significant leap forward in the field of Artificial Intelligence, offering unprecedented capabilities in understanding, generating, and interacting with human language. This talk aims to demystify these complex systems, explaining their fundamental architecture, how they are trained on vast datasets, and the underlying technologies that enable their sophisticated processing abilities. We will explore the importance of LLMs, highlighting their role in driving innovation, enhancing productivity, and opening new avenues for human-computer interaction. However, with great power comes great responsibility, and LLMs are not without their limitations and challenges. This presentation will critically examine the inherent weaknesses of LLMs, such as biases in training data, the potential for generating misleading information, and ethical concerns. We will delve into real-world examples to illustrate these failures, offering a balanced perspective on the capabilities and limitations of these models. Finally, I will overview ongoing research in my group aimed at mitigating these shortcomings, including extracting facts from generated text and interconnecting them to the source text.

Disclaimer: 90% of this abstract was generated by GPT-4.

CILC 2024: 39th Italian Conference on Computational Logic, June 26-28, 2024, Rome, Italy

✉ navigli@diag.uniroma1.it (R. Navigli)

🆔 0000-0003-3831-9706 (R. Navigli)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).