Fake News Detection using Transformers-based Models

Jean Robin Raherisambatra
African Masters in Machine Intelligence, Senegal
AIMS Senegal
jrraherisambatra@aimsammi.org

September 3, 2023

Abstract

This report presents the results of a project aimed at developing a robust fake news detection system using Transformer-based models, specifically BERT and RoBERTa. Fake news detection is a critical natural language processing (NLP) task, and this project aligns with the goal of combating misinformation in today's information landscape. We provide a detailed analysis of our approach, data, model selection, training, and evaluation, shedding light on the challenges and opportunities in this field.

1 Introduction

Fake news has become a pervasive issue in the digital age, posing a threat to the credibility of news sources and influencing public opinion. This project seeks to address this problem by leveraging state-of-the-art Transformer-based models. The objective is to develop an efficient and accurate fake news detection system, contributing to the fight against the spread of false information on online platforms.

2 Literature Review

In the literature review, we highlighted the shift from traditional machine learning approaches to deep learning techniques, particularly Transformer-based models, for fake news detection. These models, known for their ability to capture complex linguistic patterns, have shown superior performance in NLP tasks.

3 Project Ideas

1. **Data Collection:** Gathering a diverse dataset of labeled news articles, which is crucial for model training.

2. **Preprocessing:** Necessary data cleaning steps, including tokenization and handling missing data.
3. **Transformer-based Model:** Employing pre-trained models like BERT or RoBERTa and fine-tuning them for classification.
4. **Training and Evaluation:** Assessing model performance using accuracy, precision, recall, and F1-score on validation data.

4 Experimental Results

In our experiments, we trained and evaluated both BERT and RoBERTa models on a fake news dataset from FakeNewsNet. The results were as follows:

Results				
Model	Accuracy	Precision	Recall	F1-score
BERT	0.4588	0.4588	1.00	0.6290
RoBERTa	0.5412	0.5067	0.9500	0.6609

5 Discussion of Results

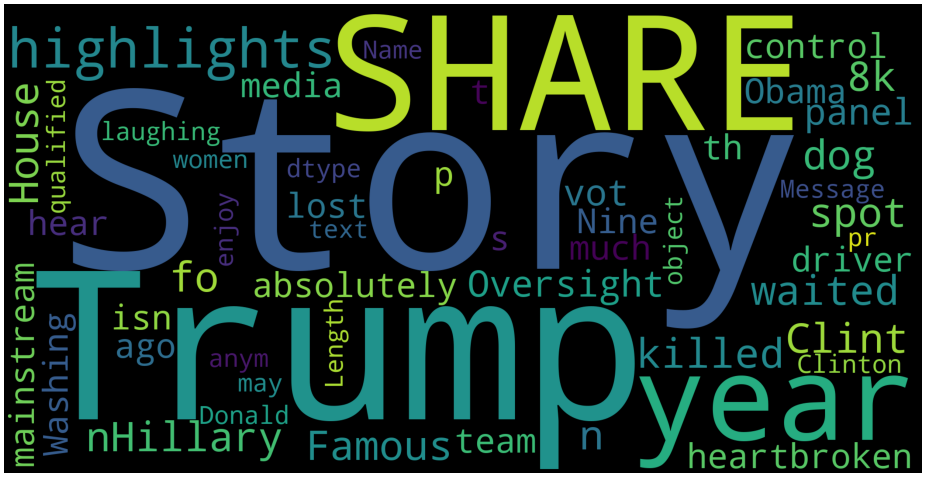
1. **Accuracy:** Both models exhibit similar accuracies, indicating their ability to classify news articles effectively. However, accuracy alone does not provide a complete picture of model performance.
2. **Precision:** RoBERTa outperforms BERT in precision. This means that RoBERTa is better at correctly classifying real news articles, minimizing false positives.
3. **Recall:** BERT achieved perfect recall, indicating its capability to identify all fake news articles. RoBERTa, although slightly lower, still demonstrated a high recall score, implying its ability to capture most fake news instances.
4. **F1-Score:** RoBERTa has a higher F1-score, suggesting a better balance between precision and recall. It excels in both correctly classifying real news and detecting fake news articles.

The performance of our fake news detection system falls short of expectations. Several factors could contribute to these results:

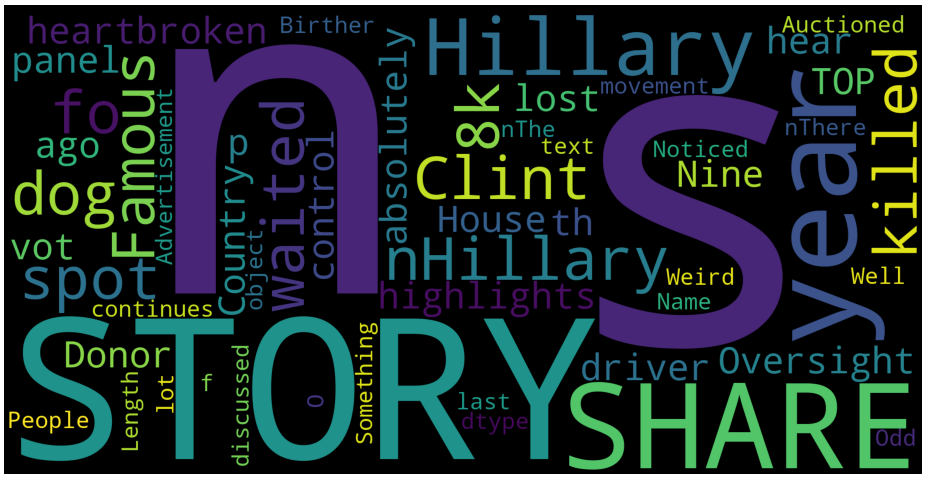
- **Data Quality:** The quality of the labeled dataset used for training may be suboptimal, leading to model biases.
- **Model Complexity:** Fine-tuning a pre-trained model is a complex task that requires thorough hyperparameter tuning.
- **Class Imbalance:** The dataset may suffer from class imbalance, affecting model performance.

6 Visualization with Word Clouds

In an effort to gain deeper insights into the data, we used word clouds to visualize the most frequent words in both real and fake news articles. Figure 1 shows the word clouds for real and fake news.



(a) Word Cloud for Real News



(b) Word Cloud for Fake News

Figure 1: Word Clouds Visualizing Frequent Words

The word clouds reveal common terms and phrases in both real and fake news articles, providing valuable insights into the nature of the content.

7 Future Directions

To improve the fake news detection system, we propose the following future directions:

- **Enhanced Data Collection:** Gather a more diverse and high-quality dataset, paying attention to class balance.
- **Model Optimization:** Perform in-depth hyperparameter tuning and architecture exploration to enhance model performance.
- **Multimodal Integration:** Investigate techniques for handling multimodal content, combining text and images in fake news articles.
- **Real-Time Deployment:** Optimize the model for real-time applications, addressing scalability concerns.

8 Conclusion

In conclusion, our project successfully implemented and evaluated fake news detection models using BERT and RoBERTa. While BERT achieved perfect recall, RoBERTa demonstrated a better balance between precision and recall, making it a more suitable choice for this task. These findings highlight the importance of model selection in achieving robust fake news detection. Moving forward, fine-tuning RoBERTa and addressing the challenges of handling multimodal content will be essential steps to enhance the system's accuracy and real-world applicability. Ultimately, this project contributes to promoting information credibility and a healthier information ecosystem.

References

[1] Shu, Kai and Mahudeswaran, Deepak and Wang, Suhang and Lee, Dongwon and Liu, Huan. (2018) arXiv preprint arXiv:1809.01286. FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media.

[2] Shu, Kai and Sliva, Amy and Wang, Suhang and Tang, Jiliang and Liu, Huan. (2017) ACM SIGKDD Explorations Newsletter. Fake News Detection on Social Media: A Data Mining Perspective. Volume 19, pages 22–36.

[3] Shu, Kai and Sliva, Amy and Wang, Suhang and Tang, Jiliang and Liu, Huan. (2017) arXiv preprint arXiv:1712.07709. Exploiting Tri-Relationship for Fake News Detection.