Data Challenge Report

Critical Text Analysis Tool

23.01.2024

Amira Chaib

Iris van den Boomen

Introduction

The goal of this project was to develop a tool that assists users in reading text more consciously, drawing attention to bias and manipulation. We envisioned a tool with critical text analysis capabilities, capable of highlighting logical fallacies and identifying lapses in journalistic standards. This tool aimed to empower users to discern biases and manipulations in the content they encounter.

Key Objectives

- Implementing a user-friendly tool for critical text analysis.
- Pointing out sections that belong to selected categories, for the MVP the required category was logical fallacies
- Explanation of highlighted sections to enhance user understanding.

Target Audience

Various user roles, including Researchers, Journalists, PR Specialists, and Legal Professionals. For more information refer to Target Audience / Stakeholders file.

Approach and Methodology

Planning Document: Developed the overall idea, identified the problem, and outlined the ethical aspects. Decided to focus on logical fallacies and journalistic standards.

Research Document: Formulated research questions, explored critical text analysis, logical fallacies, and journalistic ethics. Investigated potential models, datasets, and outlined the methods for developing the tool.

Application Development: Created a GitHub repository for version control and a GitHub Project to keep track of tasks and progress. Utilized Langchain for natural language processing, Streamlit for the user interface, and GPT-3.5 as the Large Language Model. Developed a user-friendly interface with real-time text analysis and explanation features. Considered ethical aspects by using TICT.

Achievements

- Implemented a simple interface allowing users to input text and choose categories for analysis.
- Incorporated Langchain and GPT-3.5 to provide real-time analysis and highlight sections based on logical fallacies and political intent.
- Providing explanations for the model's categorisation to enhance understandability.

- Created a mockup to visualize the tool's design and functionality.
- Identified diverse user roles, including Researchers, Journalists, PR Specialists, and Legal Professionals.
- Utilized Langchain, Streamlit, and GPT-3.5 to create an effective first prototype.
- Short presentation of the tool including a Demo.

Considerations

Since users input their own text for analysis, a secure and transparent experience is ensured. We experimented with a few models, ultimately choosing GPT-3.5 for its balance of performance and ease of use.

Future Considerations

- Expanding the tool to include more categories than logical fallacies and political intent.
- Usage of a proper frontend framework to make text-highlighting possible
- Exploring the effectiveness of different language models for specific use cases.

Conclusion

The Critical Text Analysis Tool represents a first prototype of our aim to empower users to read text more consciously. By combining advanced natural language processing techniques with user-friendly design, the tool provides a valuable resource for individuals across various professions. Continuous improvements, like a proper frontend using a non-Streamlit framework, and expansions will keep the tool relevant and effective in the changing way people consume information.