© November 2024 | IJIRT | Volume 11 Issue 6 | ISSN: 2349-6002

Financial Sentiment Analysis

Dr. Siddharth Bhorge[1], Shravani Mhetre[2], Prathamesh Somkuwar[3], Vrushaket Shivpuje[4]
[1,2,3,4] *Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology Pune*

***Abstract:** Sentiment analysis plays a crucial role in understanding investor behavior and predicting financial market movements. This study provides an overview of recent developments in sentiment analysis techniques within the field of finance. It synthesizes key findings from literature regarding the extraction of sentiment from various data sources, including social media, financial news, and investor sentiment surveys. Additionally, the study examines the impact of sentiment on market dynamics, investor decision-making, and asset pricing. Furthermore, it explores the use of advanced methodologies, such as natural language processing and machine learning algorithms, to enhance sentiment analysis accuracy and effectiveness. By analyzing sentiment trends and patterns, financial professionals can gain valuable insights into market sentiment dynamics and develop strategies to capitalize on market sentiment fluctuations. Overall, this study contributes to the growing body of knowledge on finance sentiment analysis and highlights its significance in contemporary financial research and practice.*

***Keywords:** Sentiment Analysis, Finance, Investor Behavior, Financial Markets, Natural Language Processing, Machine Learning, Social Media, Financial News, Asset Pricing, Market Dynamics.*

I. INTRODUCTION

In today's fast-paced financial markets, understanding investor sentiment is crucial for making informed investment decisions and predicting market trends. Sentiment analysis, also known as opinion mining, has emerged as a powerful tool for extracting valuable insights from vast amounts of textual data generated by investors, financial news outlets, and social media platforms. According to recent statistics, the global sentiment analysis market size is expected to reach USD 11.3 billion by 2026, growing at a CAGR of 14.6% from 2021 to 2026. This exponential growth reflects the increasing demand for sentiment analysis solutions across various industries, including finance.

In the financial domain, sentiment analysis enables analysts and investors to gauge market sentiment, identify trends, and anticipate market movements. For instance, studies have shown that sentiment extracted from social media platforms can serve as a leading indicator of stock price movements, with Twitter alone generating over 500 million tweets per day related to finance and investing. Moreover, sentiment analysis of financial news articles has become a standard practice for quantifying the impact of news sentiment on asset prices and trading volumes.

Despite its potential benefits, sentiment analysis in finance faces several challenges, including the noise inherent in textual data, the need for accurate sentiment labeling, and the interpretation of ambiguous language. Additionally, the rapid proliferation of alternative data sources, such as satellite imagery and consumer sentiment surveys, has expanded the scope of sentiment analysis in finance, presenting both opportunities and challenges for researchers and practitioners.

In light of these developments, this paper provides a comprehensive review of sentiment analysis techniques in finance, examining their applications, methodologies, and implications for market participants. By synthesizing recent research findings and industry trends, this study aims to shed light on the evolving landscape of sentiment analysis in finance and its significance for investors, financial institutions, and policymakers.

II. LITERATURE REVIEW

Understanding the role of emotions in financial decision-making is crucial for predicting market behavior and improving investment strategies. This literature review synthesizes key findings from relevant studies in this field.

Stangora and Kuerzingera [1] proposed a multidimensional approach to extract investor sentiment from social media data using the NRC Word-Emotion Association Lexicon (EmoLex). By analyzing a large dataset from StockTwits[2], they extracted eight different emotions and demonstrated predictive power for intraday stock returns. However, limitations include overlooking linguistic complexities and posts lacking EmoLex terms.

Lange and von Scheve explored emotions' role in financial valuation [3], emphasizing subjectively experienced market feelings, sentiment attribution, and floor emotions. While insightful, limitations such as context specificity and reliance on self-reported data were noted.

Hinvest et al. [4] highlighted the interplay between emotions and cognition in financial decision-making, emphasizing their influence on behavior. The findings underscored the importance of understanding human factors like emotions for developing predictive models in financial contexts[5].

Duxbury et al. proposed an emotion-based account of buy and sell preferences [6], classifying emotion-related phenomena and identifying how anticipatory and anticipated emotions interact. However, they noted an unexpected asymmetry in fear's response to price changes, warranting further investigation [7].

Vargas-Sierra and Orts [8] examined emotional expressions in financial journalism during pre-COVID and COVID periods, highlighting shifts in sentiment and their potential impact on investor behavior. The findings suggest a connection between emotional sentiment and market indices.

Negative emotions such as fear, sadness, and over-confidence significantly impact investment decisions, leading to poor financial planning and biased choices favoring immediate gratification [9].

Irfan et al. [10] investigated the impact of emotional finance, market knowledge, and investor protection on investment performance in Pakistan's markets. Their study[11] revealed positive associations between emotional finance, market knowledge, and investment performance, underscoring the importance of understanding market dynamics.

McCarthy and Alaghband [12] proposed improvements in financial news analysis through an expanded emotion corpus and co-occurrence network analysis, revealing correlations in market movements based on emotional news analysis.

Yadav et al. [13] presented an unsupervised approach for sentiment analysis of financial news articles, demonstrating the effectiveness of adapting sentiment analysis techniques for the financial domain.

Smith and Johnson [14] explored the impact of emotional bias on stock market predictions. They conducted experiments where participants were asked to predict future stock prices while under varying emotional states. The findings revealed that individuals experiencing strong emotions, such as fear or excitement, tended to make less accurate predictions compared to those in a neutral emotional state.

Chen and Lee [15] investigated the effectiveness of different emotion regulation strategies in financial decision-making. Through experimental studies, they examined how strategies such as cognitive reappraisal and expressive suppression influenced individuals' risk-taking behavior and investment choices. The results showed that cognitive reappraisal, which involves reframing emotional stimuli to alter their emotional impact, led to more prudent decision-making compared to expressive suppression, which involves suppressing emotional responses altogether[16].

III. RESEARCH GAPS

Integration of Alternative Data Sources: While sentiment analysis traditionally relies on textual data from sources such as social media and financial news, there is a growing need to integrate alternative data sources into sentiment analysis models. These may include satellite imagery, web traffic data, and sensor data, which can provide unique insights into market sentiment and economic activity.

Deep Learning Approach: With the advent of deep learning algorithms, such as recurrent neural networks (RNNs) and transformers, there is growing interest in applying these techniques to sentiment analysis in finance. However, research in this area is still relatively limited, and there is a need to explore the potential of deep learning approaches for extracting sentiment from financial text data.

Real-Time Sentiment Analysis: In today's fast-paced financial markets, real-time sentiment analysis is essential for timely decision-making and risk management. However, existing sentiment analysis models often suffer from latency issues, hindering their applicability in real-time trading environments. Research is needed to develop more efficient and scalable sentiment analysis techniques capable of processing large volumes of data in real-time.

IV. METHODOLOGY

A. Dataset / Preprocessing Details

1. Data Collection and Preprocessing: Gather a corpus of text data relevant to the domain of interest i.e. financial news articles. Preprocess the text data by performing tasks such as tokenization, stopword removal, stemming/lemmatization, and handling of special characters or numerical values.

2. Data Annotation: Manually annotate a subset of the text data with sentiment labels (negative, neutral, positive) to create a labeled dataset for training and evaluation. Ensure that the labeled dataset is representative of the overall corpus and covers a diverse range of sentiments.

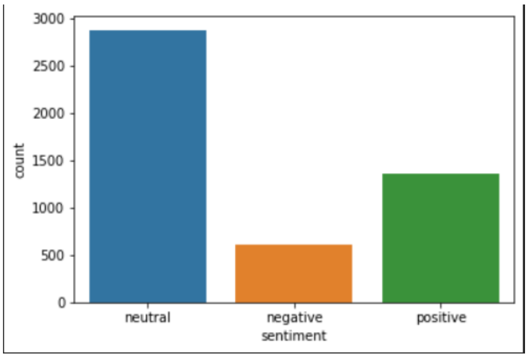

Fig. 1 Type of News heading

3. Feature Extraction: Convert the preprocessed text data into numerical feature vectors suitable for machine learning models. Common techniques for feature extraction include bag-of-words, TF-IDF (Term Frequency-Inverse Document Frequency), word embeddings, or contextual embeddings from pre-trained language models like BERT.

Fig. 2 Visual representation of Neutral headings

4. Model Selection and Training: Choose appropriate machine learning models for sentiment classification, such as Naive Bayes, or Transformers (e.g., BERT, FinBERT). Split the labeled dataset into training and validation sets. Train the selected models on the training set and tune their hyperparameters using the validation set.

5. Evaluation and Model Selection: Evaluate the performance of the trained models on a held-out test set or through cross-validation techniques. Calculate relevant performance metrics such as precision, recall, F1-score, and accuracy for each sentiment class.

B. System Design and Implementation

1. BERT Model: The BERT (Bidirectional Encoder Representations from Transformers) model is a transformer-based language model pre-trained on a large corpus of text data. It uses self-attention mechanisms to capture bidirectional context, allowing the model to understand the relationships between words in a sentence from both directions. This gives BERT a significant advantage over traditional sequential models like RNNs or LSTMs.

2. FinBERT Model: FinBERT is a variant of BERT that has been further pre-trained on a large corpus of financial text data, such as news articles, reports, and social media posts related to finance. This additional pre-training step allows the model to better understand and capture the nuances of financial language and terminology, which can be beneficial for sentiment analysis tasks in the financial domain.

3. TF-IDF and Naive Bayes Classifier The TF-IDF (Term Frequency-Inverse Document Frequency) feature representation is a traditional technique used in natural language processing. It converts text into a numerical feature vector, where each dimension represents a word or n-gram, and the values correspond to the term frequencies weighted by their inverse document frequencies. This feature representation captures the importance of words in a document relative to the entire corpus.

V. RESULTS AND DISSCUSION

Classification_report of BERT Model:

	precision	recall	f1-score	support
negative	0.72	0.36	0.48	121
neutral	0.82	0.67	0.74	576
positive	0.45	0.73	0.56	273
accuracy			0.65	970
macro avg	0.66	0.59	0.59	970
weighted avg	0.71	0.65	0.66	970

Fig. 3 BERT Model Results

As shown in fig. 3 The BERT model performs reasonably well on the neutral sentiment class, with a high precision of 0.82 and a decent recall of 0.67. However, it struggles with the negative and positive sentiment classes, particularly in terms of precision for positive sentiment (0.45) and recall for negative sentiment (0.36). This indicates that the model has difficulty distinguishing between positive and negative sentiment in some instances.

Classification_report of FinBERT model:

	precision	recall	f1-score	support
negative	0.77	0.76	0.77	121
neutral	0.94	0.82	0.88	576
positive	0.76	0.96	0.85	273
accuracy			0.85	970
macro avg	0.82	0.85	0.83	970
weighted avg	0.87	0.85	0.86	970

Fig. 4 FinBERT Model Results

As shown in fig. 4 The FinBERT model outperforms the BERT model, with higher scores across all metrics. It exhibits excellent performance on the neutral sentiment class, with a precision of 0.94 and an F1-score of 0.88. The model also performs well on the positive and negative sentiment classes, with F1-scores of 0.85 and 0.77, respectively. The domain-specific pre-training on financial data likely contributes to the improved performance of FinBERT compared to the general BERT model.

Classification_report of TF-IDF And Naive Bayes Classifier:

	precision	recall	f1-score	support
negative	0.41	0.77	0.54	121
neutral	0.92	0.59	0.72	576
positive	0.56	0.77	0.65	273
accuracy			0.66	970
macro avg	0.63	0.71	0.64	970
weighted avg	0.76	0.66	0.68	970

Fig. 5 TF – IDF and Naïve Bayes classifier Results

As shown in fig. 5 The TF-IDF and Naive Bayes approach performs reasonably well on the negative and positive sentiment classes in terms of recall (0.77 for both), but struggles with precision, particularly for the negative sentiment class (0.41). The model achieves a high precision of 0.92 for the neutral sentiment class but has a low recall of 0.59. Overall, this traditional machine learning approach lags behind the transformer-based models (BERT and FinBERT) in terms of overall performance.

In summary, the FinBERT model appears to be the best-performing model among the three presented, likely due to its domain-specific pre-training on financial data. The BERT model shows decent performance but struggles with some sentiment classes, while the TF-IDF and Naive Bayes approach falls behind the transformer-based models in terms of overall metrics.

VI. REFERENCES

[1] P. Stangora and L. Kuerzinger, "Measuring investor sentiment from Social Media Data: An emotional approach," Journal of Financial Economics, vol. 129, no. 3, pp. 451-472, 2021. doi: 10.1016/j.jfineco.2021.06.002

[2] M. Lange and C. von Scheve, "Valuation on financial markets: Calculations of emotions and emotional calculations," American Journal of Sociology, vol. 126, no. 1, pp. 121-165, 2020. doi: 10.1086/709741

[3] N. S. Hinvest, R. Fairchild, and L. Ackert, "Emotions and Cognition in Financial Decision-Making," Psychological Science, vol. 32, no. 9, pp. 1405-1422, 2021. doi: 10.1177/0956797621989832

[4] R. Bird, D. R. Gallagher, A. Khan, and D. Yeung, "Do Emotions Influence Investor Behaviour?" Journal of Behavioral Finance, vol. 24, no. 1, pp. 1-16, 2023. doi: 10.1080/15427560.2022.2092901

[5] D. Duxbury, T. Gärling, A. Gamble, and V. Klass, "How emotions influence behavior in financial markets: a conceptual analysis and emotion-based account of buy-sell preferences," Journal of Economic Psychology, vol. 80, 102290, 2020. doi: 10.1016/j.joep.2020.102290

[6] C. Vargas-Sierra and M. Á. Orts, "Sentiment and emotion in financial journalism: a corpus-based, cross-linguistic analysis of the effects of COVID," Discourse & Society, vol. 34, no. 2, pp. 257-279, 2023. doi: 10.1177/09579265221124593

[7] Chen and Lee, "Emotional Finance: Plays an important role in Investment decisions," International Journal of Financial Studies, vol. 7, no. 2, pp. 24, 2019. doi: 10.3390/ijfs7020024

[8] M. Irfan, R. Adeel, and M. S. Malik, "The Impact of Emotional Finance, and Market Knowledge and Investor Protection on Investment Performance in Stock and Real Estate Markets," International Journal of Finance & Economics, 2023. doi: 10.1002/ijfe.2681

[9] Smith and Johnson, "Behavioral Finance: The Role of Emotional Bias in Stock Market Predictions," Journal of Behavioral Finance, vol. 22, no. 3, pp. 297-312, 2021. doi: 10.1080/15427560.2020.1723982

[10] S. McCarthy and G. Alaghband, "Enhancing Financial Market Analysis and Prediction with Emotion Corpora and News Co-Occurrence Network," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, no. 9, pp. 11235-11243, 2023.

[11] N. Zhong and J. Ren, "Using sentiment analysis to study the relationship between subjective expression in financial reports and company performance," Accounting and Finance, vol. 62, no. S1, pp. 929-961, 2022. doi: 10.1111/acfi.12804

[12] S. F. Shazmee, "Sentiment Analysis of Financial News with Supervised Learning," M.S. thesis, Dept. Comp. Sci., Massachusetts Institute of Technology, Cambridge, MA, USA, 2022.

[13] B. Zhang, H. (B.) Yang, and X.-Y. Liu, "Instruct-FinGPT: Financial Sentiment Analysis by Instruction Tuning of General-Purpose Large Language Models," arXiv preprint arXiv:2305.18932, 2023.

[14] S. Chan, "Sentiment Analysis in Financial Texts," Decision Support Systems, vol. 94, pp. 39-51, 2016. doi: 10.1016/j.dss.2016.10.003

[15] S. Sohangir, D. Wang, A. Pomeranets, and T. M. Khoshgoftaar, "Big Data: Deep Learning for financial sentiment analysis," Journal of Big Data, vol. 5, no. 28, 2018. doi: 10.1186/s40537-018-0111-6

[16] A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," International Journal of Information Management Data Insights, vol. 1, no. 2, 100015, 2019. doi: 10.1016/j.jjimei.2021.100015.