Bulletin of Electrical Engineering and Informatics
Vol. 14, No. 2, April 2025, pp. 1231~1240
ISSN: 2302-9285, DOI: 10.11591/eei.v14i2.8469

The impact of BERT-infused deep learning models on sentiment analysis accuracy in financial news

Oussama Ndama, Ismail Bensassi, El Mokhtar En-Naimi
DSAI2S Research Team, Computer Science and Smart Systems Laboratory, FST of Tangier, Abdelmalek Essaâdi University, Tetouan, Morocco

Article Info

Article history:

Received Mar 15, 2024
Revised Sep 29, 2024
Accepted Oct 17, 2024

Keywords:

Bidirectional encoder representations from transformers
Deep learning models
Financial news
Natural language processing
Sentiment analysis

ABSTRACT

This study delves into the enhancement of sentiment analysis accuracy within the financial news domain through the integration of bidirectional encoder representations from transformers (BERT) with traditional deep learning models, including artificial neural networks (ANN), long short-term memory (LSTM) networks, gated recurrent units (GRU), and convolutional neural networks (CNN). By employing a comprehensive method encompassing data preprocessing, polarity analysis, and the application of advanced neural network architectures, we investigate the impact of incorporating BERT's contextual embeddings on the models' sentiment classification performance. The findings reveal significant improvements in model accuracy, precision, recall, and F1 scores when BERT is integrated, surpassing both traditional sentiment analysis models and contemporary natural language processing (NLP) transformers. This research contributes to the body of knowledge in financial sentiment analysis by demonstrating the potential of combining deep learning and NLP technologies to achieve a more nuanced understanding of financial news sentiment. The study's insights advocate for a shift towards sophisticated, context-aware models, highlighting the pivotal role of transformer-based techniques in advancing the field.

Corresponding Author:

Oussama Ndama
DSAI2S Research Team, Computer Science and Smart Systems Laboratory, FST of Tangier
Abdelmalek Essaâdi University
Tetouan, Morocco
Email: oussama.ndama@etu.uae.ac.ma

1. INTRODUCTION

In the age of information overload, the ability to accurately gauge sentiment from financial news has become paramount for investors, analysts, and policymakers. The rapid evolution of online media has not only democratized access to financial information but also introduced challenges in sifting through vast amounts of data to discern impactful insights. Within this context, sentiment analysis emerges as a critical tool, enabling the extraction of subjective information from textual data, thereby offering a nuanced understanding of market dynamics and investor sentiment. Traditional approaches, while useful, often fall short in capturing the subtleties and complexities of language used in financial news, underscoring the need for more advanced techniques.

The advent of deep learning and natural language processing (NLP) technologies, particularly Transformer-based models like bidirectional encoder representations from transformers (BERT), has marked a significant leap forward in sentiment analysis [1], [2]. These models, with their unprecedented ability to understand context and nuance in text, present a novel avenue for enhancing sentiment analysis accuracy.

This study aims to explore the potential of BERT-infused deep learning models namely, artificial neural networks (ANN), long short-term memory (LSTM) networks, gated recurrent units (GRU), and convolutional neural networks (CNN) in improving the classification of sentiment in financial news articles. By integrating the deep contextual insights provided by BERT with the specialized capabilities of these neural network architectures, we propose a sophisticated model approach designed to achieve a deeper and more accurate analysis of sentiment within the financial news domain.

Our method adopts a comprehensive approach, beginning with meticulous data preprocessing and polarity analysis. This foundation supports the construction and evaluation of both standalone deep learning models and their hybrid configurations with BERT. Our dual-pathway exploration enables the assessment of each model's individual performance in sentiment classification and investigates the synergistic potential of integrating BERT with traditional neural network architectures. The significance of this research extends to its contribution towards refining sentiment analysis techniques within the realm of financial news. By leveraging cutting-edge NLP technologies, this study aims to offer valuable insights into the dynamics of financial markets, thus enhancing decision-making processes and providing a robust analytical tool for financial analysis in the digital era. Through this comprehensive exploration, we seek to establish new benchmarks for accuracy and depth in financial sentiment analysis, highlighting the transformative impact of merging deep learning models with transformer-based NLP technologies.

Following this introduction, the paper unfolds through a structured exploration of our research journey. We commence with a literature review, situating our work within the broader discourse on financial sentiment analysis and highlighting the novel contributions of deep learning and NLP technologies. The research method section details our comprehensive approach to data preparation and model implementation, paving the way for the results and discussion segment, where we critically examine our findings and their implications for the field. The paper culminates in a conclusion that synthesizes our insights, offering reflections on the future of sentiment analysis in financial news and potential directions for subsequent research. This cohesive structure ensures a thorough examination of the methodologies, findings, and contributions of our study, advancing the dialogue on sentiment analysis innovations.

2. LITERATURE REVIEW

In the rapidly evolving field of financial sentiment analysis, recent studies have significantly contributed to understanding and advancing methodologies ranging from lexicon-based approaches to sophisticated machine learning models and cutting-edge NLP transformers. Mishev *et al.* [3] initiate this exploration by assessing various sentiment analysis models within the financial domain, highlighting the challenges posed by specialized language and the scarcity of labeled datasets. Their comprehensive evaluation platform showcases the superior efficacy of contextual embeddings over traditional lexicons and fixed encoders, even in the absence of extensive datasets, and suggests the utility of distilled NLP transformers in production environments due to their comparable performance to larger models.

Building on the foundation of domain-specific challenges, Sinha *et al.* [4] introduce SEntFiN 1.0, a meticulously annotated dataset aimed at enhancing research in fine-grained financial sentiment analysis. Their work emphasizes the complexity of extracting sentiments associated with multiple entities within a single headline, demonstrating the effectiveness of lexicon-based n-gram ensembles and the notable performance of RoBERTa and finBERT, in achieving high accuracy and F1-scores. Further, Mishev *et al.* [5] delve into the performance evaluation of word and sentence embeddings for sentiment analysis in finance headlines, identifying the BiGRU+Attention architecture combined with word embedding features as the most effective approach. Their study underlines the transformative impact of embeddings and deep neural networks in improving sentiment analysis tasks within the financial sector.

In a pioneering study, Fatouros *et al.* [6] explore the potential of ChatGPT 3.5 in financial sentiment analysis, focusing on the foreign exchange market. Their innovative use of a zero-shot prompting approach with ChatGPT significantly outperforms traditional models like FinBERT in sentiment classification and shows a higher correlation with market returns, underscoring the importance of prompt engineering in enhancing model performance. Lastly, Daudert [7] introduces a novel approach that leverages both textual and contextual information for fine-grained financial sentiment analysis. By incorporating sentiment contagion and leveraging an adapted feed-forward neural network, their method demonstrates substantial improvements over baseline models, showcasing the importance of multi-text analysis in capturing implicit sentiments.

Tan *et al.* [8] and Yadav *et al.* [9] further contribute to the field by exploring rule-based and unsupervised techniques for financial news sentiment analysis. Tan *et al.* [8] rule-based algorithm employs a polarity lexicon and sentiment composition rules to classify financial news, achieving notable accuracy. Conversely, Yadav *et al.* [9] focus on tuning existing techniques for the financial domain, proposing alternative methods that significantly enhance sentiment strength identification from financial news. Collectively, these studies underscore a critical shift towards integrating advanced NLP techniques, such as transformers and large language models, with traditional sentiment analysis approaches. This body of work not only highlights the progress in financial sentiment analysis but also sets the stage for future research that bridges deep learning advancements with domain-specific challenges.

3. METHOD

The method section of our study delineates a structured approach designed to evaluate the efficacy of BERT-infused deep learning models in enhancing sentiment analysis accuracy within financial news. This comprehensive methodological framework encompasses several pivotal stages: initial data preprocessing to refine and standardize the dataset, followed by a detailed sentiment polarity analysis to categorize financial news articles based on their sentiment. Subsequently, the study delves into the construction and evaluation of sophisticated deep learning models, including ANN, LSTM, GRU, and CNN, tailored for sentiment classification. Parallelly, we explore an innovative approach by integrating the BERT model with these neural networks, aiming to leverage BERT's deep contextual insights to enrich the sentiment analysis process. This section is set to elaborate further on the technical execution and practical application of these models, underscoring the intricate method that forms the backbone of our foray into cutting-edge sentiment analysis practices within the financial news sphere. Figure 1 presents a visual overview of our study's method architecture, showcasing the flow from data preprocessing to sentiment polarity analysis, and the application of BERT-enhanced deep learning models. This schematic serves as a guide to our comprehensive approach, highlighting how each stage interconnects.

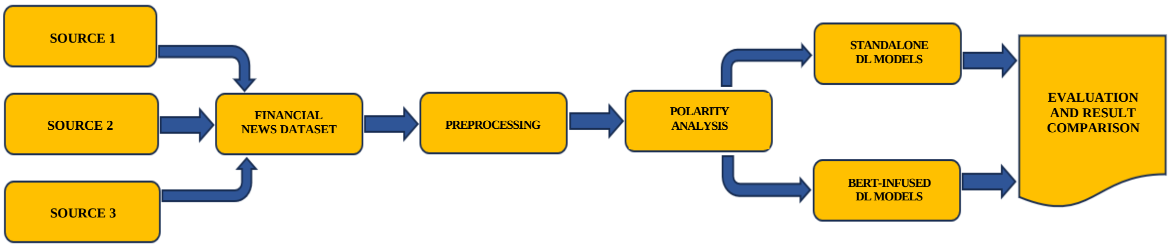

Figure 1. Proposed framework architecture for financial news sentiment analysis

3.1. Data description and preprocessing

Our analysis is based on a meticulously curated dataset from news articles by CNBC, the guardian, and reuters from 2017 to 2020, a period representing diverse economic conditions [10]. We initially harmonized and standardized the data from different sources to ensure consistency, followed by deduplication to maintain data quality. The dataset was then thoroughly cleaned to remove any duplicates and missing values, ensuring its integrity. Extensive preprocessing included converting text to lowercase, removing punctuation, filtering out stopwords, and lemmatization to refine the data for deeper analysis [11], [12]. These steps, executed using tools like the natural language toolkit (NLTK), demonstrate the precision of our dataset preparation, setting the foundation for our study of sentiment dynamics in financial news [13]. This careful curation and preparation of the data underscore the rigorous analytical framework of the study, setting the stage for an in-depth exploration of sentiment dynamics within financial news. To provide a concrete example of the outcomes of these preprocessing steps, Table 1 showcases an example of the preprocessed dataset.

Table 1. Example of the preprocessed dataset

Index	Info
0	Jim cramer better way invest covid19 vaccine gold rush
1	Toyota turn ai startup accelerate goal robot home
2	Trump torn uschina trade deal official push fulfill term
3	CEO tell trump hiring american without college degree
4	China February export seen falling two-year import reuters poll

3.2. Polarity analysis

Following the comprehensive preprocessing of the dataset, the study advances into the sentiment polarity analysis phase, a critical component aimed at deciphering the underlying sentiment in financial news. This phase employs the valence aware dictionary and sentiment reasoner (VADER) sentiment intensity analyzer from NLTK, a tool specifically designed for sentiment analysis in social media, news, and similar textual content [14], [15]. VADER is adept at computing sentiment scores by evaluating the emotional valence of textual data, making it particularly suitable for analyzing the nuances of financial news sentiment. Each news item in the dataset is assessed to generate a compound sentiment score, which ranges from -1 (indicating a strongly negative sentiment) to +1 (indicating a strongly positive sentiment). This scoring mechanism is pivotal for classifying each piece of news into one of three sentiment categories: positive, neutral, or negative. To visually represent the outcome of this sentiment analysis, Figure 2 illustrates the distribution of sentiment categories across the dataset.

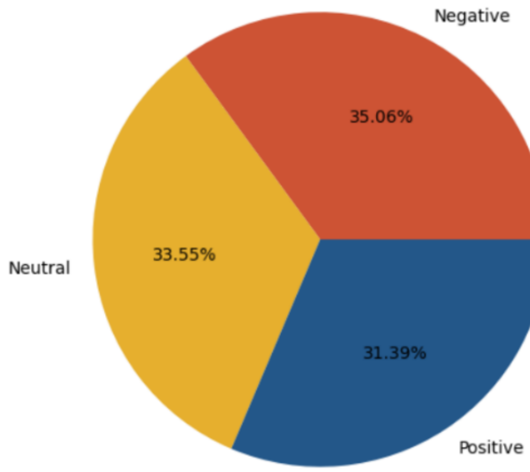

Figure 2. Sentiment categories distribution

The sentiment analysis process is meticulously crafted to ensure accuracy and depth. Initially, the ‘info’ column, which contains the preprocessed news text, is subjected to the sentiment analysis procedure. The sentiment intensity analyzer evaluates the text, assigning it a compound score based on the aggregated sentiment of the words within the text, effectively quantifying the sentiment expressed in each news article. This polarity analysis stage is crucial for understanding the sentiment dynamics within the dataset, laying the groundwork for subsequent investigations into the efficacy of BERT-infused deep learning models in enhancing sentiment analysis accuracy. By leveraging advanced sentiment analysis techniques, this phase of the study aims to uncover nuanced sentiment patterns, offering a robust foundation for exploring the transformative potential of BERT-enhanced models in financial sentiment analysis. To illustrate the outcomes of this polarity analysis, Table 2 presents examples from the dataset post-polarity analysis stage. This table offers a glimpse into the processed data, now enriched with sentiment scores, providing concrete examples of how the VADER tool categorizes each news item into distinct sentiment categories.

Table 2. Example of the dataset after polarity analysis stage

Index	Info	Polarity score	Sentiment
0	Jim cramer better way invest covid19 vaccine gold rush	0.50000	Positive
1	Toyota turn ai startup accelerate goal robot home	0.00000	Neutral
2	Trump torn uschina trade deal official push fulfill term	0.22630	Positive
3	CEO tell trump hiring american without college degree	0.00000	Neutral
4	China February export seen falling two-year import reuters poll	-0.15310	Negative

3.3. Neural networks

Building upon the foundational work of data preprocessing and polarity analysis, the study extends its exploration into the realm of advanced neural network architectures, namely ANNs, LSTM, GRU, and CNNs. These models stand at the forefront of the sentiment analysis phase, meticulously designed to delve deeper into the intricacies of sentiment classification within financial news. ANNs serve as the bedrock for understanding complex, non-linear relationships within the data, providing a versatile framework for capturing the subtleties of sentiment expression [16], [17]. LSTM and GRU networks, with their exceptional ability to process sequences of data, are particularly adept at analyzing the flow of narrative within news articles, capturing long-range dependencies that are crucial for understanding context and sentiment evolution over text [18], [19]. CNNs, renowned for their feature extraction capabilities, are leveraged to identify key phrases and patterns indicative of sentiment within the textual data, enriching the analysis with their spatial processing prowess [20], [21].

3.3.1. Artificial neural networks

In our study's exploration of neural network architectures for sentiment analysis, the ANN plays a pivotal role. Our model utilizes a sequential structure starting with an embedding layer that transforms text into dense vectors of 50 dimensions, considering a vocabulary up to max_words and sequence lengths max_sequence_length [22]. This layer effectively captures semantic relationships essential for processing text data. Following this, a flatten layer transitions the data for processing by two dense layers, with 128 and 64 units respectively, each activated by 'relu' to handle complex patterns and prevent overfitting with a dropout rate of 0.5 [23].

The architecture concludes with a dense output layer that categorizes sentiments into positive, neutral, and negative using a 'softmax' function, ensuring accurate classification based on probability [24]. This ANN architecture is meticulously designed to capture the nuances of sentiment in financial news, leveraging the power of deep learning to achieve nuanced sentiment classification. Through this model, the study seeks to uncover deeper insights into sentiment trends within the financial domain, setting a foundation for further exploration with more complex neural network architectures and BERT-enhanced models.

3.3.2. Long short-term memory

In our study, the LSTM network is pivotal for sentiment analysis in financial news, addressing the challenges of temporal dependencies in text data. The LSTM model starts with an embedding layer, transforming textual input into dense vectors of 50 dimensions, tailored for a maximum vocabulary size max_words and sequence length max_sequence_length. This setup captures the semantic essence of words, facilitating deeper analysis. At the core is the LSTM layer with 64 units, using 'tanh' and 'sigmoid' activations for handling main and recurrent transitions [25]. This enables the LSTM to retain and process information over extended sequences, essential for tracking evolving sentiment in news articles and understanding contextual nuances [26]. The model also includes a dense layer with 64 units activated by 'relu' to learn complex data relationships, and a Dropout layer with a 0.5 rate to prevent overfitting. It concludes with a dense output layer using a 'softmax' function, categorizing outputs into three sentiment categories, thus classifying news articles by their prevailing sentiment. This LSTM-based model represents a sophisticated approach in our sentiment analysis toolkit, offering enhanced capability to decipher complex sentiment dynamics within financial news. By leveraging LSTM's strength in processing sequential data, the model provides a foundation for in-depth analysis and interpretation of sentiment trends, contributing significantly to our understanding of financial market sentiments.

3.3.3. Gated recurrent units

The study enhances its sentiment analysis framework by integrating GRU models, noted for their efficiency in processing sequential data, particularly in textual sentiment analysis. Structured using a sequential framework, the GRU model begins with an embedding layer that transforms text into dense vectors with a capacity for 10,000 words and an output dimension of 50, preparing the model to capture the semantic nuances of financial news articles [27]. At its core, the GRU model features a GRU layer with 64 units, utilizing 'tanh' for activation and 'sigmoid' for recurrent transitions, effectively managing long sequence dependencies [28]. This layer is complemented by dropout rates of 0.4 and recurrent dropout rates of 0.2, optimizing the model against overfitting while maintaining generalizability. Following this, a dense layer with 64 units uses 'relu' activation to add non-linearity, enhancing the model's ability to discern complex data relationships. Additional stability is provided by a dropout rate of 0.3 and L2 regularization on the dense layer, which helps in reducing overfitting and stabilizing the model [29]. The architecture culminates in a dense output layer with three units using a 'softmax' activation function, allowing for precise classification of news into positive, neutral, or negative sentiment categories based on the generated probability distribution. Employing GRU models embodies a strategic choice in the study's analytical arsenal, leveraging their capability to effectively model temporal dependencies and nuances in sentiment within financial news text. This model, with its carefully calibrated architecture and parameters, underscores the study's comprehensive approach to employing state-of-the-art machine learning techniques for the nuanced task of sentiment analysis.

3.3.4. Convolutional neural networks

In the progression of exploring neural network architectures for sentiment analysis, the study incorporates CNNs, which are renowned for their exceptional feature extraction capabilities from sequential data, including text. Constructed with a sequential layout, the CNN model begins with an embedding layer that transforms text into 50-dimensional dense vectors, designed to handle a vocabulary of `max_words` and sequences up to `max_sequence_length`. This layer is vital for mapping words into a high-dimensional semantic space, enabling effective text processing [30], [31].

The model features a Conv1D layer with 128 filters and a kernel size of 3, activated by ‘relu’. This layer adeptly extracts key features and captures local patterns essential for identifying nuanced sentiments in financial news [32]. It’s followed by a GlobalMaxPooling1D layer that reduces dimensionality to highlight critical features for analysis [33]. The model is further refined with a dense layer of 64 units using ‘relu’ activation to learn complex relationships and a Dropout layer at a rate of 0.5 to prevent overfitting and maintain generalizability. The architecture concludes with a dense output layer that uses a ‘softmax’ activation to classify sentiments into categories based on probability. This CNN model exemplifies the methodological rigor of our study, employing advanced neural networks to extract and process text patterns, aiming to enhance the understanding of sentiment trends within the financial news domain significantly.

3.4. Transformers in sentiment analysis

The advent of transformer models has revolutionized the field of NLP, introducing a novel architecture that surpasses traditional methods in understanding the complexities of human language [34], [35]. Unlike previous models that processed text sequentially, transformers utilize a mechanism called Attention, allowing them to weigh the importance of different words within a sentence, regardless of their positional distance [36], [37]. This enables a deeper comprehension of context and semantic relationships, making transformers particularly effective for tasks like sentiment analysis. In this study, we harness the power of transformers, specifically the BERT, to elevate the accuracy of sentiment analysis within financial news. BERT, with its pre-trained model on a vast corpus of text, brings a profound understanding of language nuances, significantly enhancing our model’s ability to discern sentiment from text.

3.4.1. Bidirectional encoder representations from transformers

BERT is a crucial component of our method, enhancing the analysis of sentiment in financial news through its advanced deep learning and NLP capabilities. Developed by Google, BERT employs a unique architecture designed to pre-train deep bidirectional representations, conditioning on both left and right contexts at all layers. This method, performed on a large text corpus, allows BERT to capture an extensive range of language patterns and contexts, essential for understanding complex linguistic nuances [38]. We utilize BERT’s pre-trained models, fine-tuning them with our specific financial news dataset to better align with the nuances of financial sentiment analysis. This fine-tuning adjusts BERT’s parameters to improve its sensitivity to the unique contexts and idiomatic expressions found in financial news [39], enhancing its ability to classify sentiment accurately.

The application of BERT in our study involves processing preprocessed news articles to generate embeddings that reflect deep contextual relationships. These embeddings significantly enhance our neural network models, boosting the accuracy of our sentiment classification efforts. Ultimately, the integration of BERT not only marks a significant methodological advancement but also demonstrates the transformative potential of incorporating Transformer models alongside traditional neural network architectures, setting a new standard for depth and accuracy in financial news sentiment analysis.

3.5. The proposed model architecture

Our study’s model architecture employs a comprehensive approach, beginning with data preprocessing and polarity analysis to prepare the groundwork for advanced sentiment classification. Initially, the study constructs and evaluates various deep learning models specifically, ANN, LSTM networks, GRU, and CNN each specifically tailored for sentiment classification. These models are meticulously designed to capture and analyze the nuances of financial news sentiment, utilizing the processed and polarized dataset to accurately classify sentiment.

Simultaneously, our study adopts an innovative method by integrating the BERT model with each neural network architecture. This integration capitalizes on the deep contextual insights provided by BERT along with the specialized processing capabilities of each neural network. The goal is to enhance the accuracy of sentiment classification through the contextual embeddings generated by BERT, which offer a profound understanding of language nuances. This dual-pathway approach not only improves the models’ precision in detecting sentiments but also enriches the analysis by combining traditional neural network methodologies with advanced transformer technology.

3.5.1. Bidirectional encoder representations from transformers integration and application

In our study, BERT is methodically integrated into the model architecture to enhance sentiment analysis in financial news, employing a series of technical steps to ensure precision and effectiveness. The process begins with the dataset, consisting of financial news 'info' and their corresponding sentiment labels 'sentiment', undergoing label encoding and one-hot encoding. This preparation is critical as it formats the sentiment labels for classification tasks and prepares the data for effective processing by BERT and neural network classification.

The integration of BERT is pivotal, beginning with tokenization using the BertTokenizer from the 'bert-base-uncased' model [40]. This tokenizer adjusts the textual data into a consistent format by converting it into tokens and standardizing the input lengths to 128 tokens through padding and truncation. Such uniform tokenization is essential for maintaining dataset consistency, enabling BERT to efficiently generate contextual embeddings for each piece of news. Following tokenization, the processed data feeds into the BERT model, producing contextual embeddings that capture nuanced sentiment indicators within the text. These embeddings are crucial as they are then inputted into a neural network layer designed to process sequential data and capture temporal dependencies, key for tracking the evolution of sentiment. The neural network's output is refined through dropout layers to prevent overfitting, enhancing reliability and generalizability of the model. The final stage in our model's architecture involves a 'softmax' output layer that classifies the sentiment into one of three categories: positive, neutral, or negative, based on the refined embeddings provided by BERT.

This detailed approach not only allows for a nuanced understanding and classification of sentiment but also utilizes BERT's robust capabilities to set a new standard for accuracy and depth in sentiment analysis. Overall, by technically and strategically applying BERT alongside traditional neural networks, our study presents a sophisticated blend of deep learning and NLP technologies, significantly improving our models' ability to interpret and classify sentiment in financial news. This integration showcases the transformative potential of combining transformer models with conventional neural network architectures in advanced NLP tasks.

4. RESULTS AND DISCUSSION

The results of our study reveal a substantial improvement in the accuracy of sentiment analysis in financial news when utilizing BERT-infused deep learning models compared to their standalone counterparts. As indicated in Table 3, the baseline accuracies for ANN, LSTM, GRU, and CNN were 92.72%, 94.17%, 95.12%, and 94.03%, respectively. The infusion of BERT into these models resulted in a notable increase in performance, as shown in Table 4, where the accuracy rates for BERT-ANN, BERT-LSTM, BERT-GRU, and BERT-CNN models reached 96.36%, 96.74%, 97.02%, and 96.66%, respectively. This enhancement underscores the profound impact of integrating BERT's contextual understanding into neural network architectures, significantly enriching the models' capability to discern nuanced sentiment within textual data.

Table 3. Accuracy of deep learning models before infusing BERT

Model	Accuracy
ANN	0.9272
LSTM	0.9417
GRU	**0.9512**
CNN	0.9403

Table 4. Accuracy of deep learning models after infusing BERT

Model	Accuracy
BERT-ANN	0.9636
BERT-LSTM	0.9674
BERT-GRU	**0.9702**
BERT-CNN	0.9666

Further, our results were juxtaposed with those from existing models in the domain of financial sentiment analysis, as detailed in Table 5. The comparison delineates our BERT-GRU model's superior performance, achieving an accuracy of 97.02%, precision of 96.97%, recall of 97.07%, and an F1 score of 97.70%. This model not only outperforms traditional models such as the sentiment analyser [8], which achieved an F1 score of 75.60%, and noun verb approach [9] with an F1 score of 86.45%, but also advanced

NLP models like BART-large [3] and GPT-P2 [6], which reported lower accuracy rates of 94.70% and 79.00%, respectively. Notably, our model's excellence is highlighted against the backdrop of various methodologies, including Glove+BiGRU+Attention [5] and NLP transformers like GPT-P2 [6], underscoring the significant advancements made by integrating BERT with GRU for sentiment analysis in finance.

Table 5. Comparison results with existing models

Model	Accuracy	Precision	Recall	F1 Score
BART-large [3]	94.70	95.00	94.50	94.70
Glove+BiGRU+Attention [5]	-	-	-	89.30
GPT-P2 [6]	79.00	79.70	79.00	79.00
Sentiment analyser [8]	-	78.50	73.00	75.60
Noun verb approach [9]	79.00	78.82	95.71	86.45
Our model (BERT-GRU)	97.02	96.97	97.07	97.70

These findings not only affirm the efficacy of BERT-infused models in capturing and analyzing sentiment within financial news but also exemplify the leap in performance metrics against both traditional and contemporary NLP approaches. The discussion pivots on the essentiality of contextual embeddings provided by BERT, which, when coupled with the dynamic processing of GRU, renders a highly accurate and reliable sentiment analysis model. This synergy not only sets a new benchmark in the field but also prompts a reevaluation of existing models, advocating for a broader adoption of transformer-based techniques in financial sentiment analysis. Considering these results, the study contributes to the ongoing dialogue in NLP research, particularly in the financial domain, urging a paradigm shift towards more sophisticated, context-aware models. The success of the BERT-GRU model in our study is a testament to the transformative potential of leveraging deep learning and NLP technologies in harmony, paving the way for future innovations in sentiment analysis and beyond.

5. CONCLUSION

Our research clearly illustrates that integrating BERT with traditional neural network architectures such as ANN, LSTM, GRU, and CNN significantly boosts the accuracy of sentiment analysis within the financial news sector, setting a new standard for NLP advancements. This innovative approach leverages BERT's powerful contextual embeddings in conjunction with the dynamic capabilities of established neural networks, effectively capturing the subtle nuances of sentiment expressed in financial news. The substantial improvements in accuracy, precision, recall, and F1 scores not only enhance the performance of individual models but also establish a robust benchmark against traditional and advanced NLP models. This synthesis of deep learning and NLP technologies underlines the necessity for a shift towards more sophisticated, context-aware models, enhancing the precision and reliability of financial market analyses. Our findings advocate for this paradigm shift, highlighting the transformative potential of combining transformer-based techniques with neural network architectures for broader applications in sentiment analysis and beyond. As the field of sentiment analysis evolves amidst an era of information overload, our study provides pivotal insights and a solid foundation for future research endeavors, paving the way for continued innovation and improved analytical depth in financial sentiment analysis.

REFERENCES

[1] L. Mathew and B. V R, "Efficient Transformer based Sentiment Classification Models," *Informatica*, vol. 46, no. 8, Dec. 2022, doi: 10.31449/inf.v46i8.4332.

[2] L. Bacco, A. Cimino, F. Dell'Orletta, and M. Merone, "Explainable Sentiment Analysis: A Hierarchical Transformer-Based Extractive Summarization Approach," *Electronics*, vol. 10, no. 18, p. 2195, Sep. 2021, doi: 10.3390/electronics10182195.

[3] K. Mishev, A. Gjorgjevikj, I. Vodenska, L. T. Chitkushev, and D. Trajanov, "Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers," in *IEEE Access*, vol. 8, pp. 131662-131682, 2020, doi: 10.1109/ACCESS.2020.3009626.

[4] A. Sinha, S. Kedas, R. Kumar, and P. Malo, "SEntFiN 1.0: Entity-aware sentiment analysis for financial news," *Journal of the Association for Information Science and Technology*, vol. 73, no. 9, p. 1314–1335, Mar. 2022, doi: 10.1002/asi.24634.

[5] K. Mishev *et al.*, "Performance Evaluation of Word and Sentence Embeddings for Finance Headlines Sentiment Analysis," *ICT Innovations 2019. Big Data Processing and Mining*, vol. 1110, pp. 161–172, 2019, doi: 10.1007/978-3-030-33110-8_14.

[6] G. Fatouros, J. Soldatos, K. Kouroumali, G. Makridis, and D. Kyriazis, "Transforming sentiment analysis in the financial domain with ChatGPT," *Machine Learning with Applications*, vol. 14, p. 100508, Dec. 2023, doi: 10.1016/j.mlwa.2023.100508.

[7] T. Daudert, "Exploiting textual and relationship information for fine-grained financial sentiment analysis," *Knowledge-Based Systems*, vol. 230, p. 107389, Oct. 2021, doi: 10.1016/j.knosys.2021.107389.

[8] L. I. Tan, W. S. Phang, K. O. Chin, and A. Patricia, "Rule-Based Sentiment Analysis for Financial News," *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Hong Kong, China, 2015, pp. 1601-1606, doi:

10.1109/SMC.2015.283.
[9] A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," *Procedia Computer Science*, vol. 167, pp. 589–598, 2020, doi: 10.1016/j.procs.2020.03.325.
[10] L. Pham, "Financial News Headlines Data," *Kaggle*, 2020, [Online]. Available: https://www.kaggle.com/datasets/notlucasp/financial-news-headlines/data. (Accesed: Jul. 19, 2020).
[11] S. Pradha, M. N. Halgamuge, and N. T. Q. Vinh, "Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data," *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, Da Nang, Vietnam, 2019, pp. 1-8, doi: 10.1109/KSE.2019.8919368.
[12] V. Balakrishnan and E. Lloyd-Yemoh, "Stemming and Lemmatization: A Comparison of Retrieval Performances," *Lecture Notes on Software Engineering*, vol. 2, no. 3, pp. 262–267, 2014.
[13] J. Yao, "Automated Sentiment Analysis of Text Data with NLTK," *Journal of Physics: Conference Series*, vol. 1187, no. 5, p. 052020, Apr. 2019, doi: 10.1088/1742-6596/1187/5/052020.
[14] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014, doi: 10.1609/icwsm.v8i1.14550.
[15] T. Pano and R. Kashef, "A Complete VADER-Based Sentiment Analysis of Bitcoin (BTC) Tweets during the Era of COVID-19," *Big Data and Cognitive Computing*, vol. 4, no. 4, p. 33, Nov. 2020, doi: 10.3390/bdcc4040033.
[16] M. G. M. Abdolrasol *et al.*, "Artificial Neural Networks Based Optimization Techniques: A Review," *Electronics*, vol. 10, no. 21, p. 2689, Nov. 2021, doi: 10.3390/electronics10212689.
[17] I. Santos, L. Castro, N. Rodriguez-Fernandez, Á. Torrente-Patiño, and A. Carballal, "Artificial Neural Networks and Deep Learning in the Visual Arts: a review," *Neural Computing and Applications*, vol. 33, no. 1, pp. 121–157, Jan. 2021, doi: 10.1007/s00521-020-05565-4.
[18] S. Nosouhian, F. Nosouhian, and A. Kazemi Khoshouei, "A Review of Recurrent Neural Network Architecture for Sequence Learning: Comparison between LSTM and GRU," *Computer Science and Mathematics*, 2021, doi: 10.20944/preprints202107.0252.v1.
[19] R. Cahuantzi, X. Chen, and S. Güttel, "A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences," *Intelligent Computing*, vol. 739, pp. 771–785, 2023, doi: 10.1007/978-3-031-37963-5_53.
[20] I. G. B. J. Abasan and E. B. Setiawan, "Empowering hate speech detection: leveraging linguistic richness and deep learning," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 13, no. 2, pp. 1371–1382, Apr. 2024, doi: 10.11591/eei.v13i2.6938.
[21] S. Aich, S. Chakraborty, and H. Kim, "Convolutional neural network-based model for web-based text classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, p. 5785, Dec. 2019, doi: 10.11591/ijece.v9i6.pp5785-5191.
[22] C. Wei and Y. Lee, "Sequential forecast of incident duration using Artificial Neural Network models," *Accident Analysis & Prevention*, vol. 39, no. 5, pp. 944–954, Sep. 2007, doi: 10.1016/j.aap.2006.12.017.
[23] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," *arXiv*, doi: 10.48550/arXiv.1803.08375, 2018.
[24] D. Zhu, S. Lu, M. Wang, J. Lin, and Z. Wang, "Efficient Precision-Adjustable Architecture for Softmax Function in Deep Learning," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3382-3386, Dec. 2020, doi: 10.1109/TCSII.2020.3002564.
[25] H. Chen *et al.*, "An Efficient Hardware Architecture with Adjustable Precision and Extensible Range to Implement Sigmoid and Tanh Functions," *Electronics*, vol. 9, no. 10, p. 1739, Oct. 2020, doi: 10.3390/electronics9101739.
[26] F. Landi, L. Baraldi, M. Cornia, and R. Cucchiara, "Working Memory Connections for LSTM," *Neural Networks*, vol. 144, pp. 334–341, Dec. 2021, doi: 10.1016/j.neunet.2021.08.030.
[27] J. Yuan and Y. Tian, "An Intelligent Fault Diagnosis Method Using GRU Neural Network towards Sequential Data in Dynamic Processes," *Processes*, vol. 7, no. 3, p. 152, Mar. 2019, doi: 10.3390/pr7030152.
[28] S. Mahjoub, L. Chrifi-Alaoui, B. Marhic, and L. Delahoche, "Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks," *Sensors*, vol. 22, no. 11, p. 4062, May 2022, doi: 10.3390/s22114062.
[29] S. Shi, K. Hu, J. Xie, Y. Guo, and H. Wu, "Robust scientific text classification using prompt tuning based on data augmentation with L2 regularization," *Information Processing & Management*, vol. 61, no. 1, p. 103531, Jan. 2024, doi: 10.1016/j.ipm.2023.103531.
[30] H. Wang, J. He, X. Zhang, and S. Liu, "A Short Text Classification Method Based on N -Gram and CNN," *Chinese Journal of Electronics*, vol. 29, no. 2, pp. 248–254, Mar. 2020, doi: 10.1049/cje.2020.01.001.
[31] P. Prakash, S. K. Y. Hanumanthaiah, and S. B. Mayigowda, "CRNN model for text detection and classification from natural scenes," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, p. 839, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp839-849.
[32] D. S. Dewantara, I. Budi, and M. O. Ibrohim, "3218IR at SemEval-2020 Task 11: Conv1D and Word Embedding in Propaganda Span Identification at News Articles," *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp. 1716–1721, 2020, doi: 10.18653/v1/2020.semeval-1.225.
[33] B. Jang, M. Kim, G. Harerimana, S. Kang, and J. W. Kim, "Bi-LSTM Model to Increase Accuracy in Text Classification: Combining Word2vec CNN and Attention Mechanism," *Applied Sciences*, vol. 10, no. 17, p. 5841, Aug. 2020, doi: 10.3390/app10175841.
[34] B. AlBadani, R. Shi, J. Dong, R. Al-Sabri, and O. B. Moctard, "Transformer-Based Graph Convolutional Network for Sentiment Analysis," *Applied Sciences*, vol. 12, no. 3, p. 1316, Jan. 2022, doi: 10.3390/app12031316.
[35] K. L. Tan, C. P. Lee, K. S. M. Anbananthen and K. M. Lim, "RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network," in *IEEE Access*, vol. 10, pp. 21517-21525, 2022, doi: 10.1109/ACCESS.2022.3152828.
[36] S. Bhowmik, R. Sadik, W. Akanda, and J. R. Pavel, "Sentiment analysis with hotel customer reviews using FNet," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 13, no. 2, pp. 1298–1306, Apr. 2024, doi: 10.11591/eei.v13i2.6301.
[37] B. M. Gurusamy, P. K. Rengarajan, and P. Srinivasan, "A hybrid approach for text summarization using semantic latent Dirichlet allocation and sentence concept mapping with transformer," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 6, p. 6663, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6663-6672.
[38] J. Devlin "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv*, 2018, doi: 10.48550/arXiv.1810.04805.
[39] K. K, Z. Wang, S. Mayhew, and D. Roth, "Cross-Lingual Ability of Multilingual BERT: An Empirical Study," *arXiv*, Dec. 17, 2019, doi: 10.48550/arXiv.1912.07840.

[40] M. P. Geetha and D. K. Renuka, "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model," *International Journal of Intelligent Networks*, vol. 2, pp. 64–69, 2021, doi: 10.1016/j.ijin.2021.06.005.

BIOGRAPHIES OF AUTHORS

Oussama Ndama is a Ph.D. student in DSAI2S Research Team, C3S Laboratory, Faculty of Sciences and Technologies (FST), Tangier, Morocco. He had his Master in Computer Science and Big Data, Laureate of FST of Tangier. He is also a Business Intelligence Engineer with more than 5 years of experience in different multinational companies. The research topics of interest are smart systems, machine learning, deep learning, nlp, ann, sentiment analysis, and smart cities. He can be contacted at email: oussama.ndama@etu.uae.ac.ma.

Ismail Bensassi is a Ph.D. student in DSAI2S Research Team, C3S Laboratory, Faculty of Sciences and Technologies (FST), Tangier, Morocco. He is an Engineer in computer Science, Laureate of FST of Tangier. The research topics of interest are smart connection of user profiles in a big data context, multi-agent systems (MAS), case based reasoning (CBR), ontology, machine learning, smart cities, and eLearning/MOOC/SPOC. He can be contacted at email: bensassi.ismail@gmail.com.

Dr. El Mokhtar En-Naimi is a Full Professor in the University of Abdelmalek Essaâdi (UAE), Faculty of Sciences and Technologies of Tangier (FSTT), Department of Computer Sciences. He was Temporary Professor: from 1999 to 2003 and Permanent Professor: since 2003/2004 until Now. He is a full Professor in UAE, FST of Tangier. He was a Head of Computer Sciences Department, since October 2016 until the end of December 2020. He was responsible for a Licence of Science and Technology, LST Computer Engineering ("Licence LST-GI"), from January 2012 to October 2016. He is a Chief of DSAI2S Research Team since the academic year 2022/2023. He is also a founding member of the Both Laboratories: LIST (Laboratoire d'Informatique, Systèmes et Télécommunications) Laboratory (From 2008 To 2022) and C3S Laboratory since the academic year 2022/2023 until now, the University of Abdelmalek Essaâdi, FST of Tangier, Morocco. He is also an Expert Evaluator with the ANEAQ, since the academic year 2016/2017 until now, that an Expert of the Private Establishments belonging to the territory of the UAE and also an Expert of the Initial or Fundamental Formations and Formations Continuous at the Ministry of Higher Education, Scientific Research and Executive Training and also at the UAE University and the FST Tangier since 2012/2013 until now. He is an author/co-authors of several articles, published in The International Journals in Computer Sciences, in particular, in multi-agent systems (MAS), cases based reasoning (CBR), artificial intelligent (AI), machine learning (ML), deep learning (DL), eLearning, MOOC/SPOC, big data, data-mining, wireless sensor network, VANet, MANet, and smart city. He was/is also Director of several Doctoral Theses in Computer Sciences. He has too served as a general chair, technical program chair, technical program committee member, organizing committee member, session chair, and reviewer for many international conferences and workshops. In addition, he is an associate member of the ISCN - Institute of Complex Systems in Normandy, the University of the Havre, France, since 2009 until now. He can be contacted at email: en-naimi@uae.ac.ma.