Problem Statement

The aim of this data science project is to utilise advanced Natural Language Processing (NLP) techniques and supervised machine learning models for effectively detecting and classifying news articles into categories such as business, sports, entertainment, tech, or politics. Additionally, the project intends to implement a user-friendly news research tool designed for effortless summary retrieval. Users can input an article, and the fine-tuned LLM model will predict the category (business, sports, entertainment, tech, or politics) while providing a summary from that article.

Primary Objectives

The primary objectives of this project are to leverage Natural Language Processing (NLP) and supervised machine learning techniques to achieve two key tasks: automatically predicting the category (business, sports, entertainment, tech, or politics) of a given news article and providing a user-friendly news research tool for easy summary retrieval.

Significance and Impact

This innovative project holds substantial significance in the realm of information processing and user engagement. By harnessing advanced Natural Language Processing (NLP) techniques and supervised machine learning models, the system promises to revolutionise the way we interact with news content. The ability to automatically classify news articles into distinct categories, including business, sports, entertainment, tech, and politics, streamlines information retrieval, enhancing user experience and efficiency.

Moreover, the implementation of a user-friendly news research tool is poised to have a lasting impact on how individuals access and digest news. The tool's effortless summary retrieval feature not only caters to the fast-paced nature of modern information consumption but also accommodates users with varying levels of expertise in the respective domains.

Beyond individual user benefits, the project contributes to the broader landscape of NLP applications and machine learning advancements. The fine-tuned Language Model (LLM) plays a pivotal role in not just predicting categories accurately but also generating concise summaries, showcasing the potential for intelligent content analysis.

In a world inundated with information, the project's outcomes have the potential to streamline and enhance the way we engage with news, making it a noteworthy endeavour with far-reaching implications for both individual users and the field of data science at large

---