Article

Not peer-reviewed version

Dynamics Computational Sentiment Analysis in Financial Markets

Jessica Carter , Wyne Nasir , Sophia Lee , Ethan Parker *

Posted Date: 15 April 2024

doi: 10.20944/preprints202404.0928.v1

Keywords: Sentiment Analysis; Financial Markets; Machine Learning

Article

Dynamics Computational Sentiment Analysis in Financial Markets

Jessica Carter, Wyne Nasir, Sophia Lee and Ethan Parker *

Tufts University; carter@tufts.edu (J.C.); wynenasir@gmail.com (W.N.); sophia@tufts.edu (S.L.)
* Correspondence: eparker@tufts.edu

Abstract: The flux of opinions in newsletters and social media platforms significantly influences market perceptions regarding corporate entities and their financial instruments. This paper introduces an advanced model, the Sentiment Dynamics Analyzer (SDA), which leverages a refined hierarchical architecture of Transformers to effectively discern the sentiment embedded within financial texts, ranging from news headlines to microblogs. We have enhanced a RoBERTa model with specialized sentiment lexicons and an augmented layer of Transformer models tailored for nuanced sentence-level sentiment detection, aiming to assign a sentiment score from -1 to +1. Our evaluations demonstrate that SDA not only surpasses the previous best methods but also exhibits superior performance over robust baseline models. This success underscores the utility of integrating tailored contextual analyses with sector-specific sentiment insights in improving predictive accuracy.

Keywords: sentiment analysis; financial markets; machine learning

---

1. Introduction

Sentiment analysis, often referred to as opinion mining, has evolved from basic text processing to a complex interdisciplinary field encompassing linguistics, psychology, and computational intelligence [1]. Historically, the initial attempts to understand textual sentiment involved simple heuristics-based methods, such as counting positive and negative words using pre-defined lexicons [2,3]. Over time, with the advent of machine learning, the approaches have shifted towards more sophisticated models that can grasp the subtleties and context of language [4,5]. This evolution was markedly accelerated with the introduction of social media platforms, where the vast and dynamic nature of user-generated content presented both new challenges and data opportunities for sentiment analysis.

In recent years, the focus has shifted towards deep learning techniques, particularly the use of neural networks, to handle the complexities of sentiment analysis. Models such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units) have demonstrated significant improvements in capturing long-range dependencies in text, essential for understanding sentiments expressed in longer documents or across multiple sentences [11,12]. The breakthrough, however, came with the introduction of Transformer-based models like BERT (Bidirectional Encoder Representations from Transformers), which utilize attention mechanisms to provide a contextually enriched representation of each word in a sentence [13]. These advancements have not only increased the accuracy of sentiment analysis tools but have also broadened their applicability across different languages and domains.

Sentiment analysis explores the complex web of emotions and opinions articulated by individuals concerning various entities such as products, services, or brands [11]. This scientific field seeks to parse these textual expressions to categorize sentiments as negative, neutral, or positive. Given the intricate and volatile nature of financial markets, understanding these sentiments is crucial as they can significantly sway market behavior.

In financial contexts, each stock or financial instrument, whether it's a commodity, a currency pair in the Forex market, a cryptocurrency, or other securities, can be viewed as an entity. Often, financial narratives might involve multiple such entities, each potentially associated with distinct sentiments [2].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 April 2024 doi:10.20944/preprints202404.0928.v1

A sophisticated sentiment analysis framework is essential to disentangle and accurately reflect the sentiment directed towards each entity.

The endeavor to predict stock market movements dates back to the early 20th century when pioneering analysts like Charles Dow developed techniques that laid the groundwork for technical analysis. Historically, financial prediction relied heavily on fundamental analysis, where the intrinsic value of stocks was calculated using economic and financial factors. Technical analysts, on the other hand, focused on price movements and trading volumes to identify patterns that could predict future market behavior. These traditional methods, however, often struggled with the inherent volatility and non-linear nature of financial markets.

With the integration of computational technologies in the late 20th century, quantitative finance began to embrace statistical and mathematical models to forecast market trends. The introduction of algorithms capable of processing large datasets led to the development of quantitative trading strategies, further revolutionizing financial prediction [19]. In the modern era, artificial intelligence and machine learning have taken center stage, employing models that can analyze not only numerical data but also unstructured data like news articles and social media posts. Techniques such as natural language processing have become critical in understanding market sentiment, which is increasingly recognized as a potent predictor of financial market movements. This shift towards data-driven, automated trading strategies underscores a significant transformation in how market analysts and investors predict the future.

Financial language possesses unique characteristics that set it apart from general language usage. Terms like "rally" have dual meanings: in finance, it refers to a sharp increase in stock prices, whereas outside finance, it could mean a large public gathering [24,25]. Financial texts, from objective newsletters to subjective microblogs, demand a nuanced understanding of both common and domain-specific lexicons.

This research is driven by the hypothesis that sentiment analysis can serve as a predictive tool for gauging market reactions. Reflecting on the challenges and innovations, which focused on fine-grained sentiment analysis of financial microblogs and news [26–29], it becomes evident that there is a pressing need for a robust model capable of deep financial text analysis and accurate sentiment prediction.

Our proposed SDA system commences with generating contextual representations of text, integrating general language understanding with financial lexicon insights [33]. This is followed by applying sentiment dictionaries to assign preliminary sentiment scores to individual words. A hierarchical stack of Transformers then processes these data points to construct a comprehensive sentiment profile, taking into account the context-dependent sentiment of each word [36]. The amalgamation of these sentiment and contextual representations yields a precise sentiment score for the entity discussed within the text.

The primary contributions of this paper are twofold: (i) We extend the capabilities of Transformer models, previously validated for their efficacy in various sentiment analysis applications, to the domain of financial text analysis, and (ii) We introduce the SDA model, which uniquely combines hierarchical Transformers with sentiment dictionaries to enhance sentiment detection and prediction in financial texts, thereby setting a new benchmark in the field and surpassing previous state-of-the-art models and robust baseline systems.

2. Related Work

The field of sentiment analysis, integral to natural language processing, focuses on deciphering the embedded opinions within texts. This domain generally treats the analysis as either a classification task, sorting emotions into categories such as negative, neutral, and positive, or as a regression task, quantifying sentiments on a continuous scale from $[-1, +1]$ [19,28,29,37]. The versatility of sentiment analysis is evidenced by its numerous applications, ranging from assessing product reviews to monitoring public opinion across digital platforms.

A notable milestone in this field was the focus of SemEval-2017 Task 4, which concentrated on analyzing sentiments expressed in multilingual tweets, necessitating the differentiation between general tweet sentiment and topic-specific sentiment [37]. This task underscored the complexity of sentiment analysis, especially in dynamically generated content such as social media.

Supporting tools for sentiment analysis include various sentiment lexicons, which provide essential resources for both general and domain-specific text analysis. Among these, SENTIWORDNET 3.0 offers a comprehensive sentiment classification framework by assigning positivity, negativity, and neutrality scores to the synsets of WORDNET [43]. Similarly, the "Affective Norms for English Words" (ANEW) provides emotional ratings for a broad spectrum of English language words [44]. Further contributions to this toolkit have been made by Wilson et al., who annotated thousands of subjective expressions to aid sentiment analysis [47].

The adaptation of sentiment analysis to the intricacies of social media was exemplified by Nielsen, who developed a sentiment dictionary specifically calibrated for tweets, spanning a sentiment range from -5 (highly negative) to $+5$ (highly positive) [48]. The financial sector, too, has seen bespoke developments, such as a market sentiment dictionary encompassing a broad array of terms derived from financial social media platforms [53].

Traditional models often fall short in the complex landscape of financial news, which necessitates a nuanced understanding of domain-specific vocabulary. Krishnamoorthy's approach employs financial and non-financial performance indicators within a sentiment analysis framework tailored for financial texts, showcasing the potential of hierarchical sentiment classifiers based on association rule mining [54].

Reflecting on the advancements from SemEval-2017 Task 5, which delved into the fine-grained sentiment analysis of financial microblogs and news, several entries highlighted the effectiveness of combining domain-specific sentiment dictionaries with advanced modeling techniques [19,26–29]. Techniques ranged from integrating linguistic and sentiment lexicons with ensemble regression models to employing convolutional and long-short term memory networks augmented by feature-based models, underscoring the diverse methodologies capable of capturing the nuanced sentiment of financial texts.

Recent strides in sentiment analysis have leveraged BERT-based models, which have significantly refined the analysis by integrating domain-specific and task-specific training enhancements [4,5,56]. These models have been instrumental in improving sentiment detection accuracy, especially in domain-specific contexts like financial news analysis.

Drawing inspiration from these developments, our proposed SDA method employs a BERT-based framework to enhance the sentiment analysis of financial texts. By integrating comprehensive general and financial sentiment dictionaries, our model achieves a deeper understanding of the sentiments expressed about entities within texts. This enriched analysis framework is further bolstered by a stack of Transformer encoder layers, which synthesize the contextual word representations with sentiment scores to derive an accurate sentiment score for each entity, quantified on a scale from $[-1, +1]$.

Building on this foundation, our methodology, detailed in the subsequent section, combines these state-of-the-art techniques to set a new standard in sentiment analysis, particularly tailored for the financial domain.

3. Method

Our research introduces the Sentiment Dynamics Analyzer (SDA), an innovative sentiment analysis framework that leverages a modified BERT architecture (Bidirectional Encoder Representations from Transformers) [57] tailored specifically for financial texts. Recognizing the inherent complexities of financial language, the SDA model incorporates additional Transformer layers and specialized sentiment dictionaries to enhance the granularity and accuracy of sentiment detection within financial documents.

3.1. System Overview

The foundational step in our approach involves the utilization of a pre-trained BERT model to initiate token representations of the textual data. This model, known for its efficacy in handling a wide range of NLP tasks, provides a robust starting point for feature extraction. Following this, we enrich these representations with sentiment scores derived from comprehensive sentiment dictionaries, specifically designed to capture the nuanced expressions typical of financial discourse.

Subsequent to the initial feature extraction, our model integrates n additional Transformer encoder layers, designed to refine and contextualize the sentiment representations associated with each token. This stacked Transformer architecture allows for a deeper analysis of the syntactic structures within the text, enabling a more nuanced understanding of sentiment dependencies and interactions. The process flow from tokenization to sentiment scoring is depicted. The tokenization itself is a two-step process leveraging BERT's built-in tokenizer for initial word separation, followed by the WordPiece [58] mechanism for subword tokenization. This ensures that even complex financial terms are effectively broken down into understandable units, as demonstrated by the segmentation of 'currencies' into subtokens ('cu', '##rre', '##ncies').

For sentiment quantification, SDA employs scores from two key lexical resources: SentiWordNet [59] and NTUSD [53]. SentiWordNet provides a structured sentiment scoring of synsets across dimensions of positivity, negativity, and neutrality, ideal for general linguistic analysis. NTUSD, on the other hand, offers insights grounded in financial lexicon, with its extensive dataset of labeled posts from financial social media, encompassing a wide range of words, hashtags, and emojis pertinent to market sentiment. Each token's representation is augmented by these scores, thereby enhancing the sentiment predictive power of the model.

Enhancing the model further, each Transformer encoder layer in our stack, adhering to the architecture proposed by Vaswani et al. [13], integrates multi-head attention mechanisms with sinusoidal position embeddings to maintain the sequence's temporal characteristics. These layers process the extended token representations—initially sized at 768 dimensions from BERT and expanded to 772 dimensions with sentiment scores—through both self-attention mechanisms and feed-forward networks, each supplemented with residual connections and layer normalization.

The culmination of this intricate process is a sophisticated aggregation of the contextual and sentiment-informed representations, particularly focusing on the [CLS] token—a common practice in BERT-based models for classification tasks. This aggregated representation is then processed through a final feed-forward layer, which synthesizes the insights derived from both the BERT-based model and the additional Transformer layers to produce a sentiment score, ranging from $[-1, 1]$. This score reflects a refined sentiment analysis, tailored to the specific challenges and nuances of financial text analysis.

3.2. Model Architecture

The core of the SAT model is based on a modified Transformer architecture, which is specifically adapted for sentiment analysis in the financial domain. The model processes textual data through multiple layers of Transformer blocks, each designed to capture different aspects of linguistic and sentiment information that are predictive of stock market behavior.

Tokenization and Embedding: Initially, financial texts are tokenized using a BERT-like tokenizer that includes subword tokenization to handle the financial jargon and neologisms common in stock market reports. The mathematical representation of the tokenization can be expressed as:

$$\mathbf{X} = \text{Tokenize}(\text{Text}),$$

where $\mathbf{X}$ represents the sequence of token indices.

Embedding Layer: Following tokenization, each token is converted into a high-dimensional vector using an embedding matrix $\mathbf{E}$. This is combined with position embeddings to maintain the sequence information, essential for processing the order-dependent nature of language. The embedding process is described by:

$$\mathbf{E}_{\text{total}} = \mathbf{E}_{\text{token}}(\mathbf{X}) + \mathbf{E}_{\text{pos}},$$

where $\mathbf{E}_{\text{total}}$ is the total embedding matrix for the input sequence, $\mathbf{E}_{\text{token}}$ is the token embedding matrix, and $\mathbf{E}_{\text{pos}}$ is the positional embedding matrix.

Transformer Blocks: The embedded tokens are then processed through a series of Transformer blocks. Each block consists of a multi-head self-attention mechanism followed by a position-wise fully connected feed-forward network. The self-attention mechanism allows the model to weigh the importance of different words relative to each other, which is critical for understanding contextual sentiments. The output of each Transformer block can be formulated as:

$$\mathbf{H}^{(l+1)} = \text{Transformer\_Block}(\mathbf{H}^{(l)}),$$

where $\mathbf{H}^{(l)}$ is the output of the l-th Transformer block, and l ranges from 1 to L, with L being the total number of Transformer blocks in the model.

Sentiment Scoring: After processing through the Transformer layers, the representation of the sequence's start token (often denoted as [CLS] in BERT-like models) is used to predict the sentiment score. This score is intended to reflect the overall sentiment of the text concerning the market or specific stocks. The sentiment score S is computed as:

$$S = \sigma(\mathbf{W} \cdot \mathbf{H}^{(L)}_{[\text{CLS}]} + b),$$

where σ is the sigmoid activation function, $\mathbf{W}$ is the weight matrix, b is the bias term, and $\mathbf{H}^{(L)}_{[\text{CLS}]}$ is the final layer representation of the [CLS] token.

3.3. Training and Optimization

The SAT model is trained end-to-end with a dataset of labeled financial texts, where each text is annotated with a sentiment score reflecting its predicted impact on stock prices. The training objective is to minimize the mean squared error (MSE) between the predicted sentiment scores and the actual scores, which is given by:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(S_i - \hat{S}_i)^2,$$

where N is the number of samples, S_i is the true sentiment score, and $\hat{S}_i$ is the predicted sentiment score for the i-th sample.

This comprehensive approach enables the SAT model to not only understand the general sentiment of the financial text but also its implications for stock market dynamics, providing a powerful tool for traders and financial analysts.

4. Experimental Setup

4.1. Datasets

Our experiments employ the datasets designed for SemEval-2017 Task 5, meticulously curated by the SSIX project [60]. These datasets encompass two principal types of financial data: microblogs and news headlines, reflecting a broad spectrum of market sentiments.

The microblog dataset comprises a vast array of financially pertinent messages harvested from Twitter and StockTwits, platforms that serve as vibrant hubs for traders and investors to exchange insights. Particularly, StockTwits messages are rich in shorthand financial jargon and stock symbols, directly reflecting market sentiments. Conversely, the news headlines dataset aggregates a diverse set of articles from esteemed sources including AP News and Reuters, offering a more structured and formal examination of market trends.

Each dataset entry is meticulously annotated with a sentiment score ranging from -1 (indicative of a negative or bearish outlook) to $+1$ (signifying a positive or bullish outlook), depending on the implied market trend within the message or headline. Table 1 provides a statistical overview of these datasets, highlighting the variance in text length and sentiment distribution, which underscores the challenges in analyzing such diverse and unstructured data.

Table 1. Representative samples from the SSIX datasets showing microblogs and news headlines with corresponding sentiment annotations.

Source		Entity	Text	Sentiment
Microblogs	Twitter	$AAPL	Anticipating a positive earnings report	0.65
	StockTwits	$TSLA	Concerns over quarterly revenue figures	-0.30
	StockTwits	$AMZN	Exceeds market expectations	0.75
	Twitter	$NFLX	Steady growth in subscriber base	0.40
News headlines		GM	GM to unveil new electric vehicle lineup	0.50
		Tesla	Tesla faces regulatory scrutiny	-0.45
		Netflix	Netflix announces a strategic shift towards more original content	0.35

4.2. Evaluation Metrics

To objectively assess the performance of sentiment analysis models, we adopt the same evaluation metrics as used in SemEval-2017 Task 5, with a focus on cosine similarity as the primary measure of accuracy. Cosine similarity quantitatively evaluates the alignment between the predicted sentiment scores and the expert-annotated gold standards, defined mathematically as:

$$cosine(G, P) = \frac{\sum_{i=1}^{n} G_i \times P_i}{\sqrt{\sum_{i=1}^{n} G_i^2} \times \sqrt{\sum_{i=1}^{n} P_i^2}} \tag{1}$$

The effectiveness of the models is further quantified by a weighted score that takes into account the proportion of scored instances, encouraging models to provide predictions across all data points:

$$score(G, P) = \frac{|P|}{|G|} \times cosine(G, P) \tag{2}$$

4.3. Baseline Models and Comparative Analysis

For our comparative analysis, we consider three leading BERT-based language models as baselines: BERT, RoBERTa, and DistilBERT [57,61,62]. These models serve as the foundational frameworks upon which our Sentiment Dynamics Analyzer (SDA) model is built. The SDA model extends these architectures by integrating additional Transformer encoder layers and a sophisticated sentiment lexicon to enhance the sentiment prediction capabilities specifically tuned for financial texts. Each baseline model's performance is benchmarked using the architecture, where the sentiment scores are derived from the [CLS] token's representation processed through a dedicated feed-forward layer.

4.4. Optimization and Training

The selection of the optimal model architecture was based on comprehensive preliminary tests across various configurations of Transformer encoder layers, ranging from 1 to 6. These tests were aimed at determining the most effective depth for capturing the nuances of financial sentiment. All models were trained using batch normalization, mean squared error loss, and the Adam optimizer, with hyperparameters fine-tuned to optimal levels.

Preprocessing involved standardizing the input data by removing any hyperlinks from the texts, ensuring that the models focused purely on the textual content without external influence.

5. Results and Analyses

This section presents the results from experimental evaluations of the baseline models alongside our proposed Sentiment Dynamics Analyzer (SDA) model. The performances are assessed using the microblog and news headline datasets, with results displayed in Tables 2 and 3. We compare the top-performing systems from SemEval-17 Task 5, several baseline models, and various configurations of our SDA model.

Table 2. Comparison of cosine similarity scores across different models on the microblog dataset, averaged over five runs. The highest scores are highlighted in bold.

Models	Cosine Similarity
Jiang et al. [26]	0.778
Ghosal et al. [27]	0.751
Deborah et al. [63]	0.735
BERT	0.789
DistilBERT	0.784
RoBERTa	0.832
SDA+1×Transf	0.833
SDA+2×Transf	0.832
SDA+3×Transf	0.831
SDA+4×Transf	0.834
SDA+5×Transf	**0.841**
SDA+6×Transf	0.833

Our analysis reveals that all BERT-based models surpass the SemEval-2017 top performers, demonstrating the evolving effectiveness of BERT-based architectures in financial text sentiment analysis. Notably, RoBERTa, upon which our SDA is based, shows superior results due to its optimized training strategy and the absence of a next-sentence prediction objective, enhancing its relevance for tasks focused purely on sentiment analysis [61].

Table 3. Performance metrics showing cosine similarity scores for various models evaluated on the news headlines dataset, averaged across five experimental runs. Best results are indicated in bold.

Models	Cosine Similarity
Mansar et al. [28]	0.745
Kar et al. [29]	0.744
Rotim et al. [64]	0.733
BERT	0.790
DistilBERT	0.782
RoBERTa	0.839
SDA+1×Transf	0.844
SDA+2×Transf	0.847
SDA+3×Transf	**0.848**
SDA+4×Transf	0.842
SDA+5×Transf	**0.848**
SDA+6×Transf	0.843

Incorporating additional Transformer layers in the SDA model has evidently enriched the sentiment analysis capabilities, as reflected by the consistently high performance across different levels of Transformer integration. The SDA model with five Transformer layers (SDA+5×Transf) consistently achieves or ties for the highest scores, underscoring the effectiveness of deeper Transformer stacks in capturing nuanced sentiment expressions in financial texts.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 April 2024 doi:10.20944/preprints202404.0928.v1

Both datasets present unique challenges: news headlines are generally longer and more structured, while microblogs often contain more informal, concise, and sometimes cryptic expressions of sentiment. This disparity underscores the adaptability of the SDA model, which performs robustly across both types of text.

5.1. Case Study

To further elucidate the practical implications of our findings, a detailed analysis of specific cases from both the microblogs and news headlines is provided in Table 4. This case study highlights the predicted sentiment scores alongside the gold standard (GS) for various models, illustrating the nuanced performance of each system in real-world scenarios.

Table 4. Case study comparing the sentiment predictions of different models against the gold standard for selected microblog and news headline texts.

Source	Entity	Text	GS	BERT	DistilBERT	RoBERTa	SDA +5×Transf
stocktwits	$UNG	Bought more $UNG puts	-0.875	-0.027	-0.139	-0.429	**-0.497**
stocktwits	$C	Cautious outlook as earnings approach	-0.093	**-0.082**	-0.060	-0.042	-0.074
stocktwits	$GDX	Strong day for Junior Gold Miners	0.750	0.707	0.659	0.698	**0.709**
News headline	Barclays	Barclays braces for a substantial forex penalty	-0.834	-0.604	-0.379	-0.652	**-0.672**
	Royal Mail	Minor increase in stamp prices announced	-0.05	0.198	0.132	**-0.018**	-0.003
	Ashtead	Share buyback announced after a profitable year	0.588	0.453	0.300	0.724	**0.639**

The analysis reveals that while our SDA model generally predicts sentiment scores closer to the gold standard, some challenges remain, particularly with neutral or mixed sentiment texts. This suggests an avenue for further refinement in handling texts with subtle sentiment cues or competing sentiment signals within the same text.

6. Conclusion and Future Work

This research focused on advancing sentiment analysis within the financial sector by leveraging state-of-the-art machine learning techniques. Sentiment analysis in this domain seeks to evaluate the sentiment related to entities such as companies or stocks, assigning a score within the range of $[-1, +1]$, where -1 indicates extremely negative sentiments and $+1$ represents highly positive sentiments. Our findings confirm that BERT-based models, renowned for their effectiveness in natural language processing tasks, are particularly adept at handling the nuanced language of finance [4,5,56]. Notably, all BERT-based baseline models demonstrated superior performance compared to the leading systems from the SemEval-2017 Task 5 on financial texts.

In our study, we introduced the Sentiment Dynamics Analyzer (SDA), an advanced deep learning framework that integrates a fine-tuned RoBERTa encoder with multiple Transformer encoder layers. This architecture was specifically designed to enhance sentiment analysis by merging general linguistic representations with targeted sentiment evaluations. The effectiveness of this approach is evidenced by the precise sentiment scores it produced, which were facilitated by the integration of sentiment dictionaries into the Transformer layers.

Our comprehensive evaluations revealed that the SDA model, particularly the configuration with five additional Transformer layers (SDA+5×Transf), consistently provided the most accurate sentiment representations. This model configuration not only achieved the highest cosine similarity scores across both structured and unstructured texts but also demonstrated significant improvements over the top-performing entries from SemEval-2017, with enhancements of 13.8% for news headlines and 8% for microblogs. These results underscore the robustness of our model in processing and analyzing complex financial datasets. This conclusion not only encapsulates the achievements of the current research but also sets a clear pathway for subsequent investigations, aiming to bridge the gap between sentiment analysis and practical financial forecasting.

Looking ahead, several areas of potential research have been identified. Future studies could explore sentiment analysis within lengthy news articles that discuss multiple entities with varying sentiments. This presents a considerable challenge due to the complex sentence structures and the diversity of sentiments that may be expressed within a single article. Additionally, addressing the imbalance in sentiment representation within datasets, particularly by increasing the proportion of negative examples, could further improve model training and accuracy.

Another promising direction involves extending our sentiment analysis framework to predict stock price movements based on sentiment trends derived from financial news and microblogs. By correlating sentiment scores with market performance, it may be possible to develop predictive models that offer valuable insights into stock price fluctuations. Ultimately, our goal is to refine and expand the capabilities of sentiment analysis to provide more nuanced and actionable insights for investors and financial analysts.

References

1. Sun, C.; Huang, L.; Qiu, X. Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *arXiv preprint arXiv:1903.09588* **2019**.
2. Sonkiya, P.; Bajpai, V.; Bansal, A. Stock price prediction using BERT and GAN, 2021, [arXiv:q-fin.ST/2107.09055].
3. Fei, H.; Zhang, M.; Ji, D. Cross-Lingual Semantic Role Labeling with High-Quality Translated Training Corpus. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 7014–7026.
4. Hoang, M.; Bihorac, O.A.; Rouces, J. Aspect-Based Sentiment Analysis using BERT. Proceedings of the 22nd Nordic Conference on Computational Linguistics; Linköping University Electronic Press: Turku, Finland, 2019; pp. 187–196.
5. Xu, H.; Liu, B.; Shu, L.; Yu, P. BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis. 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); Association for Computational Linguistics: Minneapolis, Minnesota, 2019; pp. 2324–2335. doi:10.18653/v1/N19-1242.
6. Wu, S.; Fei, H.; Li, F.; Zhang, M.; Liu, Y.; Teng, C.; Ji, D. Mastering the Explicit Opinion-Role Interaction: Syntax-Aided Neural Transition System for Unified Opinion Role Labeling. Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence, 2022, pp. 11513–11521.
7. Shi, W.; Li, F.; Li, J.; Fei, H.; Ji, D. Effective Token Graph Modeling using a Novel Labeling Strategy for Structured Sentiment Analysis. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2022, pp. 4232–4241.
8. Fei, H.; Zhang, Y.; Ren, Y.; Ji, D. Latent Emotion Memory for Multi-Label Emotion Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 7692–7699.
9. Wang, F.; Li, F.; Fei, H.; Li, J.; Wu, S.; Su, F.; Shi, W.; Ji, D.; Cai, B. Entity-centered Cross-document Relation Extraction. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, 2022, pp. 9871–9881.
10. Zhuang, L.; Fei, H.; Hu, P. Knowledge-enhanced event relation extraction via event ontology prompt. *Inf. Fusion* **2023**, *100*, 101919.
11. Tammewar, A.; Cervone, A.; Riccardi, G. Emotion Carrier Recognition from Personal Narratives. *Accepted for publication at INTERSPEECH* **2021**.
12. Fei, H.; Ren, Y.; Ji, D. Retrofitting Structure-aware Transformer Language Model for End Tasks. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 2151–2161.
13. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All You Need. Proceedings of the 31st International Conference on Neural Information Processing Systems; Curran Associates Inc.: Red Hook, NY, USA, 2017; NIPS'17, p. 6000–6010.

14. Fei, H.; Wu, S.; Li, J.; Li, B.; Li, F.; Qin, L.; Zhang, M.; Zhang, M.; Chua, T.S. LasUIE: Unifying Information Extraction with Latent Adaptive Structure-aware Generative Language Model. Proceedings of the Advances in Neural Information Processing Systems, NeurIPS 2022, 2022, pp. 15460–15475.
15. Qiu, G.; Liu, B.; Bu, J.; Chen, C. Opinion word expansion and target extraction through double propagation. *Computational linguistics* **2011**, *37*, 9–27.
16. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D.; Liang, X. Enriching contextualized language model from knowledge graph for biomedical information extraction. *Briefings in Bioinformatics* **2021**, *22*.
17. Wu, S.; Fei, H.; Ji, W.; Chua, T.S. Cross2StrA: Unpaired Cross-lingual Image Captioning with Cross-lingual Cross-modal Structure-pivoted Alignment. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 2593–2608.
18. Wu, S.; Fei, H.; Qu, L.; Ji, W.; Chua, T.S. NExT-GPT: Any-to-Any Multimodal LLM. *CoRR* **2023**, *abs/2309.05519*.
19. Cortis, K.; Freitas, A.; Daudert, T.; Huerlimann, M.; Zarrouk, M.; Handschuh, S.; Davis, B. SemEval-2017 Task 5: Fine-Grained Sentiment Analysis on Financial Microblogs and News. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 519–535. doi:10.18653/v1/S17-2089.
20. Fei, H.; Li, F.; Li, C.; Wu, S.; Li, J.; Ji, D. Inheriting the Wisdom of Predecessors: A Multiplex Cascade Framework for Unified Aspect-based Sentiment Analysis. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI, 2022, pp. 4096–4103.
21. Wu, S.; Fei, H.; Ren, Y.; Ji, D.; Li, J. Learn from Syntax: Improving Pair-wise Aspect and Opinion Terms Extraction with Rich Syntactic Knowledge. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 2021, pp. 3957–3963.
22. Li, B.; Fei, H.; Liao, L.; Zhao, Y.; Teng, C.; Chua, T.; Ji, D.; Li, F. Revisiting Disentanglement and Fusion on Modality and Context in Conversational Multimodal Emotion Recognition. Proceedings of the 31st ACM International Conference on Multimedia, MM, 2023, pp. 5923–5934.
23. Fei, H.; Liu, Q.; Zhang, M.; Zhang, M.; Chua, T.S. Scene Graph as Pivoting: Inference-time Image-free Unsupervised Multimodal Machine Translation with Visual Scene Hallucination. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 5980–5994.
24. Goonatilake, R.; Herath, S. The volatility of the stock market and news. *International Research Journal of Finance and Economics* **2007**, pp. 53–65.
25. Van de Kauter, M.; Breesch, D.; Hoste, V. Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with Applications* **2015**, *42*, 4999–5010. doi:https://doi.org/10.1016/j.eswa.2015.02.007.
26. Jiang, M.; Lan, M.; Wu, Y. ECNU at SemEval-2017 Task 5: An Ensemble of Regression Algorithms with Effective Features for Fine-Grained Sentiment Analysis in Financial Domain. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 888–893. doi:10.18653/v1/S17-2152.
27. Ghosal, D.; Bhatnagar, S.; Akhtar, M.S.; Ekbal, A.; Bhattacharyya, P. IITP at SemEval-2017 Task 5: An Ensemble of Deep Learning and Feature Based Models for Financial Sentiment Analysis. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 899–903. doi:10.18653/v1/S17-2154.
28. Mansar, Y.; Gatti, L.; Ferradans, S.; Guerini, M.; Staiano, J. Fortia-FBK at SemEval-2017 Task 5: Bullish or Bearish? Inferring Sentiment towards Brands from Financial News Headlines. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 817–822. doi:10.18653/v1/S17-2138.
29. Kar, S.; Maharjan, S.; Solorio, T. RiTUAL-UH at SemEval-2017 Task 5: Sentiment Analysis on Financial Data Using Neural Networks. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 877–882. doi:10.18653/v1/S17-2150.
30. Li, J.; Xu, K.; Li, F.; Fei, H.; Ren, Y.; Ji, D. MRN: A Locally and Globally Mention-Based Reasoning Network for Document-Level Relation Extraction. Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 2021, pp. 1359–1370.

31. Fei, H.; Wu, S.; Ren, Y.; Zhang, M. Matching Structure for Dual Learning. Proceedings of the International Conference on Machine Learning, ICML, 2022, pp. 6373–6391.
32. Cao, H.; Li, J.; Su, F.; Li, F.; Fei, H.; Wu, S.; Li, B.; Zhao, L.; Ji, D. OneEE: A One-Stage Framework for Fast Overlapping and Nested Event Extraction. Proceedings of the 29th International Conference on Computational Linguistics, 2022, pp. 1953–1964.
33. Astudillo, R.F.; Ballesteros, M.; Naseem, T.; Blodgett, A.; Florian, R. Transition-based Parsing with Stack-Transformers. Findings of EMNLP, 2020. doi:10.18653/v1/2020.findings-emnlp.89.
34. Fei, H.; Li, F.; Li, B.; Ji, D. Encoder-Decoder Based Unified Semantic Role Labeling with Label-Aware Syntax. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, pp. 12794–12802.
35. Li, B.; Fei, H.; Li, F.; Wu, Y.; Zhang, J.; Wu, S.; Li, J.; Liu, Y.; Liao, L.; Chua, T.S.; Ji, D. DiaASQ: A Benchmark of Conversational Aspect-based Sentiment Quadruple Analysis. Findings of the Association for Computational Linguistics: ACL 2023, 2023, pp. 13449–13467.
36. Georgiou, E.; Papaioannou, C.; Potamianos, A. Deep Hierarchical Fusion with Application in Sentiment Analysis. Proc. Annual Conference of the Int'l Speech Communication Association (INTERSPEECH). ISCA, 2019, pp. 1646–1650. doi:10.21437/Interspeech.2019-3243.
37. Rosenthal, S.; Farra, N.; Nakov, P. SemEval-2017 Task 4: Sentiment Analysis in Twitter. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 502–518. doi:10.18653/v1/S17-2088.
38. Wu, S.; Fei, H.; Cao, Y.; Bing, L.; Chua, T.S. Information Screening whilst Exploiting! Multimodal Relation Extraction with Feature Denoising and Multimodal Topic Modeling. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 14734–14751.
39. Fei, H.; Wu, S.; Ren, Y.; Li, F.; Ji, D. Better Combine Them Together! Integrating Syntactic Constituency and Dependency Representations for Semantic Role Labeling. Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, 2021, pp. 549–559.
40. Wu, S.; Fei, H.; Zhang, H.; Chua, T.S. Imagine That! Abstract-to-Intricate Text-to-Image Synthesis with Scene Graph Hallucination Diffusion. *Advances in Neural Information Processing Systems* **2024**, *36*.
41. Fei, H.; Wu, S.; Ji, W.; Zhang, H.; Chua, T.S. Empowering dynamics-aware text-to-video diffusion with large language models. *arXiv preprint arXiv:2308.13812* **2023**.
42. Qu, L.; Wu, S.; Fei, H.; Nie, L.; Chua, T.S. Layoutllm-t2i: Eliciting layout guidance from llm for text-to-image generation. Proceedings of the 31st ACM International Conference on Multimedia, 2023, pp. 643–654.
43. Baccianella, S.; Esuli, A.; Sebastiani, F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC; Calzolari, N.; Choukri, K.; Maegaard, B.; Mariani, J.; Odijk, J.; Piperidis, S.; Rosner, M.; Tapias, D., Eds. European Language Resources Association, 2010.
44. Bradley, M.; Lang, P. Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. Technical Report C-1, 1999.
45. Fei, H.; Ren, Y.; Ji, D. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction. *Information Processing & Management* **2020**, *57*, 102311.
46. Li, J.; Fei, H.; Liu, J.; Wu, S.; Zhang, M.; Teng, C.; Ji, D.; Li, F. Unified Named Entity Recognition as Word-Word Relation Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, pp. 10965–10973.
47. Wilson, T.; Wiebe, J.; Hoffmann, P. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing; Association for Computational Linguistics: USA, 2005; HLT '05, p. 347–354. doi:10.3115/1220575.1220619.
48. Nielsen, F.Å. A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs. Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages, Heraklion, Crete, Greece, May 30, 2011; Rowe, M.; Stankovic, M.; Dadzie, A.; Hardey, M., Eds. CEUR-WS.org, 2011, Vol. 718, *CEUR Workshop Proceedings*, pp. 93–98.
49. Fei, H.; Chua, T.; Li, C.; Ji, D.; Zhang, M.; Ren, Y. On the Robustness of Aspect-based Sentiment Analysis: Rethinking Model, Data, and Training. *ACM Transactions on Information Systems* **2023**, *41*, 50:1–50:32.
50. Zhao, Y.; Fei, H.; Cao, Y.; Li, B.; Zhang, M.; Wei, J.; Zhang, M.; Chua, T. Constructing Holistic Spatio-Temporal Scene Graph for Video Semantic Role Labeling. Proceedings of the 31st ACM International Conference on Multimedia, MM, 2023, pp. 5281–5291.

51. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D. Nonautoregressive Encoder-Decoder Neural Framework for End-to-End Aspect-Based Sentiment Triplet Extraction. *IEEE Transactions on Neural Networks and Learning Systems* **2023**, *34*, 5544–5556.
52. Zhao, Y.; Fei, H.; Ji, W.; Wei, J.; Zhang, M.; Zhang, M.; Chua, T.S. Generating Visual Spatial Description via Holistic 3D Scene Understanding. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 7960–7977.
53. Chung-Chi Chen, H.H.H.; Chen, H.H. NTUSD-Fin : A Market Sentiment Dictionary for Financial Social Media Data Applications. Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018); El-Haj, M.; Rayson, P.; Moore, A., Eds.; European Language Resources Association (ELRA): Paris, France, 2018.
54. Krishnamoorthy, S. Sentiment analysis of financial news articles using performance indicators. *Knowledge and Information Systems* **2017**, *56*, 373–394.
55. Fei, H.; Li, B.; Liu, Q.; Bing, L.; Li, F.; Chua, T.S. Reasoning Implicit Sentiment with Chain-of-Thought Prompting. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2023, pp. 1171–1182.
56. Sousa, M.G.; Sakiyama, K.; Rodrigues, L.d.S.; Moraes, P.H.; Fernandes, E.R.; Matsubara, E.T. BERT for Stock Market Sentiment Analysis. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), 2019, pp. 1597–1601. doi:10.1109/ICTAI.2019.00231.
57. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 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); Association for Computational Linguistics: Minneapolis, Minnesota, 2019; pp. 4171–4186. doi:10.18653/v1/N19-1423.
58. Wu, Y.; Schuster, M.; Chen, Z.; Le, Q.V.; Norouzi, M.; Macherey, W.; Krikun, M.; Cao, Y.; Gao, Q.; Macherey, K.; others. Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144* **2016**.
59. Baccianella, S.; Esuli, A.; Sebastiani, F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC; Calzolari, N.; Choukri, K.; Maegaard, B.; Mariani, J.; Odijk, J.; Piperidis, S.; Rosner, M.; Tapias, D., Eds. European Language Resources Association, 2010.
60. Davis, B.; Cortis, K.; Vasiliu, L.; Koumpis, A.; McDermott, R.; Handschuh, S. Social Sentiment Indices Powered by X-Scores. 2nd International Conference on Big Data, Small Data, Linked Data and Open Data, ALLDATA; , 2016.
61. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach, 2019, [arXiv:cs.CL/1907.11692].
62. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, 2020, [arXiv:cs.CL/1910.01108].
63. S, A.D.; Rajendram, S.M.; Mirnalinee, T.T. SSN_MLRG1 at SemEval-2017 Task 5: Fine-Grained Sentiment Analysis Using Multiple Kernel Gaussian Process Regression Model. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 823–826. doi:10.18653/v1/S17-2139.
64. Rotim, L.; Tutek, M.; Šnajder, J. TakeLab at SemEval-2017 Task 5: Linear aggregation of word embeddings for fine-grained sentiment analysis of financial news. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 866–871. doi:10.18653/v1/S17-2148.