


Chapter 3


Exploring the Core Principles, Architectures, and Advancements of Large Language Models

Suriya Murugan

 <https://orcid.org/0000-0002-8130-7228>

Sri Eshwar College of Engineering, India

Anandakumar Haldorai

 <https://orcid.org/0000-0001-9975-6462>

Sri Eshwar College of Engineering, India

ABSTRACT

Language is necessary for human interaction because it helps in communication, learning, and the creation of rules and agreements that advance society. Although computers have historically had difficulty understanding and producing human language, advances in artificial intelligence, especially with regard to large language models (LLMs), have significantly enhanced computers' ability to grasp and produce natural language. LLMs, which are based on deep learning architectures such as transformers, have revolutionized natural language processing (NLP) by facilitating complex text generation and comprehension. Even with their successes, LLMs have drawbacks, like as theoretical issues and possible political and cultural biases in the training set. This chapter explores the underlying theories and architectural frameworks of LLMs, as well as the principles that guide their operation. The impact of these potent models on how humans engage with information and technology is highlighted, and it also covers current study on the difficulties and moral issues surrounding their use.

DOI: 10.4018/979-8-3693-8387-2.ch003

INTRODUCTION

The advent of language processing technology has been a recent milestone in artificial intelligence (AI), allowing us to create more intelligent systems with a deeper comprehension of language than ever before. “Large Language Models (LLMs),” also known as large pre-trained transformer language models, significantly increase the capabilities of text-based systems. The purpose of LLMs is to comprehend and produce text that is similar to that of humans. They are trained on enormous datasets that contain text from books, journals, websites, and several other sources. They are constructed using neural networks, more especially deep learning (DL) architectures, (Bommasani et al., 2022a).

Figure 1. Outline of proposed workflow



As outlined in Figure 1, this chapter will delve into the theoretical underpinnings of LLMs, examining the principles that drive their success and the foundational theories that enable their remarkable performance. While challenges persist, ongoing research and a commitment to ethical AI practices pave the way for a future where large language models continue to shape our interaction with information and technology. It is essential to recognize the potential, challenges, and responsibilities

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/exploring-the-core-principles-architectures-and-advancements-of-large-language-models/390562

Related Content

Legal and Policy Frameworks for Securing Large Language Models

Dilshad Ahmad Mhia-aldin (2026). *Securing Large Language Models Against Emerging Threats* (pp. 39-80).

www.irma-international.org/chapter/legal-and-policy-frameworks-for-securing-large-language-models/394790

Understanding and Applying Machine Learning Models

Leena Suresh Moreand Binod Kumar (2024). *The Pioneering Applications of Generative AI* (pp. 274-309).

www.irma-international.org/chapter/understanding-and-applying-machine-learning-models/350788

Exception Handling

(2023). *Principles, Policies, and Applications of Kotlin Programming* (pp. 100-115).

www.irma-international.org/chapter/exception-handling/323933

Functions

(2024). *Advancements, Applications, and Foundations of C++* (pp. 160-216).

www.irma-international.org/chapter/functions/345782

Diving Deep Into the World of LLMs

Mohamad Saleh (2026). *Leveraging LLMs for Business Innovation: Practical Solutions and Future Trends* (pp. 1-40).

www.irma-international.org/chapter/diving-deep-into-the-world-of-llms/401810