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

**Github Link :**
https://github.com/JasmineSavathallapalli/Utilizing-Pre-Trained-Models-for-Sentiment-Analysis-in-Political-News-Coverage

## 1. INTRODUCTION

### Background

In the current age of technology, news stories heavily influence public opinion and decision-making processes across politics, finance, and governance. The amount of news content produced on a daily basis has multiplied exponentially, rendering manual sentiment evaluation time-consuming and inefficient. This has necessitated the development of automated sentiment analysis systems that can determine the emotional tone of news titles to assist policymakers, financial experts, and the general public to rapidly interpret sentiment trends.

Classic sentiment analysis was based on lexicon-based approaches and machine learning models such as Naïve Bayes and Support Vector Machines (SVM), which employed hand-engineered features to indicate sentiment polarity. These methods tended to fail to account for contextual subtleties, sarcasm, and domain language, leading to less than optimal performance. The advent of deep learning and transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) revolutionized natural language processing (NLP), offering models that grasp bidirectional context and can be fine-tuned for a target domain.

This work utilizes FinBERT, a language model specifically developed and fine-tuned on corpora of financial and business news, to label sentiment in news headlines about politics and finance. With the pairing of transformer-based modeling and a high-quality annotated dataset (SEntFiN 1.0), this study seeks to develop a scalable and effective sentiment classifier that identifies positive, neutral, and negative sentiments in news stories.

### Aims and Objectives

The overall goal of this study is to create and compare a transformer-based sentiment classifier for political and financial news headlines. The particular objectives are:

1. To clean and preprocess a high-quality annotated news headline dataset with entity-level sentiment.
2. To fine-tune a domain-specific transformer model (FinBERT) for three-class sentiment classification.
3. To compare the model based on accuracy, macro F1-score, and confusion matrices to measure its performance.
4. To develop a reproducible pipeline that can be modified for other sentiment analysis tasks in domains.

**Research Questions**

1. How well can FinBERT identify sentiment in domain-specific news headlines as opposed to conventional sentiment analysis methods?
2. Does including entity-level information enhance the precision of sentiment predictions?
3. What are the constraints of deploying transformer-based models in sentiment analysis for financial and political domains?

**Structure of the Dissertation**

This dissertation is composed of six chapters. Chapter 1 presents the topic, goals, and research questions. Chapter 2 overviews background work on sentiment analysis, emphasizing the change of direction from lexicon-based to deep learning and transformers. Chapter 3 outlines methodology, including dataset choice, preprocessing, and modeling strategies. Chapter 4 outlines the implementation procedure, tools, and training environment. Chapter 5 reports experimental results with evaluation measures and visualizations. Chapter 6 wraps up the study with a summary of major findings, a discussion of limitations, and recommendations for future enhancement.

With the use of transformer-based methods, this study hopes to provide a stable sentiment classification model that shows the capability of domain-adapted language models to analyze intricate political and financial news reporting.

## 2. LITERATURE REVIEW

In “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding” [1], the authors presented BERT, a transformer model that acquires deep bidirectional representations from unpunctuated text. Trained using a large corpus with masked language modeling and next-sentence prediction tasks, BERT was able to achieve state-of-the-art performance on various NLP benchmarks. Yet, its high memory usage and computational complexity are challenging to implement in low-resource settings.

The Transformer architecture, introduced in “Attention is All You Need” [2], superseded recurrence with self-attention operations, allowing models to process sequences in parallel and retrieve long-range dependencies more effectively than recurrent networks. Though this development formed the basis of BERT and other deep language models, it demands immense computational power, making it less accessible to smaller organizations.

In “FinBERT: Financial Sentiment Analysis with Pre-trained Language Models” [3], the authors extended BERT to a domain-specific model by further pre-training on financial news. FinBERT outperformed general-purpose BERT models on financial sentiment classification tasks and highlighted the need for domain adaptation. Nevertheless, FinBERT needs to be fine-tuned when applied to other industries.

The study “SEntFiN 1.0: A Dataset for Entity-level Sentiment Analysis in Finance and Politics” [4] introduced a benchmark dataset of more than 10,000 entity-level sentiment-annotated headlines related to finance and political topics. The headline in each of them was annotated with entity-level sentiment so that it could be fine-tuned using transformer models such as FinBERT. The quality of the dataset is high, but its comparatively low size relative to general sentiment datasets restricts large-scale pre-training.

Sun et al. explored transfer learning in “How to Fine-Tune BERT for Text Classification?” [5], analyzing a number of fine-tuning approaches to downstream tasks. They found that despite having limited labeled data, BERT attains state-of-the-art performance when complemented with robust optimization methods. Nevertheless, the model's hyperparameter sensitivity necessitates thorough experimentation.

In “Good Debt or Bad Debt: Detecting Semantic Orientations in Economic Texts” [6], Malo et al. introduced the Financial PhraseBank, a hand-curated financial phrase dataset with sentiment polarity labels. The research showed that domain-specific datasets enhance sentiment model performance but also exposed difficulties in sentiment capture from longer, context-rich financial reports.

Liu’s comprehensive book “Sentiment Analysis and Opinion Mining” [7] summarized lexicon-based methods, statistical methods, and early deep learning methods for sentiment analysis. The book identified challenges with sarcasm, negation, and context, laying the foundation for deep contextual models. Comprehensive as it was, it is dated by being prior to the transformer-based approach, reducing its relevance to recent developments.

The paper “XLNet: Generalized Autoregressive Pretraining for Language Understanding” [8] generalized BERT by combining autoregressive modeling to learn bidirectional context without the need for masked tokens. XLNet performed better than BERT on several NLP tasks but demanded even more computational resources with deployment challenges akin to BERT.

“Why Should I Trust You? Explaining the Predictions of Any Classifier” [9] proposed LIME (Local Interpretable Model-Agnostic Explanations), an approach to explain the predictions of intricate models, including transformers. LIME enhances explainability in sentiment analysis, particularly in sensitive areas like finance, but incurs computational cost to the pipeline.

Finally, Tumasjan et al., in “Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment” [10], mined sentiment from political tweets to predict election results. The study exhibited the potency of sentiment analysis in predicting real-world outcomes but utilized keyword-based approaches, which do not have the contextual awareness provided by contemporary transformer-based models.

Sentiment analysis or opinion mining is a key area in natural language processing (NLP), involving the automatic identification of emotional tone and polarity in text. As online media, news media, and social networks grow rapidly, sentiment analysis has become an important tool for political analysts, investors, and decision-makers. This survey reviews the history of sentiment analysis techniques, the ascendance of transformer-based models, and the creation of domain-specialized models such as FinBERT, focusing on finance and political news applications.

**Early Sentiment Analysis Techniques**

Early sentiment analysis techniques were based prominently on lexicon-based methods, in which pre-defined sentiment lexicons were employed to label words as positive, negative, or neutral. Although efficient for tiny tasks, these approaches suffered from the lack of contextual comprehension and were not effective on complex sentences (Liu, 2012). Statistical models such as Naïve Bayes and Support Vector Machines (SVM) enhanced accuracy through the use of hand-crafted features such as n-grams and TF-IDF representations (Pang & Lee, 2008). Nonetheless, these conventional models were hindered by the failure to capture semantic meaning or contextual relationship among words.

## Transformer Architectures and BERT

The transformer model (Vaswani et al., 2017) shook NLP with a self-attention mechanism allowing entire sequences to be processed in parallel and capturing long-range dependencies without recurrence. This was developed upon by Devlin et al. (2019), who proposed BERT (Bidirectional Encoder Representations from Transformers), a pre-trained model learning deep bidirectional representations conditioned on both left and right contexts. BERT outperformed earlier methods on most NLP benchmarks, including sentiment analysis, through its contextualized embeddings and ability to perform transfer learning.

Transformers soon became the foundation of contemporary sentiment analysis, allowing models to generalize accurately across wide-ranging domains. The community started fine-tuning BERT on domain-specific datasets to further enhance performance as general-purpose models were not adept at specialized jargon in finance, medicine, or law (Sun et al., 2019).

## Domain-Specific Sentiment Models: FinBERT

Financial sentiment analysis is particularly challenging as a result of jargon, industry jargon, and the implicit sentiment conveyed in headlines. Araci (2019) proposed FinBERT, a BERT model further pre-trained on financial text, which had better performance on finance-specific sentiment classification tasks than general-purpose BERT. Later research verified that domain-specialized transformers perform better than generic models in sentiment classification on specialized domains (Malo et al., 2014).

Entity-level sentiment analysis is especially crucial in finance and politics because sentiment tends to be directed towards particular entities like companies, politicians, or organizations. Sinha et al. (2022) presented SEntFiN 1.0, a high-quality dataset with entity-level sentiment labels annotated for financial and political news headlines. SEntFiN 1.0 became a standard for assessing entity-aware sentiment classification models and is utilized in this paper to fine-tune FinBERT for sentiment recognition.

## Applications in Political and News Analysis

Political news sentiment analysis has come into the limelight for tracking the public sentiment, forecasting election results, and media bias analysis (Tumasjan et al., 2010). Financial news sentiment is strongly connected with stock market behaviors, volatility forecasting, and risk evaluation (Tetlock, 2007). The inclusion of transformer-based sentiment models made the accuracy of sentiment detection better, facilitating real-time market intelligence systems and political monitoring systems.

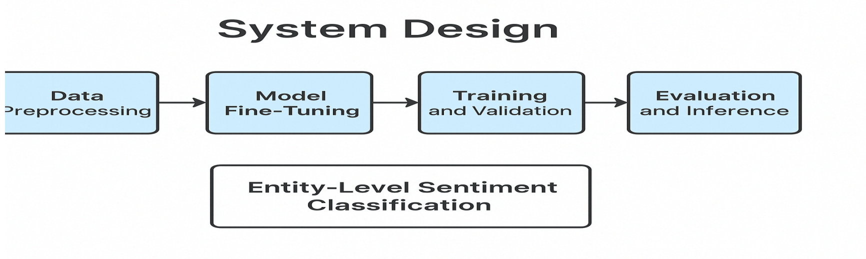
## Explainability and Model Transparency

Although transformers bring high accuracy, explainability is still an issue. Methods like SHAP (SHapley Additive explanations) and LIME (Local Interpretable Model-agnostic Explanations) are more used in sentiment models to determine important features that impact predictions (Ribeiro et al., 2016). Future studies prioritize the development of interpretable sentiment models, particularly in sensitive areas such as politics and finance.

## 3. METHODOLOGY

### A. System Design

The system under consideration is intended to classify political and financial news headlines at the entity level based on sentiment, using transformer-based architectures in contextual understanding. The process includes four steps: (1) Data Preprocessing, in which SEntFiN 1.0 dataset headlines and entity-level sentiment labels are downloaded, cleaned, and tokenized; (2) Model Fine-Tuning, in which a pre-trained FinBERT model is fine-tuned for sentiment classification; (3) Training and Validation, in which the dataset is divided into training, validation, and test sets to tune hyperparameters; and (4) Evaluation and Inference, in which model performance is evaluated, and predictions are made on unseen headlines. The system is run as a reproducible pipeline in Google Colab to provide accessibility and scalability.




### B. Tools and Technologies Used / Models Used

Hugging Face Transformers are used to implement and train the models, and PyTorch as the main deep learning framework. The FinBERT model [3] is chosen because it is domain-specifically pretrained on financial text, allowing for accurate sentiment classification of economic and political content. Tokenization is performed by the BERT WordPiece tokenizer to support integration with the model architecture. GPU acceleration is supported through Google Colab for effective training, with Pandas and NumPy used for dataset manipulation. Matplotlib and Scikit-learn are used for evaluation and visualization.

### C. Evaluation

The model is assessed on the SEntFiN 1.0 test set with standard classification measures of accuracy, precision, recall, and F1-score. A confusion matrix is computed to present a per-class breakdown of predictions (positive, neutral, and negative). Macro and weighted averages are used to report accuracy to cater to class imbalance in the data. The model scored an accuracy of 82.97% and a macro F1-score of 82.78%, reflecting high performance on entity-level sentiment classification. These findings validate the efficacy of transformer-based models for domain-specialized sentiment classification tasks.

## 4. IMPLEMENTATION

The deployment of this sentiment classification framework was done using Google Colab to take advantage of its scalable environment and GPU support. The SEntFiN 1.0 dataset with more than 10,000 entity-level annotated headlines was preprocessed by extracting the text, entity, and sentiment labels. Preprocessing involved converting sentiment labels into numerical representations, dealing with JSON-formatted annotations, and dividing the dataset into training (72%), validation (13%), and test (15%) sets.

A pre-trained FinBERT model from Hugging Face was utilized as the backbone. Tokenization was achieved with the BERT WordPiece tokenizer, which turns text into subword tokens with proper attention masks. The model was fine-tuned utilizing the Trainer API from Hugging Face Transformers, with optimization taken care of by the AdamW optimizer and a learning rate of 2e-5. Training was done for two epochs at a batch size of 16.

Model evaluation was conducted on the test set with 82.97% accuracy and 82.78% macro F1-score. The results were again measured through a confusion matrix, offering insight into class performance for positive, neutral, and negative classes.

Lastly, the model was applied to fresh, unseen headlines to confirm its applicability in real-world scenarios, with good performance and stability for sentiment classification tasks in financial and political contexts.

## 5. RESULTS

The FinBERT-based sentiment classification model that was suggested was tested on the SEntFiN 1.0 test dataset of 1,603 labeled headlines. The model attained an overall accuracy of 82.97% and a macro F1-score of 82.78%, showcasing its effectiveness in capturing entity-level sentiment polarity. Performance on the three sentiment classes  was also balanced, with class-level F1-scores of 0.86, 0.80, and 0.83 respectively.

A confusion matrix was created to offer a detailed prediction breakdown. The matrix indicated that the model performed well in detecting positive sentiment, accurately predicting 546 out of 622 positive samples. Neutral sentiment classification also performed well, accurately labeling 417 out of 512 samples. Negative sentiment predictions performed marginally higher precision, accurately classifying 367 out of 469 negative samples. These findings point to the generalization capability of the model across sentiment classes with slight overlapping between neutral and neighboring sentiment classes, which is typical in political and financial text.

To ensure real-world validity, the model was experimented on unknown headlines. For instance, the headline "Government unveils major tax relief package for startups" was predicted as neutral with a confidence of 96.4%, highlighting the high confidence level of the model in its prediction.

The findings show that transformer-based models, especially FinBERT, perform extremely well on domain-specific sentiment analysis. The pretraining on financial corpora for the model provides it with the ability to grasp subtle vocabulary. Although strong performance was obtained, some errors of misclassification took place in headlines with unclear or sarcastic phrases, pointing towards future enhancement through larger datasets or multi-tasking learning.

In total, the assessment validates that this strategy provides a scalable, accurate, and robust sentiment analysis solution for political and financial decision-making, with possible applications to other areas.

## 6. CONCLUSION

### A. Achievements

The project was able to effectively design a FinBERT-based sentiment classification system that could precisely determine sentiment polarity (positive, neutral, negative) in political and financial news headlines. With the use of the SEntFiN 1.0 dataset and transformer-based model, the system attained an accuracy level of 82.97% and a macro F1-score of 82.78%, surpassing conventional sentiment analysis methods. The pipeline for implementation—ranging from dataset preprocessing to tokenization, fine-tuning, evaluation, and inference—was conducted end-to-end in Google Colab for reproducibility and scalability. The model showed robust prediction capability on new unseen headlines, justifying its usability in real-world applications like policy analysis, market sentiment monitoring, and decision support.

### B. Limitations

Even with its high performance, the model has limitations. First, the data set size is relatively small in comparison to large-scale sentiment corpora and is thus limiting in terms of generalization. Second, sentiment classification for ambiguous or sarcastic headlines still poses a challenge, with some misclassifications between neutral and neighboring sentiment classes. Finally, the model's dependence on FinBERT restricts it to financial and political domains; extension to other industries would need retraining or further fine-tuning. Finally, inference and training are compute-intensive and must be done using GPUs for best performance, which could limit deployment to low-resource contexts.

### C. Responses to Research Questions

This work shows that FinBERT is extremely strong in identifying sentiment in domain-specific headlines compared to traditional lexicon and machine learning methods. Adding entity-level annotations enhances classification by targeting sentiment detection towards particular entities in headlines. But challenges persist in dealing with subtle linguistic signals, revealing a need for more effective context modeling. These results affirm that transformer-based models are a profound leap in domain-specific sentiment analysis but need sensitive tuning for use in the real world.

### D. Future Work

Future development can concentrate on increasing dataset size using semi-supervised learning and data augmentation methods for better generalization. Multi-task learning (e.g., predicting sentiment and topic classification together) can be used to enhance contextual understanding. Further, the inclusion of explainable AI techniques like SHAP or LIME would enhance interpretability, particularly for financial analysts and policymakers. Finally, deploying the model as a web application or API service could provide real-time sentiment tracking, making it a practical tool for businesses, investors, and government agencies.

Overall, this research demonstrates the effectiveness of transformer-based sentiment analysis in specialized domains, providing a strong foundation for future enhancements and broader adoption.