

The Feasibility Study of a Large Language Model (LLM)

Introduction

In recent years, large language models (LLMs) like ChatGPT have revolutionized natural language processing, demonstrating unprecedented capabilities in understanding and generating human-like text. These models have far-reaching implications across various domains, from healthcare and education to creative industries and scientific research. This master's thesis assessed the feasibility of developing a model comparable to GPT-4.

Research Gap

The study identifies a significant research gap in the development of large language models (LLMs) within Europe. While American tech giants lead in LLM development, there is a notable absence of European-developed models of comparable scale. This gap puts European companies at a competitive disadvantage and risks creating dependence on American technologies. It also potentially hinders Europe's ability to drive innovation in the field of artificial intelligence. This gap underscores the need for a comprehensive feasibility study to assess the potential for developing a European LLM, addressing both technical challenges and strategic importance.

Technical Feasibility

The study demonstrates that developing a large language model is technically feasible, albeit with significant resource requirements. The project demands approximately 15 million GPU hours using NVIDIA H100 graphics cards, highlighting the substantial computational power needed. With an estimated training time of 6 months, the scale and complexity of the undertaking are evident. The development process requires advanced software tools, frameworks, and high-quality datasets for successful model training. A multidisciplinary team of 14 experts, including AI researchers, data scientists, and software engineers, is deemed necessary to navigate the technical challenges effectively.

Financial Feasibility

The financial analysis reveals that developing a LLM requires substantial initial investment but also offers significant revenue potential. The study estimates initial capital costs at 53,298,400 EUR, covering both computational resources and the dedicated development team. Despite this substantial upfront investment, the project demonstrates strong financial viability. Projections indicate a cumulative revenue of 245,436,826 EUR by the end of the first year of deployment, highlighting the model's potential for rapid return on investment. The financial model predicts reaching the break-even point in the 12th month, with an impressive ROI of 270.57% over 18 months.

Risk and Challenges

The study identifies significant risks and challenges associated with LLM development including intense competition from tech giants and innovative startups, rapidly evolving regulatory environments, and potential public backlash over AI ethics concerns. Technical risks encompass development complexity, the need for continuous innovation, and potential technological disruptions. Ethical considerations, such as the risk of perpetuating biases or generating harmful content, require mitigation strategies and ongoing monitoring. It also points out legal and regulatory uncertainties, particularly around AI-specific regulations, necessitating proactive engagement with policymakers and legal experts.

Abstract

This master's thesis investigates the feasibility study of a large language model (LLM) comparable to GPT-4, the state-of-the-art model developed by OpenAI. The thesis analyzes the technical requirements, financial aspects, ethical considerations, and practical applications to provide a comprehensive overview of the opportunities and challenges associated with developing such a model.

Key Words

large language model, LLM, GPT-4, feasibility study, artificial intelligence, machine learning, natural language processing

