

# Research on the Nonlinear Impact of Investor Sentiment on Stock Returns Based on Deep Learning and Text Mining

Fangchen Liu

Faculty of Mathematics, University of Waterloo, Waterloo, Canada  
Email: davidliu1526@gmail.com

**How to cite this paper:** Liu, F. C. (2025). Research on the Nonlinear Impact of Investor Sentiment on Stock Returns Based on Deep Learning and Text Mining. *American Journal of Industrial and Business Management*, 15, 1676-1684.

<https://doi.org/10.4236/ajibm.2025.1511087>

**Received:** August 12, 2025

**Accepted:** November 10, 2025

**Published:** November 13, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

This paper focuses on the nonlinear correlation between investor sentiment and stock returns and conducts in-depth research with the aid of deep learning and text mining techniques. First of all, sort out the relevant theoretical cornerstones, covering behavioral finance, market efficiency theory and herd effect theory, to provide theoretical support for the research. Secondly, analyze and study the data context, including the characteristics of text data and stock market data, as well as the necessity of integrated correlation analysis. In terms of extracting investor sentiment indicators, compare the advantages and disadvantages of the sentiment dictionary method, the machine learning method and the deep learning method. Subsequently, an in-depth exploration was conducted on the manifestations, theoretical explanations and implications for investment decisions of the nonlinear impact of investor sentiment on stock returns. Empirical results show that investor sentiment has a significant nonlinear impact on stock returns, and the degree of influence varies among different emotional states. This research provides investors with more accurate sentiment analysis tools to help them predict the trend of the stock market more scientifically and optimize investment decisions.

## Keywords

Investor Sentiment, Stock Income, Nonlinear Influence, Deep Learning, Text Mining

---

## 1. Introduction

As an important component of the financial market, the stock market's fluctuations are influenced by multiple factors. Among them, investor sentiment, as one

of the important factors affecting the fluctuations of the stock market, has received extensive attention in recent years (Ding et al., 2019). Investor sentiment reflects investors' optimistic or pessimistic attitude towards the market. This attitude will further affect investors' decision-making behavior, thereby influencing the trend of the stock market. However, traditional investor sentiment research is often based on limited data sources such as questionnaires or market indices, making it difficult to comprehensively and accurately reflect the true emotional state of investors (Tetlock, 2007).

As big data and artificial intelligence technologies have advanced at a rapid pace, deep learning and text mining techniques have offered fresh perspectives and innovative approaches for conducting research on investor sentiment (Bollen et al., 2011). Deep learning models possess the capability to process extensive and high-dimensional datasets, enabling them to discern intricate patterns and relationships within the data. Meanwhile, text mining technology is adept at extracting valuable insights from vast amounts of textual data, thereby furnishing abundant data resources for the analysis of investor sentiment (Antweiler & Frank, 2004). Therefore, this paper aims to utilize deep learning and text mining techniques to explore the nonlinear impact of investor sentiment on stock returns and provide investors with more in-depth sentiment analysis tools.

## 2. A Review of Research on the Relationship between Investor Sentiment and Stock Returns

Numerous research endeavors have delved into examining how investor sentiment influences stock returns. In the early stages, these studies predominantly relied on data sources like questionnaires or market indexes, and they uncovered a definite connection between investor sentiment and stock returns (Brown et al., 2004). Nevertheless, these investigations are frequently constrained by the data sources they employ and the methodologies they utilize, which hinders their ability to precisely portray the actual condition of investor sentiment as well as its underlying mechanism of impact on stock returns (Baker & Wurgler, 2006).

Over recent years, as big data and artificial intelligence technologies have advanced, an increasing number of studies have started to employ deep learning and text mining methods for the purpose of analyzing investor sentiment (Renault, 2017). These research efforts gather textual data from sources like social media platforms and news articles. They then employ natural language processing techniques to extract indicators of investor sentiment and analyze these indicators alongside stock market data. The empirical findings reveal that investor sentiment exerts a substantial influence on stock returns, and this influence frequently exhibits a nonlinear nature (Siganos et al., 2014). However, the existing studies still have certain limitations in aspects such as model construction, data sources and the extraction of sentiment indicators, which require further in-depth exploration (Pagolu et al., 2016).

### **3. The Theoretical and Data Basis of the Relationship between Investor Sentiment and Stock Returns**

#### **3.1. The Theoretical Cornerstone of the Relationship between Investor Sentiment and Stock Returns**

Behavioral finance incorporates psychological principles into financial studies, overcoming the constraints imposed by the “rational person” hypothesis in conventional finance. It posits that investors are not fully rational and are swayed by elements like cognitive distortions and emotional shifts during their decision-making process. Investor sentiment, as an important research content of behavioral finance, reflects investors’ optimistic or pessimistic attitude towards the market. This emotional state can distort investors’ judgment of stock value, causing stock prices to deviate from their intrinsic value. For instance, overconfident investors might overestimate the future returns of stocks, thereby driving up stock prices. However, overly pessimistic investors may underestimate the value of stocks, triggering a decline in stock prices. The theory of behavioral finance provides an important theoretical basis for understanding the impact of investor sentiment on stock returns and reveals the mechanism of investors’ irrational behavior in the stock market.

The market efficiency theory posits that within an efficient market, stock prices are capable of completely mirroring all accessible information. Nevertheless, the presence of investor sentiment might undermine the market’s efficiency. When investor sentiment tends to be broadly optimistic, a substantial influx of funds surges into the market, propelling stock prices higher. At this time, stock prices may be higher than their intrinsic value, and the market shows an overreaction. Conversely, when investors are pessimistic, a large amount of funds flows out, stock prices fall, and stock prices may be lower than their intrinsic value, resulting in an insufficient market response. The market inefficiency stemming from investor sentiment gives rise to an intricate connection between stock returns and sentiment, thus offering a theoretical foundation for exploring the nonlinear influence of investor sentiment on stock returns. The herd effect refers to the fact that when information is uncertain, investors tend to imitate others’ investment decisions instead of making investments based on their own independent analysis. In the stock market, when the sentiment of some investors changes, it may trigger the herd behavior of other investors, leading to the rapid spread and amplification of market sentiment. For instance, when some well-known investors or institutions show optimism about a certain stock and purchase it in large quantities, other investors may follow suit and buy it, driving up the stock price significantly. The herd effect makes the impact of investor sentiment on stock returns collective and contagious, further intensifying the volatility and nonlinearity of the stock market.

#### **3.2. Data Context regarding the Study on the Link between Investor Sentiment and Stock Returns**

As the Internet has become more widespread and social media has advanced, a

vast quantity of textual data has surfaced, furnishing abundant data resources for research into investor sentiment (Zhang et al., 2011). Text data, such as user comments and posts on social media platforms and financial news on news reporting websites, can reflect investors' views and sentiments towards the market in real time. These text data have the characteristics of strong timeliness, wide coverage and rich information. Compared with traditional data, they can better capture the subtle changes in investors' sentiments. Through the mining and analysis of text data, an in-depth understanding of investors' emotional states and their impact on the stock market can be achieved, providing a new perspective and method for studying the relationship between investors' emotions and stock returns.

Stock market data encompasses a wide range of elements, such as share prices, trading volumes, price-to-earnings ratios, and price-to-book ratios, among others. These data points are capable of mirroring the overall functioning of the stock market as well as the performance of individual stocks. Stock prices are the most direct indicator reflecting stock returns, while trading volume can reflect the market's activity level and investors' participation. Assessment metrics like the price-earnings ratio and the price-to-book ratio can assist investors in figuring out whether the stock price is fair or reasonable. Stock market data of different types are inter-related and jointly constitute an important data basis for studying the relationship between investor sentiment and stock returns. Through the comprehensive analysis of these data, a more comprehensive understanding of the influence mechanism of investor sentiment on stock returns can be achieved.

To thoroughly examine the connection between investor sentiment and stock returns, it is essential to combine and perform a correlation analysis on both text data and stock market data. For the purpose of analyzing how shifts in sentiment influence stock returns, the investor sentiment details within the text data must align with the return information present in the stock market data. During the process of data integration, issues such as the temporal synchronization of data and the consistency of data need to be considered. Correlation analysis can explore the potential relationship between investor sentiment and stock returns through statistical methods, machine learning algorithms, etc. For instance, the distribution characteristics of stock returns under different emotional states can be analyzed, or the correlation between emotional changes and fluctuations in stock returns can be explored. Through data integration and correlation analysis, the nonlinear impact of investor sentiment on stock returns can be revealed more accurately.

## **4. Extraction of Investor Sentiment Indicators and Theoretical Analysis**

### **4.1. Methods for Extracting Investor Sentiment Indicators**

An emotional dictionary is a kind of dictionary that contains both positive and negative words. By calculating the quantity and proportion of positive and negative words in the text, the emotional score of the text can be obtained. In investor sen-

timent analysis, a specialized financial sentiment dictionary can be constructed, including both positive and negative terms related to the stock market. Then, sentiment analysis is conducted on the preprocessed text data to calculate the sentiment score of each text, and thereby obtain the investor sentiment indicators at each time point. This approach is straightforward and easy to grasp, yet it might be constrained by the extent of coverage and the level of accuracy offered by the sentiment dictionary.

Machine learning methods can automatically identify the sentiment tendencies in text by training models. Frequently employed machine learning algorithms encompass Support Vector Machine (SVM), Naive Bayes, Random Forest, and the like. In the context of investor sentiment analysis, labeled textual data can be utilized to train machine learning models, empowering them to precisely differentiate between texts with positive, negative, and neutral tones. Subsequently, the trained model is applied to carry out sentiment classification on fresh textual data, and the percentage of each text type is computed to serve as an addition to the sentiment index. This method can automatically learn the emotional features in the text, with high accuracy and flexibility, but it requires a large amount of labeled data for model training.

Deep learning methods have achieved remarkable results in text sentiment analysis. Widely adopted deep learning models comprise Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), along with their variations (like LSTM, GRU), and so on. These models are capable of handling long sequence text data and capturing semantic information and emotional features in the text. In investor sentiment analysis, deep learning models can be used to perform sentiment classification or sentiment score prediction on text data to obtain more accurate sentiment indicators. Deep learning methods have strong learning ability and generalization ability, but they require a large amount of computing resources and time for model training.

#### **4.2. Theoretical Analysis of Investor Sentiment Indicators**

Through statistical analysis of the extracted investor sentiment indicators, one can understand the distribution characteristics, mean values, variances and other statistics of the sentiment indicators. These statistics can reflect the overall level and fluctuations of investor sentiment. For instance, a higher average value of sentiment indicators may indicate that the overall market sentiment is relatively optimistic, while a larger variance may suggest that market sentiment fluctuates more intensely.

Through the application of the correlation analysis approach, we can delve into the connection between investor sentiment indicators and stock returns. By computing the correlation coefficient between these sentiment indicators and stock returns, we can initially ascertain whether a linear relationship exists between the two. Nevertheless, given that the link between investor sentiment and stock returns might be nonlinear, correlation analysis merely offers a preliminary insight. Further

analysis can adopt nonlinear correlation coefficients or visualization methods, such as scatter plots and heat maps, to present the relationship between the two more intuitively.

Investor sentiment is dynamically changing and its changes may be influenced by multiple factors. By analyzing the dynamic changes of emotional indicators, one can understand the changing trends and cyclical characteristics of emotions. For example, the time series analysis method can be utilized to model the emotional indicators and predict the future changing trend of emotions. Additionally, we can analyze how sentiment indicators evolve across various market phases (like bull markets and bear markets) to investigate the connection between investor sentiment and market trends.

## **5. The Nonlinear Impact of Investor Sentiment on Stock Returns**

### **5.1. Manifestations of Nonlinear Influence**

Investor sentiment's influence on stock returns might exhibit a threshold effect. In other words, once the sentiment indicator hits a specific level, the impact of sentiment on stock returns will undergo a notable change. For instance, when market sentiment is extremely optimistic, investors may overbuy stocks, driving up stock prices significantly. When market sentiment is extremely pessimistic, investors may sell stocks excessively, causing a significant drop in stock prices. This threshold effect can be verified by setting different emotional thresholds and analyzing the impact of emotions on stock returns within different threshold intervals.

Shifts in investor sentiment might not be instantly mirrored in stock returns; instead, they could exhibit a certain delay or lag effect. That is, the current emotional state may have an impact on future stock returns. This lag effect might be due to the fact that the decision-making process of investors requires a certain amount of time, or the dissemination and digestion of market information also takes a certain amount of time. By scrutinizing the time-series connection between sentiment indicators and stock returns, we can delve into whether a lag effect exists and gauge its intensity.

The influence of investor sentiment on stock returns might exhibit an asymmetric impact. In other words, the degree to which positive (optimistic) and negative (pessimistic) sentiments impact stock returns could vary. For instance, pessimism may have a greater negative impact on stock returns because investors are more likely to engage in panic selling when they are pessimistic. The existence of the asymmetric effect can be verified by comparing the changes in stock returns under optimistic and pessimistic sentiments.

### **5.2. Theoretical Explanation of Nonlinear Influence**

The psychological factors of investors are one of the important reasons that lead to the nonlinear impact of investor sentiment on stock returns. Psychological fac-

tors such as investors' cognitive biases and emotional fluctuations can affect their interpretation of market information and decision-making behaviors. For example, investors may be overly confident and ignore risks when they are optimistic. When pessimistic, one may be overly fearful and magnify the risks. The effect of this psychological factor can lead to the influence of emotions on stock returns, showing nonlinear characteristics.

The operating mechanism of the stock market can also affect the nonlinear impact of investor sentiment on stock returns. For instance, factors such as market liquidity and transaction costs can influence investors' trading behaviors. When market liquidity is poor, changes in investors' sentiment may be more likely to cause significant fluctuations in stock prices. When transaction costs are high, changes in investors' sentiment may have a relatively small impact on stock returns. In addition, market regulatory policies, information disclosure systems, etc. will also have an impact on investors' emotions and behaviors, and thereby affect stock returns.

External disturbances, like alterations in the macroeconomic landscape, policy modifications, and unforeseen occurrences, can exert a substantial influence on investor sentiment and the stock market. These external factors might cause abrupt shifts in investor sentiment, subsequently prompting irrational volatility in the stock market.

For example, when an economic crisis erupts, it can result in highly pessimistic investor sentiment and spark a stock market plunge. Conversely, the implementation of beneficial policies could foster optimistic investor sentiment and propel stock prices higher.

The impact of these external shocks will amplify the nonlinear effect that investor sentiment has on stock returns.

### **5.3. Implications of Nonlinear Influence for Investment Decision-Making**

Investors should recognize the nonlinear impact of investor sentiment on stock returns and enhance their risk awareness. When market sentiment is extremely optimistic, be vigilant against possible market bubbles and avoid excessive investment. When market sentiment is extremely pessimistic, one should remain calm and avoid blind selling. At the same time, the investment portfolio should be reasonably adjusted according to the changes in sentiment indicators to reduce investment risks.

Due to the nonlinear and uncertain impact of investor sentiment on stock returns, short-term speculative behavior often faces greater risks. Therefore, investors should establish the concept of long-term investment, pay attention to the fundamentals and long-term value of stocks, and avoid being affected by short-term emotional fluctuations. Through long-term investment, the impact of emotional fluctuations on investment returns can be smoothed out, achieving stable appreciation of assets.

Diversified investment is one of the effective strategies to reduce investment risks. Investors can diversify the impact of investor sentiment on a single stock or industry by investing in stocks of different industries and regions, as well as allocating to other asset classes (such as bonds, funds, etc.). Diversified investment can reduce the overall risk of the investment portfolio and enhance the stability of investment returns.

## 6. Conclusion

This research carries out a comprehensive examination of the nonlinear influence of investor sentiment on stock returns, leveraging deep learning and text mining approaches. Behavioral finance, market efficiency theory, and herd effect theory collectively offer a robust theoretical foundation for investigating the connection between investor sentiment and stock returns. Behavioral finance elucidates how investors' irrational behaviors shape their judgments of stock value and contribute to price fluctuations. Market efficiency theory posits that investor sentiment can potentially give rise to market inefficiency, thereby establishing an intricate relationship between stock returns and sentiment. The herd effect underscores the collective and contagious nature of investor sentiment, exacerbating the volatility and nonlinearity of the stock market. These theories are mutually complementary and together establish a theoretical framework for comprehending the nonlinear impact of investor sentiment on stock returns.

With the proliferation of the Internet and social media, text data has emerged as a rich and up-to-date source for researching investor sentiment. Stock market data encompasses various dimensions, including stock prices, trading volumes, price-earnings ratios, and price-to-book ratios, offering a holistic view of the stock market's operations. By integrating and conducting correlation analysis on text data and stock market data, we can more precisely uncover the nonlinear impact of investor sentiment on stock returns. When it comes to extracting investor sentiment indicators, the sentiment dictionary method is straightforward and easy to understand. However, its effectiveness is constrained by the dictionary's coverage and accuracy. Machine learning methods are capable of automatically learning emotional features, boasting high accuracy and flexibility, yet they demand a substantial amount of labeled data. Deep learning methods exhibit strong learning and generalization capabilities, but they necessitate significant computing resources and time for model training. Each extraction method has its own strengths and weaknesses, and the appropriate method can be chosen based on research requirements and data availability.

This study, by employing deep learning and text mining techniques, delves deeply into the nonlinear impact of investor sentiment on stock returns, equipping investors with more sophisticated sentiment analysis tools and a scientific basis for making investment decisions. Nevertheless, this study is not without limitations, such as constraints related to data sources and the complexity of model construction. Future research endeavors can further broaden data sources, refine the model struc-

ture, enhance the accuracy and reliability of the research, and offer more valuable insights for investors.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, *59*, 1259-1294. <https://doi.org/10.1111/j.1540-6261.2004.00662.x>
- Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*, 1645-1680. <https://doi.org/10.1111/j.1540-6261.2006.00885.x>
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science*, *2*, 1-8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Brown, G. W., & Cliff, M. T. (2004). Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance*, *11*, 1-27. <https://doi.org/10.1016/j.jempfin.2002.12.001>
- Ding, W., Mazouz, K., & Wang, Q. (2019). Investor Sentiment and the Cross-Section of Stock Returns: New Theory and Evidence. *Review of Quantitative Finance and Accounting*, *53*, 493-525. <https://doi.org/10.1007/s11156-018-0756-z>
- Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2016). Sentiment Analysis of Twitter Data for Predicting Stock Market Movements. In *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPEs)* (pp. 1345-1350). IEEE. <https://doi.org/10.1109/scopes.2016.7955659>
- Renault, T. (2017). Intraday Online Investor Sentiment and Return Patterns in the U.S. Stock Market. *Journal of Banking & Finance*, *84*, 25-40. <https://doi.org/10.1016/j.jbankfin.2017.07.002>
- Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's Daily Sentiment and International Stock Markets. *Journal of Economic Behavior & Organization*, *107*, 730-743. <https://doi.org/10.1016/j.jebo.2014.06.004>
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*, 1139-1168. <https://doi.org/10.1111/j.1540-6261.2007.01232.x>
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting Stock Market Indicators through Twitter "I Hope It Is Not as Bad as I Fear". *Procedia-Social and Behavioral Sciences*, *26*, 55-62. <https://doi.org/10.1016/j.sbspro.2011.10.562>