Project Proposal: Stance Detection System for combating Fake News

Group member: LuMingrui

Student ID: A0307208L

1. Introduction

The rise of fake news presents a significant challenge in modern society, with misinformation spreading rapidly across digital platforms. Nowadays the most effective approach to combating fake news is to leverage artificial intelligence, particularly machine learning and natural language processing, to automate the fact-checking process. The goal of my project is to develop a stance detection system that classifies the relationship between the headline and body of news articles, serving as a foundational step in detecting fake news.

Project Goals:

- Develop a machine learning-based stance detection system that classifies news articles into four categories: *agree*, *disagree*, *discuss*, and *unrelated*.
- Utilize NLP techniques to enhance the system's ability to process and analyze text data, improving efficiency and scalability in fake news detection.
- Build upon existing research and methodologies to create an effective and practical solution for real-world applications.

2. Project Background And Market Research

Fake news has become a critical problem, leading to widespread misinformation and confusion. This project addresses the need for an automated system that can assist in identifying the stance of news articles relative to their headlines. Currently, fact-checking is a labor-intensive process, but AI-based systems like the stance detection classifier have the potential to significantly reduce the workload of human fact-checkers by pre-filtering and analyzing content.

From a market perspective, the demand for automated fact-checking tools and fake news detection systems has surged, driven by the rapid proliferation of misinformation across social media platforms, online news outlets, and other digital channels. A lot of company have entered the space, offering solutions to mitigate the spread of fake news. Companies like Factmata, Full Fact, and Logically have developed AI-powered systems designed to detect and flag misleading or false information. These tools often rely on machine learning and NLP to analyze news articles, social media posts, and other online content.

Although the progress made by these competitors, many existing solutions focus on detecting misinformation by verifying factual claims, rather than addressing the stance detection task. This presents a potential gap in the market that this project aims to address. By focusing specifically on stance detection, my proposed system offers a unique solution that can be used as a pre-filter in larger fact-checking pipelines. Rather than directly determining whether a claim is true or false, the system will first analyze the relationship between a headline and its corresponding body text, streamlining the fact-checking process for

human reviewers or other automated systems.

3. Project Scope

The scope of my project is to focus on the development and implementation of a stance detection system that classifies the relationship between a news headline and its body text. The system will be evaluated on its ability to categorize news content into four distinct stances: *agree*, *disagree*, *discuss*, or *unrelated*.

Key Aspects of Intelligent Reasoning:

- Use of machine learning and NLP techniques for text classification.
- Exploration of feature engineering techniques like word overlap, sentiment analysis, and TF-IDF vectors to enhance the model's accuracy.
- Integration of advanced NLP models such as BERT to understand and classify complex relationships between text segments.

The project will focus primarily on contributing to the field of intelligent reasoning systems by improving existing stance detection methods. The scope does not include deployment in real-world environments but will provide a proof of concept that can be extended in the future.

4. Data Collection and Preparation

The dataset for this project will be derived from the Fake News Challenge (FNC-1) competition, which consists of labeled pairs of headlines and body texts from news articles. The stance between the headline and the body is categorized as either *agree*, *disagree*, *discuss*, or *unrelated*. The data will be pre-processed to remove noise, normalize text, and prepare it for model training.

Key steps in data preparation will include:

- **Tokenization and Lemmatization**: Breaking down the text into tokens and reducing them to their base forms to ensure uniformity.
- **Stopword Removal**: Removing common but uninformative words to reduce data complexity.
- **Feature Engineering**: Implementing techniques such as word overlap, refuting features, and sentiment polarity to extract meaningful insights from the text.

Challenges related to class imbalance such as there is significantly fewer examples in the *disagree* class will be addressed through techniques such as oversampling or undersampling.

5. System Design

The system will be structured as a machine learning pipeline with the following components:

1. **Preprocessing Module**: Cleans, tokenizes, and normalizes the input text.
2. **Feature Extraction Module**: Generates features such as TF-IDF vectors, sentiment scores, and word embeddings such as Word2Vec and BERT embeddings.

3. **Classification Module**: Applies and try different machine learning models such as Gradient Boosting Classifier, SVM, and deep learning-based models like BERT for stance classification.
4. **Evaluation Module**: Assesses the model's performance using accuracy, precision, recall, and F1-score, with a focus on improving stance detection.

My system design aims to implement both traditional machine learning and deep learning approaches, tuning each individually and comparing their performance to achieve the most robust performance.

6. Potential Challenges

Several challenges remain in developing a fully functional stance detection system such as:

- **Class Imbalance**: Addressing the skewed distribution of data across different stance categories, which may lead to biased model performance.
- **Complexity in Text Relationships**: Headlines and body text often have subtle relationships that may not be captured by basic feature engineering techniques, requiring try more advanced models like BERT.

In my future work, I will focus on enhancing the system by experimenting with different models, fine-tuning model parameters, and testing on larger datasets.

7. Conclusion

The stance detection system that I proposed in this project provides a foundation for combating the spread of fake news by automating the initial step of identifying the relationship between a headline and its corresponding body text. By leveraging machine learning and NLP techniques, this system has the potential to enhance the efficiency of fact-checking workflows and provide a scalable solution for news organizations and social media platforms. Although challenges such as class imbalance and the complexity of text relationships remain, the system offers a promising direction for future development and real-world application in the fight against misinformation.