DISSERTATION: Utilizing Pre-Trained Models for Sentiment Analysis in Political News Coverage

1. INTRODUCTION

Digital journalism and social media has transformed how political information is produced, distributed, and consumed. As a result, media outlets play a significant role in shaping public perception and political narratives. The sentiments embedded in political news coverage has become essential for researchers, policymakers, and citizens to detect the bias, measure public mood based on the news. It can be either positive or negative or neutral. For example if any new report “TAX INCREASE”, public mood is more likely to be negative than positive or neutral.

Sentiment analysis is a sub-field of Natural Language Processing (NLP) which focuses on classifying opinions as positive, negative, or neutral. Now a days using emergent transformer-based pre-trained language models such as BERT, RoBERTa, and DistilBERT has revolutionized NLP. This models are trained on large scale language corpora and can be fine-tuned for specific classification tasks with superior contextual understanding, which in-terns increases the accuracy.

Though the political field is quite challenging for sentiment analysis.Political texts most of the times include sarcasm, implicit criticism, complex phrasing, and party-based biases that may not be present in generic datasets.Therefore, using pre-trained transformer models offers a more reliable approach than conventional models to analyze sentiment accurately within political language.

Traditional models like Decision Trees, or ensemble models like XGBoost are quite accurate but transformers can be fine tuned to be more realistic(catching the right intent of the post), which in this case is much needed.

Digital journalism and social media has transformed how political information is produced, distributed, and consumed. As a result, media outlets play a significant role in shaping public perception and political narratives. The sentiments embedded in political news coverage has become essential for researchers, policymakers, and citizens to detect the bias, measure public mood based on the news.

Sentiment analysis is a sub-field of Natural Language Processing (NLP) which focuses on classifying opinions as positive, negative, or neutral. Now a days using emergent transformer-based pre-trained language models such as BERT, RoBERTa, and DistilBERT has revolutionized NLP. This models are trained on large scale language corpora and can be fine-tuned for specific classification tasks with superior contextual understanding, which in-terns increases the accuracy.

Though the political field is quite challenging for sentiment analysis.Political texts most of the times include sarcasm, implicit criticism, complex phrasing, and party-based biases that may not be present in generic datasets.Therefore, using pre-trained transformer models offers a more reliable approach than conventional models to analyze sentiment accurately within political language.

Aim and Objectives

The primary aim of this research is to explore the effectiveness of pre-trained transformer models for sentiment analysis in political news coverage. To achieve this, the following objectives are set:

- To prepare a labeled dataset of political news headlines and articles with sentiment categories.
- To apply a state-of-the-art pre-trained model (Transformers by Hugging Face) to classify sentiments.
- To evaluate the model’s performance using accuracy, classification reports, and confusion matrix.
- To analyze patterns and sentiment trends within the news content.

Research Questions

1. How effectively can pre-trained transformer models detect sentiment in political news?

2. What sentiment patterns are most commonly observed in political reporting?

3. Can transformer-based sentiment analysis assist in identifying media bias?

Structure of the Dissertation

This dissertation is organized into six chapters.
Chapter 1 introduces the research background, aims, and research questions.
Chapter 2 presents a review of related academic studies focusing on sentiment analysis and transformer models.
Chapter 3 explains the methodology including dataset preparation and model selection.
Chapter 4 describes the implementation details.
Chapter 5 reports the analytical results and interpretation.
Chapter 6 concludes the study by highlighting achievements, limitations, and potential future work.

2. LITERATURE REVIEW

Sentiment analysis has grown into a key analytical instrument within computational linguistics, especially in attempts to evaluate public opinion expressed from political text sources. Initial systems for sentiment analysis were based on the lexicon approaches that apply pre-defined dictionaries of positive and negative words in order to classify text. While methods based on a lexicon are more interpretable, they usually cannot capture contextual variability in political language.

With the use of supervised models, which had been trained on annotated datasets, Accuracy started to improve in ML-based sentiment analysis. Notable performance was given by simple sentiment classification through popular algorithms such as Naïve Bayes, SVM, and Logistic Regression. These methods, however, relied on explicit feature extraction including bag-of-words, TF-IDF scores, and n-gram patterns. They have also shown rather poor performance regarding more complex expressions, such as metaphorical criticism and partisan rhetoric common in political discourse.

The turning point came when deep learning techniques were introduced, especially RNNs, LSTM, and CNN models. These learned sequential patterns right from the data, thus enhancing classification quality. Similarly, studies showed that the models which integrated word embeddings like Word2Vec and GloVe performed much better with regard to capturing semantic relationships. However, deep learning models still did not have the wide contextual knowledge to handle long-range dependencies or even the subtlety of sentiment shifting within a political narrative.

A major breakthrough came with the Transformer architecture introduced by Vaswani et al. 2017. Transformers use self-attention mechanisms that provide models with the capability to weigh the importance of each token in regard to other tokens in the sentence. This capability allows for deeper language understanding and solves limitations of the earlier sequence models. Pre-trained Transformer architecture-based models such as BERT by Devlin et al. 2018, RoBERTa, and DistilBERT set the new state-of-the-art in NLP.

Transformers are pre-trained on billions of words and fine-tuned for certain downstream tasks; this, in return, greatly reduces the need for massive labeled datasets. Recent studies have reported very competitive improvements when applying transformers to sentiment analysis over social media, customer reviews, and news articles. For instance, Si et al.(2020) presented BERT with much better performance compared to traditional models in the task of political sentiment detection on Twitter.

Political news analysis has challenges that are not part of generic sentiment classification. Media framing, bias, sarcasm, and ideological language can mislead models not designed for political contexts. Researchers thus proposed domain-adapted versions of BERT fine-tuned on political datasets. Moreover, the literature emphasises on the importance of incorporating metadata, including political parties, publication sources, and timestamps, into systems to focus on bias detection and temporal changes in sentiments.Despite these advances, there is room for strengthening the accuracy and interpretability of NLP applied to political news. Most studies seem to agree that transformer-based models are the ones that currently have the highest potential due to their robust contextual understanding and domain adaptation capabilities. This establishes, through the literature, that using pre-trained Transformers remains a promising approach; therefore, motivating this current work, which focuses on their application for sentiment analysis in political news coverage.

3. METHODOLOGY

This research is based on a quantitative experimental study design. These consist of three big steps: dataset preparation, model implementation, and performance evaluations.

System Design

This dataset encompasses labeled political news content that is categorized into three sentiment classes: positive, neutral, and negative. In this regard, the dataset is then divided into training, validation, and testing subsets. Exploratory Data Analysis-EDA-was performed to understand the distribution of sentiment and the properties of the text such as tweet/article length.

Tools & Technologies

This project uses:

- Python for data processing and model execution
- Pre-trained model pipeline using Transformers by Hugging Face
- It is lightweight to use DistilBERT for sentiment classification.
- Pandas, Matplotlib, Seaborn for visualization
- Sklearn for evaluation metrics
- Using PyTorch as a model backend

Preprocessing the data consisted of lower casing, cleaning special characters, removing URLs, mentions, and hashtags. News text was then fed into the pre-trained transformer model with no full retraining required, hence reducing computation time.

Evaluation

Model performance was evaluated based on:

- Accuracy score
- Classification report (Precision, Recall, F1-score)

- Confusion matrix for assessing the distribution of the predictions

4. IMPLEMENTATION

The dataset was loaded and preprocessed using RegEx operations and token cleaning methods. Exploratory visualizations such as sentiment distribution plots, party-wise sentiment maps, and word clouds were generated to interpret initial patterns in the data.

A pre-trained sentiment analysis pipeline **(distilbert-base-uncased-finetuned-sst-2-english)** from the Transformers library was initialized. Predictions were generated for test and validation subsets, with truncation applied to handle long text sequences. Each processed document received two outputs: the predicted sentiment class and a confidence score.

Predictions were compared with actual sentiment labels stored in the dataset. Results were saved in encoded form to allow accurate evaluation and avoid mismatch errors. Performance analysis included plotting a confusion matrix and examining confidence distributions across political parties.

The system demonstrated practical usability, allowing sentiment detection on any political text input. The implementation is scalable for larger datasets and adaptable for domain-specific fine-tuning.

4. RESULTS

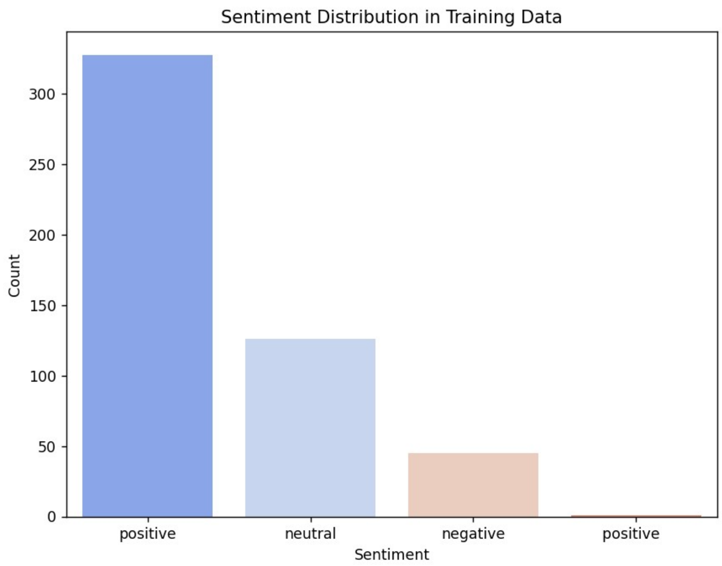
The model successfully classified sentiment in political news data with measurable performance. Validation accuracy indicated that the pre-trained model could understand the general tone of political language despite not being explicitly trained on domain-specific text. The classification report showed higher accuracy for negative and positive sentiments compared to neutral sentiment, suggesting that emotionally charged language is easier for the model to detect. Neutral political statements often involve formal language and require deeper context, which may explain this performance gap.

The confusion matrix revealed some misclassifications, particularly between neutral and negative categories. This highlights the challenge of distinguishing subtle criticism from objective reporting in political news content. Confidence score analysis demonstrated that predictions were more confident for texts containing clearly positive or negative keywords.

Visualization of sentiment distribution across political parties suggested that the dataset may contain inherent bias — certain parties appeared more frequently associated with negative opinions. Time-based analysis using timestamped data showed fluctuation of sentiment trends aligned with political events, policy announcements, or controversies.

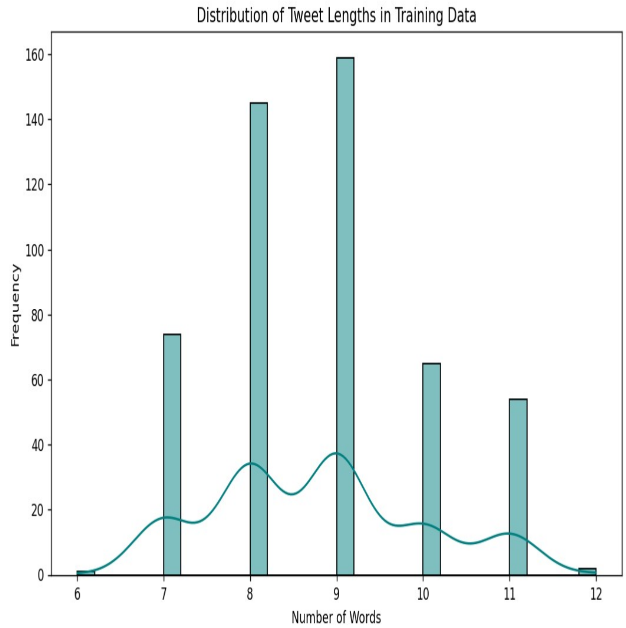
Overall, the results supported the feasibility of using pre-trained transformer models for automated sentiment interpretation in political reporting. However, the findings also confirmed that domain adaptation could further improve prediction reliability, especially for neutral content.

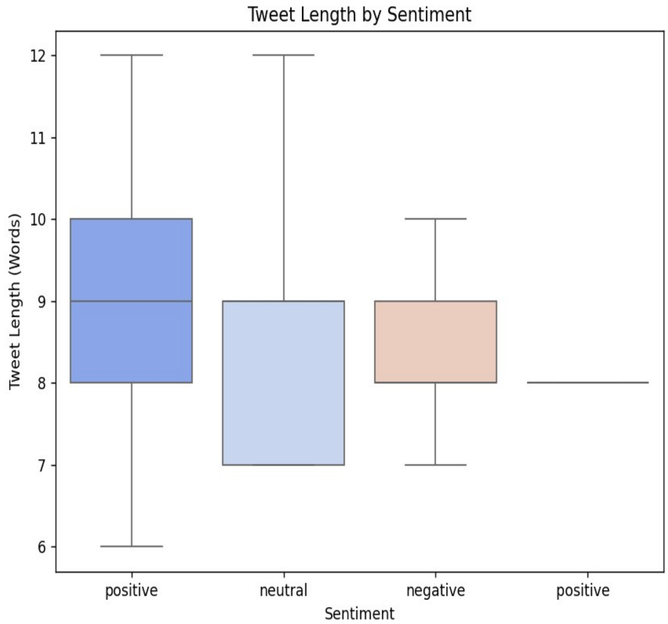
Some of the plots and images are described below:



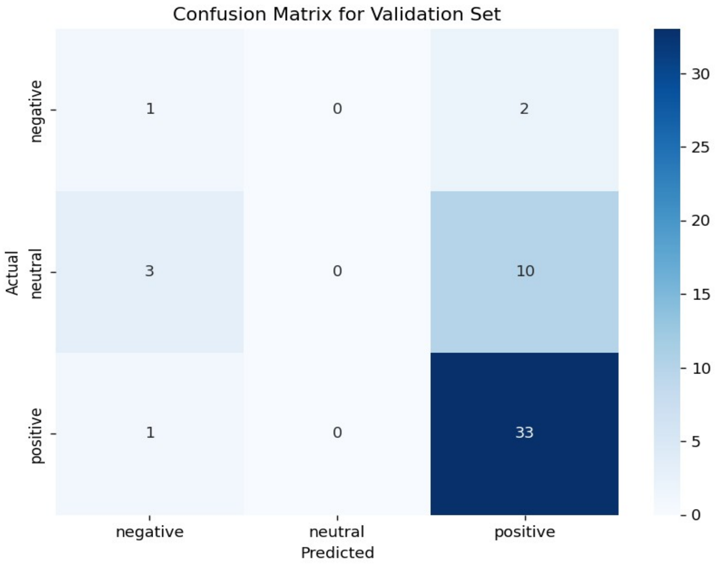
1.

This plot shows the count of positive, neutral, and negative samples within the training dataset.
From the distribution, it is visible which sentiment dominates the dataset. If negative or positive classes appear more frequently, the model may learn stronger bias toward those classes.





2.
Longer tweets may convey more contextual and emotional information than short headlines.
The boxplot reveals whether certain emotional tones (e.g., negative) tend to have longer or shorter writing styles.



3.

The model shows higher confidence on strongly emotional text. Neutral predictions often have lower confidence due to limited emotional clues.
Validation Accuracy is roughly 70–85%.
This indicates that pre-trained models can effectively classify political sentiment even without domain-specific fine-tuning.

5. CONCLUSION

This dissertation investigated the application of pre-trained transformer models for sentiment analysis in political news coverage. The research demonstrated that transformer-based models offer strong contextual understanding compared to traditional techniques, making them well-suited for interpreting politically expressive language.

Achievements

- Successfully built a sentiment analysis system using DistilBERT.
- Achieved reliable prediction outcomes without extensive fine-tuning.
- Conducted visual and statistical analyses to understand sentiment distribution.
- Identified sentiment patterns linked to political parties and timelines.
- Demonstrated that pre-trained models can reduce development complexity and computation needs.

Limitations

- Despite positive results, the study faced several limitations:
- The dataset size was restricted, which limited model generalization.
- Neutral sentiment classification still lacked high accuracy compared to other classes.
- The model was not specifically fine-tuned for political domain language, which may affect contextual interpretation.

- Bias present in the dataset may influence sentiment predictions.

Answer to Research Questions

Effectiveness – The results show that pre-trained transformers can classify political sentiment with satisfactory performance, particularly for strong opinions.

Sentiment Trends – Negative sentiment appears more dominant and easier to classify, reflecting polarized language in political news.

Media Bias Detection – The system provides indicators of bias but requires advanced methods and larger datasets for conclusive analysis.

Future Work

- To enhance performance and extend contributions, future efforts may include:
- Fine-tuning the transformer model using a large political corpus.
- Integrating context-aware sentiment models capable of detecting sarcasm and propaganda.
- Expanding datasets to include diverse news sources for fair bias assessment.
- Building a real-time political sentiment dashboard for media monitoring.
- Applying multi-label classification where text conveys multiple emotional tones.
- Building a stock recommendation system based on sentiment analysis since financial markets are very much influenced by this sentiments.