Detecting Media Bias in the Coverage of the 2024 U.S. Presidential Elections Using a Hybrid BERT-BiLSTM Model

The Paper's Structure

Abstract

Recognizing political bias in the media one consumes is essential for being an informed citizen. One common situation where media outlets might lean towards a certain end of the political spectrum is during the elections. This paper aims to contribute to understanding political bias in written media, particularly in the context of pivotal events. We will present both a BERT+LSTM and a BERT+CNN+BiLSTM model for detecting such biases during the 2024 U.S. presidential elections. These models classify news articles in 3 categories: favoring Harris (the left-wing, Democrat candidate), favoring Trump (the right-wing, Republican candidate), and neutral. Their performance will be evaluated on 3 custom datasets consisting of articles extracted from U.S. newspapers, as well as on the "SemEval-2019 Task 4" Dataset (Kiesel et al., 2019). The performance on the latter dataset will determine how well our models generalize to other types of media biases.

1 Introduction

Taken individually, BERT achieves state-of-the-art results on many text classification datasets, while LSTMs and CNNs achieve state-of-the-art results in the branch of sentiment analysis (Otter et al., 2019).

Additionally, BERT+CNN+BiLSTM architectures have been used for several classification tasks. However, no significant studies have been conducted on how they can be used to detect political bias in the media. Hence, the present paper takes **an approach that has not been used before** for this natural language processing application.

Notable usages of BERT+CNN+BiLSTM architectures include detecting fake news (Alghamdi et al., 2022) and performing sentiment analysis on movie reviews (He et al., 2024 and Gupta et al., 2022).

Our **hypothesis** is that combining the BERT (Bidirectional encoder representations from transformers) (presented in Devlin et al., 2019) encoder with a BiLSTM (Bidirectional Long Short-Term Memory) model – using the former's output as input to the latter, and potentially adding CNN layers between them – will create a media bias detector specialized in analyzing the written coverage of the 2024 U.S. elections that achieves state-of-the-art performance.

This investigation is driven by the following research questions:

1. How effective is a hybrid model combining BERT and BiLSTM – and potentially CNNs – in detecting media bias in written coverage of the 2024 U.S. presidential elections, compared to other text classification models?
2. What are the performance differences amongst various architectures based on BERT (BERT with just a classification layer, BERT with LSTM, BERT with CNN and BiLSTM) when applied to the detection of political bias in written media coverage of pivotal events?

3. How well do such models generalize on political media bias in general?

2 Related Work

To the author's knowledge, as of November 1, 2024, no studies have been done on building a neural network-based model specifically designed to detect bias in written-text media coverage of the 2024 U.S. presidential elections.

There have been a considerable number of studies on detecting either political bias in written media or political ideology of certain people, without focusing on the 2024 elections. Their approaches vary:

a) *Bayesian classifier* (Fang et al., 2015) or *Bayesian Hidden Markov Models* (Sim et al., 2013)

d) *statistical models such as support vector machines* (Lin et al., 2006)

c) *recursive neural networks* (Iyyer et al., 2014)

b) *fine-tuning Transformers models, such as BERT, on custom datasets* (Raza et al., 2022)

c) *LSTMs* (Baly et al., 2020)

3 Methodology

3.1 Parameters for BERT

For all three models, we will subsequentially try the parameters that the authors of BERT recommend in Devlin et al., 2019:

a) *Batch size*: 16 and 32
b) *Learning rate*: 5e-5, 3e-5, 2e-5
c) *Number of epochs*: 2, 3, and 4

3.2 BERT plus a classification output layer

This will be the simplest model we will try. Devlin et al., 2019 claims that it is enough to add an additional output layer to BERT to create well-performing models. We will add a dense output layer that predicts the class choosing from 3 options: “left” (favoring Harris), “neutral,” and “right” (favoring Trump).

3.3 BERT together with a LSTM

BERT will take the raw text as input and generate contextual embeddings. BERT’s embeddings will be fed into a LSTM. LSTMs capture context over longer sequences, as they incorporate elements called “gates” specifically for this purpose (Siami-Namini et al., 2019). We will feed their output into a dense classification layer that will choose from the three output classes.

3.4 BERT, CNN, and BiLSTM

BiLSTM is an upgrade to LSTM. As opposed to LSTMs, BiLSTMs look both in the future and in the past, while LSTMs only exploit historical context (Liu et al., 2019). By combining BERT and a BiLSTM, we hope to obtain a model whose understanding of the input is better than those of the individual models.

The raw input text will go through BERT. Then will follow several CNN layers. Using this approach, Kaur et al., 2023 obtained a classification model that “performs better than the state-of-the-art baseline approach.” The output of the CNN layers will be taken as input by the BiLSTM. In the end, a dense classification layer will choose the output class.

4 Experiments

4.1 Datasets

Lazaridou et al., 2016 builds a dataset consisting of political articles about the world and the UK using Guardian and The Telegraph articles. Similarily, we will extract articles from American newspapers to build two datasets. Additionally, we will build and use an extended version of the NewB (introduced in Wei, 2020) dataset.

4.1.1 Allsides.com

We will use the chart on Allsides.com to determine the political orientation of a newspaper. It organizes newspapers into 5 categories: left, lean left, center, lean right, and right.

4.1.2 Lean-left and lean-right newspapers dataset

The first dataset will use articles from lean-left and lean-right newspapers. We will extract sentences about all aspects of the presidential race. We will consider the sentences to be slightly biased towards the newspaper’s orientation. We will use American newspapers, both Democrat- (e.g. New York Times) and Republican-leaning (e.g. Wall Street Journal). By training on this dataset, we increase the models’ ability to recognize subtler biases. During the test stage, we expect weaker results on this dataset, compared to the left and right dataset described below, due to the aforementioned subtleties.

4.1.3 Left and right newspapers dataset

We will build a similar dataset using articles from left and right newspapers. We can say with a higher confidence that the sentences extracted from these have a certain bias.

4.1.4 Extended NewB

NewB (Wei, 2020): collection of over 200,000 sentences about Republican (right-wing) candidate Donald Trump extracted from 9 mainstream newspapers. The sentences are labeled by source. We expect the sentences to lean in the same direction on the political spectrum as the newspapers they originated from. We will extend it by adding sentences about Harris extracted from the same 9 newspapers.

4.2 Evaluation metrics

We will use the following metrics to compare the performance of the models: *accuracy*, *precision*, *recall,* and *macro F1* (suggested by Baly et al., 2020).

4.3 Results

We will present the above metrics for all models, both on the custom datasets we used for training and on the "SemEval-2019 Task 4: Hyperpartisan News Detection" dataset (Kiesel et al., 2019).

The latter dataset will offer a glimpse into how well our model can generalize to other types of media political bias detection, as it is not related to the 2024 elections. It expects a binary output signifying whether the input text is radical or not, regardless of whether to the left or to the right. To test our 3-class classifiers on this binary classification dataset, we will add a layer that maps classes "left" and "right" to radical, and class "neutral" to non-radical. If our models are successful on this dataset, our process could be viewed as solving a transfer learning task.

4.4 Comparisons

We will use the metrics described above to compare how the models perform on the 3 custom datasets and on the SemEval dataset.

5 Conclusion

In the "Conclusion" section of the paper, we will decide whether the hypothesis has been proved or not, based on an overview of the metrics described previously.

References

1. Johannes Kiesel, Maria Mestre, Rishabh Shukla, Emmanuel Vincent, Payam Adineh, David Corney, Benno Stein and Martin Potthast. 2019. SemEval-2019 Task 4: Hyperpartisan News Detection. In Proceedings of the 13th International Workshop on Semantic Evaluation, pages 829–839.
2. Daniel W. Otter, Julian R. Medina, and Jugal K. Kalita. 2019. A Survey of the Usages of Deep Learning for Natural Language Processing. In IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. XX, NO. X, JULY 2019.
3. Alghamdi Jawaher, Lin Yuqing and Luo Suhuai. 2022. Modeling Fake News Detection Using BERT-CNN-BiLSTM Architecture. 2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR).
4. Aixiang He and Mideth Abisado. 2024. Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model. In IEEE Access 12 (2024), pages 45229-45237.
5. Bhart Gupta, P. Prakasam and T. Velmurugan. 2022. Integrated BERT embeddings, BiLSTM-BiGRU and 1-D CNN model for binary sentiment classification analysis of movie reviews. In Multimedia Tools and Applications, volume 81, issue 23, pages 33067 – 33086.
6. Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186.
7. Anjie Fang, Iadh Ounis, Philip Habel, Craig Macdonald and Nut Limsopatham. 2015. Topic-centric Classification of Twitter User's Political Orientation. In SIGIR '15: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 791 – 794.
8. Yanchuan Sim, Brice D. L. Acree, Justin H. Gross and Noah A. Smith. 2013. Measuring Ideological Proportions in Political Speeches. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 91–101.
9. Wei-Hao Lin, Theresa Wilson, Janyce Wiebe and Alexander Hauptmann. 2006. Which Side are You on? Identifying Perspectives at the Document and Sentence Levels. In Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X), pages 109–116.
10. Mohit Iyyer, Peter Enns, Jordan Boyd-Graber and Philip Resnik. 2014. Political Ideology Detection Using Recursive Neural Networks. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1113–1122.
11. S. Raza, D.J. Reji and C. Ding. 2022. Dbias: detecting biases and ensuring fairness in news articles. In Int J Data Sci Anal 17, pages 39–59.

12. Ramy Baly, Giovanni Da San Martino, James Glass and Preslav Nakov. 2020. We Can Detect Your Bias: Predicting the Political Ideology of News Articles. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 4982–4991.
13. S. Siami-Namini, N. Tavakoli and A. S. Namin. 2019. The Performance of LSTM and BiLSTM in Forecasting Time Series. In 2019 IEEE International Conference on Big Data (Big Data), pages 3285-3292.
14. Gang Liu and Jiabao Guo. 2019. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. In Neurocomputing, volume 337, pages 325-338.
15. Kamaljit Kaur and Parminder Kaur. 2023. BERT-CNN: Improving BERT for Requirements Classification using CNN. In Procedia Computer Science, volume 218, pages 2604-2611.
16. Konstantina Lazaridou and Ralf Krestel. 2016. Identifying Political Bias in News Articles. In Bull. IEEE Tech. Comm. Digit. Libr., volume 12, n. pag.
17. Jerry Wei. 2020. NewB: 200,000+ Sentences for Political Bias Detection. In ArXiv abs/2006.03051, n. pag.