This Poster was made with the help of AI

Oxford Brookes University

Can ChatGPT be used to detect “FakeNews”?

June 27, 2024 | GenAI

Introduction

In today's digital landscape, fake news undermines academic integrity and critical thinking skills in higher education. Our research project uses Large Language Model (LLM) technology to develop an advanced model for detecting fake news, aiming to mitigate misinformation's spread in academic and societal contexts.

Literature Review

This literature review examines the critical role of artificial intelligence (AI) in combating fake news, particularly focusing on advanced language models like our project's Llama-based system. Key points include:

Information Literacy (IL) is crucial in today's world, serving as a powerful tool against misinformation. Our AI model aims to enhance IL by providing an automated, accurate fake news detection system.

Advanced language models, like the Llama series we're using, are transforming information access. Just as ChatGPT is guiding universities, our model seeks to assist librarians and students in identifying misleading information

Research recommends incorporating AI into IL activities. Our project embodies this by developing a tool that can be integrated into library resources for real-time fake news detection.

The rise of AI necessitates improvements in IL training. Our project contributes by providing a practical tool that encourages critical evaluation of news sources.

Current research, like our project, is exploring AI models that use deep learning and natural language processing to detect fake news by analyzing news sample structures and features.

Methodology

Our research project aims to develop a robust fake news detection system using Meta's evolving Llama models. Initially focusing on UK-based data, we encountered a scarcity of regional datasets, leading us to adapt US resources like BuzzFeed and PolitiFact from repositories such as Papers With Code and Kaggle.

A distinctive feature of our methodology is rigorously testing our AI's resilience. We introduce fabricated news articles to observe how it handles errors, helping us refine its ability to flag inconsistencies. To determine news authenticity, our model employs sophisticated techniques: analyzing linguistic cues, cross-referencing claims, and evaluating author credibility. We've even tested Llama 3 with internet access to explore how real-time fact-checking enhances its true/false determinations.

Recognizing the global nature of fake news, we're also tackling language barriers. Using custom Spanish datasets, one standard, another with broken source links—we test our model's adaptability to non-English news. This approach helps us understand how well it transfers learned patterns across languages, a crucial factor in today's interconnected world.

Our transition from Llama 2 to Llama 3 has yielded significant improvements, particularly in cross-lingual performance and error handling. This progression suggests that more advanced models are better equipped to navigate the complexities of fake news detection, adapting to diverse linguistic structures and cultural contexts. By focusing on error response, evidence-based validation, and language adaptability, our project strives to create a universally effective tool against misinformation.

Conclusion

Our LLM-based model effectively detects misinformation across languages and scenarios. While results are promising, there's room to expand dataset diversity and refine complex pattern detection. This work highlights AI's role in enhancing information literacy and preserving academic integrity.

By:
Aaron Worsley
Nikhil Kamaraj Ranjana
Simon Llewellyn
Jamelia Dominguez