Journal of Management and Informatics (JMI)
Vol. 4 No. 1 April 2025
E-ISSN: 2961-7731; P-ISSN: 2961-7472, Pages 741-756
Publication Link: https://jmi.stekom.ac.id/index.php/jmi/issue/view/10

AI-Driven Sentiment Analysis of Retail Investor Behavior during Market Volatility: A Study of Twitter Data in Southeast Asia

Sutriani Dewi Sriasih*[1], Farhat Abdul Razak[1], Hussein al Ikhsan[1]
Email: dewisria@unismuh.ac.id; farhat.abd.ra@gmail.com; hiialikhsan@gmail.com
[1]*Universitas Muhammadiyah Makassar, Indonesia*
*Corresponding Author

Abstract
In recent years, retail investor participation in Southeast Asian capital markets has surged, contributing to increased market volatility and making sentiment analysis a critical factor in understanding price dynamics. This study investigates the relationship between social media sentiment and stock market fluctuations by focusing on Twitter data during periods of market volatility in Indonesia, Thailand, and Malaysia. The objective is to examine how collective investor emotions, as expressed through social media, correlate with daily stock index movements. Employing an exploratory quantitative approach, the study integrates Natural Language Processing (NLP) methods, both lexicon-based tools such as VADER and advanced transformer-based models like BERT and GPT, to classify over 150,000 tweets into positive, negative, and neutral sentiments. Sentiment scores were then aggregated and statistically tested using Pearson correlation with daily stock index returns, specifically the IDX Composite, SET Index, and FTSE Bursa Malaysia. The findings reveal a significant negative correlation between negative sentiment and market returns, particularly in the IDX Composite ($r = -0.61$, $p < 0.05$), indicating that pessimistic sentiment is associated with market downturns. Thailand's SET Index and Malaysia's FTSE Index showed moderate to weak negative correlations, with $r = -0.43$ and $r = -0.27$, respectively. These results highlight the sensitivity of emerging markets to emotionally driven retail behavior. The study concludes that AI-based sentiment analysis offers a valuable early warning tool for market volatility and can complement traditional financial indicators. It recommends developing AI-based sentiment dashboards and enhancing digital financial literacy to mitigate emotional reactivity among retail investors.

Keywords: *AI Sentiment, Retail Investors, Market Volatility, Twitter Data, Stock Index.*
Received on February2025; Revised on March 2025; Accepted on March 2025; Published on April 2025

I. INTRODUCTION

Over the past decade, Southeast Asia's capital markets have undergone a significant transformation, one of which is marked by the rising participation of retail investors. In Indonesia, for instance, the number of retail investors has grown by more than 90% since 2020, according to data from the Indonesian Central Securities Depository (KSEI). A similar trend has been observed in Thailand and Malaysia, where individual investors have emerged as dominant actors in daily trading volumes. This increased retail participation has major implications for market dynamics, including sharper price fluctuations and more volatile investor sentiment. The surge in participation also reflects a shift in market structure from one previously dominated by large institutions to a more open environment accessible to the broader public through digital platforms. At the same time, the emergence of social media platforms such as Twitter (now X) as venues for sharing financial opinions has introduced "digital emotion" as a new force in price formation.

DOI: 10.51903/jmi.v4i1.179

Twitter enables the rapid dissemination of information and opinions, fostering an open discussion ecosystem that often shapes collective perceptions of market conditions. Investors are now influenced not only by fundamental or technical data but also by narratives, opinions, and social signals that evolve in real time across digital spaces.

Twitter/X, in particular, has become a massive financial discussion arena, where hashtags such as #IHSG, #LQ45, #Crypto, and #SahamIndonesia capture conversations among thousands of users each day. On this platform, investors share opinions, predictions, and emotional responses to market events, generating a dynamic and often unstructured flow of information. This information flow reflects the pattern of financial sentiment flow, a stream of market perceptions shaped by opinions, concerns, and collective expectations that spread rapidly and widely among user communities. According to (Albahli et al., 2022), social media platforms such as Twitter possess significant potential in conveying real-time indicators of investor behavior, although most applications to date have remained focused on the consumer sector and have not been extensively explored within the context of Southeast Asian capital markets. Another study by (Maqsood et al., 2022) adds that collective sentiment expressed on social media has a statistically significant relationship with stock market volatility, particularly during periods of high uncertainty, though their research primarily focuses on developed markets. These findings underscore the urgent need to develop new approaches based on Artificial Intelligence (AI), such as sentiment modeling using NLP techniques and Large Language Models (LLMs), to capture the dynamics of digital emotion that frequently influence retail investment decisions, especially within the highly volatile contexts of emerging markets.

Although previous studies have demonstrated the importance of sentiment analysis in understanding market dynamics, most have remained focused on transactional data and traditional market indicators. For example, (Mendoza-Urdiales et al., 2022) found that sentiment on Twitter could predict stock market trends in the United States; however, their study was limited to developed markets and did not consider the context of retail investors in emerging economies. Similarly, (Figà-Talamanca & Patacca, 2022) found that negative emotions on social media contributed to increased volatility in the S&P 500 index, but their methodology did not incorporate more advanced LLM-based technologies. On the other hand, (Aziz et al., 2022), who examined the Indonesian market, revealed that the correlation between sentiment and stock indices only became significant during crisis periods, and a comprehensive analysis of market volatility in general has yet to be conducted. Other scholars, such as (Li et al., 2023), focused on the interaction between investor behavior and price fluctuations but overlooked the role of social media opinion as one of the key drivers of these dynamics. In addition, research by (Kirtac & Germano, 2024) highlights the potential of AI in modeling market sentiment, though their study

is limited to lexicon-based approaches and does not delve deeply into the use of LLMs such as GPT or BERT in the context of emerging markets. Consequently, this study aims to fill these research gaps by analyzing retail investor sentiment on Twitter/X using AI-driven sentiment analysis and examining its correlation with fluctuations in major stock indices in Southeast Asia, particularly in Indonesia, Thailand, and Malaysia.

Based on the aforementioned background and research gaps, this study is designed to identify and analyze retail investor sentiment as expressed on Twitter during periods of market volatility in the Southeast Asian region. The primary focus is on how perceptions, emotions, and reactions of individual investors emerge and evolve in digital spaces and how these dynamics relate to movements in key stock indices such as the IDX Composite, SET Index, and FTSE Malaysia. By employing an AI-based sentiment analysis approach, this study seeks to construct a model capable of capturing daily sentiment fluctuations and quantifying them systematically. One of the key questions it aims to address is whether there exists a significant correlation between positive or negative investor sentiment and market index changes during unstable periods. The study also seeks to test the hypothesis that investor sentiment gathered from social media can serve as a leading indicator of market movements, particularly in the context of emerging markets that are more susceptible to emotional responses. It is hoped that the findings of this study will contribute to the development of AI-based market monitoring models and assist market participants and regulators in understanding the potential psychological impact of retail investors on stock price volatility.

II. LITERATURE REVIEW

A. Theoretical Framework

1. Retail Investor Behavior from a Behavioral Finance Perspective

In the study of behavioral finance, retail investor behavior often diverges from the rationality assumptions that underpin classical financial theory. Individual investors tend to be influenced by psychological biases when faced with market uncertainty, including overconfidence bias, herding behavior, and heuristic-based decision-making. According to (Addo et al., 2025), retail investors consistently exhibit a tendency toward excessive trading, which ultimately undermines the performance of their investment portfolios. They argue that such decisions are frequently driven by an inflated confidence in their own ability to predict market trends. This phenomenon stands in stark contrast to the institutional approach, which tends to be more data-driven and risk-oriented. Understanding the psychological characteristics of retail investors is thus essential for explaining market behavior that does not always conform to rational expectations.

One of the foundational theories in behavioral finance is Prospect Theory, introduced by Kahneman and Tversky, which posits that individuals make decisions not based on final outcomes but rather on perceived gains and losses relative to a reference point. This theory highlights the concept of loss aversion, or the tendency for individuals to perceive losses as more painful than gains of equivalent value. Research by (Bhattacharya & Sardashti, 2022) reinforces this notion, showing that retail investors are more reactive to market losses, often resulting in panic selling when markets display negative signals. This effect is further exacerbated under conditions of high volatility, when emotional pressure intensifies and rational judgment declines. In such contexts, the rapid circulation of information on social media has the potential to accelerate the collective perception of risk. These circumstances reveal a strong linkage between the psychological mechanisms of investors and an ever-evolving informational environment.

In addition to loss aversion, retail investors are also prone to overreaction bias, which refers to exaggerated responses to particular news or market events. According to (Yang, 2024), markets often experience a mean reversion effect following such overreactions, indicating that stock prices do not immediately reflect intrinsic values due to emotionally driven investor behavior. The researcher also notes that individual investors are more susceptible to this bias due to limitations in information access and investment experience. Overreaction is not limited to negative news; it also applies to speculative or sensational information, often amplified by discourse on digital platforms such as Twitter. In such cases, rapidly evolving public opinion plays a key role in shaping collective perceptions of both risk and potential gains. Emotional responses to specific market stimuli can unfold swiftly and disseminate widely through social networks.

Moreover, studies on herding behavior suggest that retail investors often follow the majority without conducting independent analysis. Research by (Sin Huei et al., 2022) found that herding behavior is more prevalent in emerging markets, including those in Asia, where financial literacy and equitable access to information remain significant challenges. They observed that during periods of market instability, retail investors tend to rely on social signals and follow prevailing public opinion rather than evaluating fundamental data. This pattern illustrates how investment decisions are frequently shaped by social pressure rather than objective analysis. In the context of social media, herding behavior may be further reinforced by platform algorithms that amplify exposure to popular content. This scenario underscores how the dissemination of opinions in the digital sphere can accelerate the formation of expectations and collective actions among retail investors.

2. Digital Sentiment Analysis and Artificial Intelligence-Based Approaches

Digital sentiment analysis is a branch of NLP that aims to identify and classify emotions or opinions embedded in text. This method has been widely applied to analyze unstructured data from social media platforms, product reviews, and online forums. According to (Velu et al., 2023) and (Catelli et al., 2022), early sentiment analysis approaches largely relied on lexicon-based techniques and supervised learning methods such as Naive Bayes or Support Vector Machines, which assess word polarity either individually or based on their frequency. These models typically depend on curated lists of positive and negative terms, as well as manually annotated datasets to train classification systems. While effective for explicit or direct expressions of sentiment, these approaches encounter limitations when applied to more complex contexts, such as irony or ambiguity. Such limitations reflect the early developmental stage of sentiment analysis before the advent of more context-aware AI models.

With the advancement of NLP technologies, deep learning-based approaches have increasingly been adopted for sentiment analysis, particularly to address the challenges posed by the unstructured and stylistically diverse nature of social media data. (Abimbola et al., 2024) note that models such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks possess the capability to retain contextual information from word sequences, thereby enhancing the interpretation of implicit sentiments. These models rely on feature representations learned automatically from large datasets, without the need for predefined dictionaries or hand-crafted rules. The layered structure of deep networks enables the detection of complex patterns in data, including shifts in tone or emotion across evolving discourse. In practical applications, these models can process both long-form narratives and short social interactions such as tweets, making them particularly suited to analyzing dynamic emotional expressions in digital communication.

More recent progress in NLP is marked by the emergence of transformer-based models such as BERT, which was developed to capture sentence-level context more comprehensively. As explained by (Bello et al., 2023) and (Alshattnawi et al., 2024), BERT employs a bidirectional contextual learning approach, allowing it to interpret the meaning of words based on both preceding and succeeding context. This approach yields richer language representations by considering the full sentence structure rather than merely focusing on local word relationships. Trained on vast and diverse datasets, BERT is capable of handling informal linguistic styles and social expressions frequently found on social media. In practice, BERT delivers superior performance across various NLP tasks, including the classification of complex sentiment. Its application to Twitter and other social media data enables a more accurate large-scale evaluation of digital emotions.

In addition to BERT, the Generative Pre-trained Transformer (GPT) model also demonstrates high proficiency in understanding and generating natural language, including in sentiment analysis tasks. According to (Oralbekova et al., 2023), GPT utilizes an autoregressive approach that learns language patterns through sequential token prediction, thereby enabling it to recognize sentence structure and meaning over extended contexts. This capacity makes GPT particularly effective in handling the variety of expressions and emotional nuances commonly encountered in online communication. GPT also supports zero-shot and few-shot learning approaches, which allow sentiment analysis to be performed without specialized training for each new dataset. In the context of social media data, this flexibility facilitates the identification of emotions and opinions across large volumes of text without the need for extensive manual adjustment. GPT's strength in detecting implicit meaning, including sarcasm and other indirect forms of expression, positions it as a critical tool for the study of retail investor behavior through digital sentiment analysis.

B. Prior Research

1. The Relationship Between Twitter Sentiment and Stock Returns Across Various Markets

Social media, particularly Twitter, has emerged as an alternative source of financial information that reflects investors' opinions and emotions in real time. According to (Dumiter et al., 2023), public sentiment gathered from Twitter can be utilized to predict the directional movement of stock market indices such as the Dow Jones Industrial Average (DJIA). The researchers employed NLP techniques to analyze approximately 10 million tweets and found that public emotional states such as "anxiety" and "calmness" were significantly correlated with market fluctuations occurring two to six days later. These findings suggest that social media data are not merely reflective in nature but may also possess predictive power in relation to market behavior. The analysis demonstrates how collective sentiment expressed on Twitter can provide relevant signals for understanding the dynamics of the stock market. This pattern illustrates the critical role of collective emotions in shaping expectations and risk perceptions within the financial market context.

A subsequent study by (Koukaras et al., 2022) extended this analysis by examining the correlation between tweet volume and stock price volatility of individual firms. In this research, tweets related to companies listed on the S&P 500 were collected and analyzed to assess changes in sentiment and corresponding market activity. The results revealed that spikes in the volume of tweets containing negative sentiment frequently coincided with increased price volatility, although they did not always lead to consistent changes in return direction. The researchers also noted that speculative or emotionally charged tweets exerted a stronger influence than neutral,

informational ones. Social interactions and opinion dissemination through social media appeared to generate psychological dynamics that significantly affected investment decisions. These findings indicate that Twitter may serve as a mirror of market sentiment not always captured by conventional economic indicators.

In the context of emerging markets, a study by (Sinaga et al., 2022) highlighted the use of Twitter data to observe retail investors' reactions to economic news and stock market movements in Indonesia. The researchers identified that during crisis periods, such as the COVID-19 pandemic, negative sentiment on Twitter rose significantly, followed by large-scale sell-offs in the domestic capital market. This study affirms that broadly expressed negative sentiment on social media can trigger collective reactions that intensify short-term market volatility. Moreover, it was found that the frequency of tweets with heightened emotional content often peaked during periods of extreme market stress, reflecting the psychological tension experienced by retail investors. In this context, Twitter functions as a space where uncertainty and concern are articulated elements that may not always be evident in financial statements or fundamental analyses. The findings underscore that social media has become a crucial element for understanding local market dynamics amid global financial pressures.

Other scholars, such as (Liu et al., 2023), have explored the relationship between social media sentiment and stock returns using more advanced machine learning approaches. In this study, learning algorithms were employed to classify the sentiment of tweets into positive, negative, and neutral categories and to link these classifications with the price movements of technology stocks in the United States. The researchers found that positive sentiment was consistently associated with short-term stock price increases, while negative sentiment tended to correlate with intraday declines in stock value. The predictive models developed in this research demonstrated improved accuracy compared to conventional methods, primarily because the algorithms were capable of detecting complex linguistic patterns. This study highlights the potential of AI-based approaches to enhance the effectiveness of sentiment analysis in financial contexts. Twitter data, in this regard, offers broad access to investor emotions that can be leveraged to understand the micro-level dynamics of stock price movements.

2. Modeling Investor Behavior Using AI in the Context of Market Volatility

AI has emerged as an effective approach for understanding investor behavior, particularly during periods of market volatility. According to (Rezaei et al., 2025), financial news and market narratives exert a significant influence on stock price movements, and such data can be quantitatively modeled to represent shifts in collective behavior. The researchers employed text analysis of economic news columns to assess market sentiment and link it to fluctuations in the

capital market. Their findings indicate that market perceptions shaped by the media contribute to widespread buying or selling behavior, even when the underlying information remains ambiguous. The models developed from textual data enable systems to recognize sentiment patterns that often precede price changes. The application of AI in this context demonstrates its ability to filter out public opinion "noise" and extract statistically relevant behavioral signals.

Another study by (Chhajer et al., 2022) elaborates on the use of various AI approaches, such as Support Vector Machines and Decision Trees, to predict stock market behavior based on data from social media and online news. The models constructed in this research integrate textual data with historical price variables to map sentiment tendencies in relation to market direction. The researchers highlight that the integration of public opinion data with predictive algorithms enhances the understanding of short-term market fluctuations, particularly during unstable conditions. In times characterized by uncertainty, investors' emotional responses tend to intensify and become more fragmented, requiring AI systems to be trained on data that reflect such contextual complexity. Sentiments formed on digital platforms such as Twitter offer a dynamic reflection of the market's psychological state. Ultimately, the use of AI-based predictive models allows for the identification of correlations between sentiment shifts and nonlinear price behavior.

(Tran et al., 2023) argue that employing deep learning models to capture investor behavior during market crises yields more representative results compared to conventional statistical methods. They utilized neural network architectures to evaluate investor responses to negative news and discovered that spikes in negative emotions often coincided with increased volatility in major indices. In this context, AI is positioned not merely as a sentiment classification tool but as a system capable of anticipating shifts in market dynamics based on historical behavioral patterns. The researchers noted that in fast-changing environments, highly flexible models such as LSTM are more effective at capturing nonlinear patterns that reflect emotional market fluctuations. Because market responses to uncertainty are not always rational, interpreting investor behavior through digital signals requires an approach that is sensitive to both temporal and emotional contexts. The structure of such models strengthens the argument that AI can be utilized to understand behavioral dimensions that are not visible through purely quantitative data.

Furthermore, a study by (Liapis & Kotsiantis, 2023) investigated how transformer-based models, such as BERT, are employed to classify and predict market sentiment from social media data during periods of heightened volatility. The researchers emphasized that attention-based models like BERT are particularly adept at capturing sentence-level context, especially within the informal language often used by retail investors on platforms like Twitter. The sentiment classification results were then linked to volatility indices, such as the VIX, to measure the

intensity of the relationship between market emotions and pricing pressures. This study observed that rising negative sentiment on social media frequently preceded sharp directional changes in the market, suggesting a link between collective emotional responses and market dynamics. BERT was able to identify linguistic nuances such as irony and uncertainty, which are commonly present in public financial discourse. This approach reflects the potential of AI to map the relationship between digital perceptions and price behavior in situations of market instability. A comparative summary of the methodologies and key findings from these studies is presented in Table 1 to provide an overview of the research approaches and insights.

Table 1. Comparison of Previous Studies on Modeling Investor Behavior Using AI in the Context of Market Volatility

Researchers	Data Source	Research Focus	Key Findings
(Rezaei et al., 2025)	Economic news	Market returns in relation to news sentiment	Public emotions expressed in news influence stock price movements
(Chhajer et al., 2022)	Social media and online news	Market prediction based on the combination of public opinion and historical price data	Integrating sentiment and price data improves predictive accuracy.
(Tran et al., 2023)	Investor reactions to negative news	Identification of emotional patterns during market crises	Negative emotions are associated with spikes in index volatility
(Liapis & Kotsiantis, 2023)	Twitter + VIX	Market sentiment classification and its relationship with volatility indices	Increased negative sentiment often precedes sharp market shifts.

III. RESEARCH METHOD

This study adopts an exploratory quantitative approach aimed at identifying patterns and relationships between retail investor sentiment on the social media platform Twitter and fluctuations in major stock market indices across Southeast Asia. This approach enables the researcher to investigate phenomena that have not been extensively examined through quantitative methods, particularly regarding the role of digital emotions in influencing market behavior. The exploratory method is deemed appropriate, as the interaction between social media sentiment and stock markets remains complex and not fully delineated within the traditional financial literature. In the context of emerging markets, especially in Southeast Asia, limitations in digital literacy and data infrastructure also shape the characteristics of investor sentiment expressed online. Therefore, this analysis is not only descriptive in nature but also seeks to capture nonlinear patterns that may not be observable through conventional data-based studies. The approach underscores the importance of contextual understanding of retail investor behavior, particularly during periods of market instability.

The primary data in this study are drawn from two sources, each representing the sentiment and market dimensions in quantitative terms. The first source comprises sentiment data from social media, specifically public tweets from the Twitter/X platform, which were collected based on relevant hashtags such as #IHSG, #IDX, #cryptoID, #sahamindonesia, and #LQ45. These hashtags were selected because they were actively used by the retail investor community to discuss market issues, blue-chip stocks, and digital assets during the observation period. Data collection was conducted over a three-month period characterized by high market volatility, thereby allowing for the capture of a wide range of emotional expressions. The second source consists of stock price index data from three Southeast Asian countries: the IDX Composite (Indonesia), the SET Index (Thailand), and the FTSE Malaysia Index (Malaysia), all retrieved from reputable public sources such as Yahoo Finance and TradingView. These two types of data were then combined to form an analytical base that enables the statistical testing of relationships between sentiment variables and market index variables. A detailed description of the data elements used in this study is presented in Table 2, which outlines the data volume, types of hashtags, and coverage of the indices analyzed.

The analytical process in this study involves a series of technical steps designed to transform raw data into information that can be statistically examined. The initial phase involves data retrieval or web scraping of public tweets from the Twitter/X platform, using tools such as Twint and snscrape. These tools were chosen due to their ability to access historical data without the quantity limitations imposed by Twitter's official API. Data collection was tailored to specific keywords, namely hashtags frequently used by the retail investor community, in order to ensure content relevance to capital market contexts. Once the data had been collected, the next step was preprocessing, which entails transforming raw textual content into a more structured format. Preprocessing steps included the removal of symbols and hyperlinks, tokenization, stopwords removal, and stemming, returning words to their root forms, thus rendering the data more consistent for subsequent sentiment analysis.

After the data were cleaned and prepared, the next phase involved the sentiment classification of the tweet content using two main approaches. The first was lexicon-based NLP, employing tools such as VADER and TextBlob to assign sentiment scores based on the polarity of words within sentences. This method operates by referencing a predefined list of words weighted for positive or negative sentiment and calculating a final score based on the composition of words in each tweet. The second approach involved classification using LLMs such as GPT and BERT, which are capable of understanding context at a deeper level and detecting complex emotional expressions, including sarcasm or ambiguity. These LLMs were applied using a zero-shot classification scheme, enabling the models to classify sentiment without requiring specific

training on the dataset in question. Once each tweet had been assigned a sentiment score, all scores for a single day were aggregated into a daily sentiment score, representing the prevailing emotional tendency of investors for that day. These daily scores were then temporally aligned with the corresponding stock market index values to facilitate correlation testing between sentiment dynamics and market fluctuations during the same period.

The computational model employed in this study is designed to measure and analyze the relationship between retail investor sentiment expressed on social media and the movement of stock market indices. The calculation begins with the aggregation of daily sentiment scores derived from the classification results of all tweets on each observation day. This score is computed as the average of all sentiment values generated by tweets on a given day, as formulated in Equation (1):

$$S_t = \frac{1}{n}\sum_{i=1}^{n} Sentiment(tweet_i) \quad (1)$$

In this equation, St represents the average sentiment score on day t, while n denotes the total number of tweets analyzed on the same day. The sentiment value of each tweet is determined based on classification results obtained through two approaches: lexicon-based methods and LLM-based classification. This aggregation enables the construction of a daily sentiment indicator that can be used to analyze collective emotional tendencies of investors in response to market conditions. The outcome of this process is a sentiment time series, which is then aligned with daily stock price index data for subsequent statistical analysis.

To examine the linear relationship between investor sentiment and stock index values, the Pearson Correlation Coefficient is employed. This metric assesses the strength and direction of association between two numerical variables by comparing the deviations of daily sentiment scores from their mean with the deviations of index values from their mean, as shown in Equation (2):

$$r = \frac{\sum(S_t - \bar{S})(I_t - \bar{I})}{\sqrt{\sum(S_t - \bar{S})^2 \sum(I_t - I)^2}} \quad (2)$$

In this formula, St denotes the daily sentiment score, and It refers to the stock index value on day t, while $\bar{S}$ and $\bar{I}$ are the respective means of the sentiment and index datasets over the entire analysis period. The resulting Pearson correlation coefficient yields a value between -1 and 1, indicating whether the relationship between the two variables is positive, negative, or linearly uncorrelated. By employing this method, the study seeks to observe the extent to which emotional dynamics on social media interact with market fluctuations in a quantitative and structured manner.

IV. RESULT

A. Results

The sentiment analysis conducted during the observation period reveals that negative sentiment consistently dominated retail investor behavior, particularly during episodes of financial market turbulence. This pattern was especially evident when markets came under pressure due to external dynamics such as central bank interest rate policies, fluctuations in global commodity prices, and heightened geopolitical tensions. These conditions coincided with various global and regional economic issues, including reports on inflation, uncertainty surrounding monetary policy, and fears of a global recession. The presence of such factors tended to amplify pessimism among retail investors, as reflected in both the social media discourse and more cautious trading volumes. This indicates that investor sentiment is strongly influenced by broad macroeconomic dynamics, with risk perception playing a critical role in investment decision-making. In this context, sentiment analysis serves as a vital tool for capturing real-time market perceptions of evolving economic information.

To illustrate the relationship between sentiment and stock market performance, Figure 1 presents the daily trend of sentiment scores and stock indices throughout January 2025. This visualization aids in identifying shifts in sentiment that appear to align with daily stock index fluctuations. The movement patterns in the graph demonstrate a close correlation between sentiment fluctuations and stock index changes, with declines in sentiment scores often coinciding with downturns in market indices. These declines frequently occurred in the aftermath of negative economic news, which subsequently shaped collective investor perceptions. As sentiment worsens, investors tend to engage in sell-offs, exerting downward pressure on stock prices. The graph provides visual evidence that non-financial information, such as emotions and opinions expressed on social media, holds potential as an early indicator of market price movements.

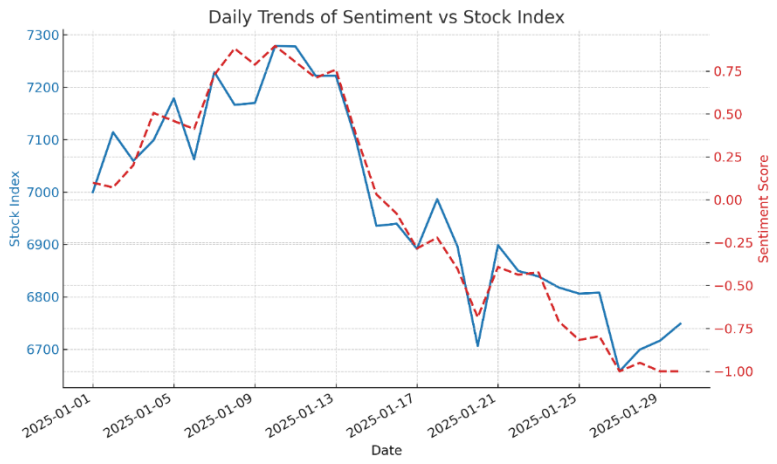

Figure 1. Daily Trend of Sentiment and Stock Index

Journal of Management and Informatics (JMI)
Vol. 4 No. 1 April 2025
E-ISSN: 2961-7731; P-ISSN: 2961-7472, Pages 741-756

Subsequently, a quantitative analysis was conducted to assess the strength of the relationship between negative sentiment and daily stock market returns. Correlation techniques were employed to determine the extent to which sentiment movements could explain variability in market performance. The results of the correlation analysis showed a coefficient of r = -0.61 for the IDX Composite index, indicating a strong negative relationship between the two variables. This finding implies that when market perception tends toward negativity, stock indices are more likely to experience a decline in value. The correlation is statistically significant ($p < 0.05$), suggesting that rising pessimism among retail investors is associated with deteriorating daily index performance during periods of crisis. These results also reflect the market's sensitivity to widely disseminated public opinion, particularly in the context of heightened economic uncertainty.

Gambar 2 menunjukkan sebaran antara skor sentimen dan return harian, di mana distribusi titik-titik data memperlihatkan kecenderungan hubungan negatif. Pola ini menunjukkan bahwa semakin rendah skor sentimen, semakin besar kemungkinan return harian mengalami penurunan. Sebaran data yang terkonsentrasi pada kuadran dengan nilai sentimen rendah dan return negatif memperkuat indikasi adanya keterkaitan yang sistematis antara persepsi pasar dan performa saham harian. Meskipun terdapat beberapa outlier yang menyimpang dari pola umum, mayoritas data mendukung hipotesis bahwa sentimen pesimis cenderung berasosiasi dengan kondisi pasar yang menurun. Kondisi ini menegaskan bahwa sentimen investor yang terekam dalam media sosial dapat mencerminkan ekspektasi pasar yang secara langsung memengaruhi keputusan transaksi harian. Visualisasi ini memberikan landasan awal untuk menguji hubungan statistik yang lebih kuat antara kedua variabel tersebut secara kuantitatif.

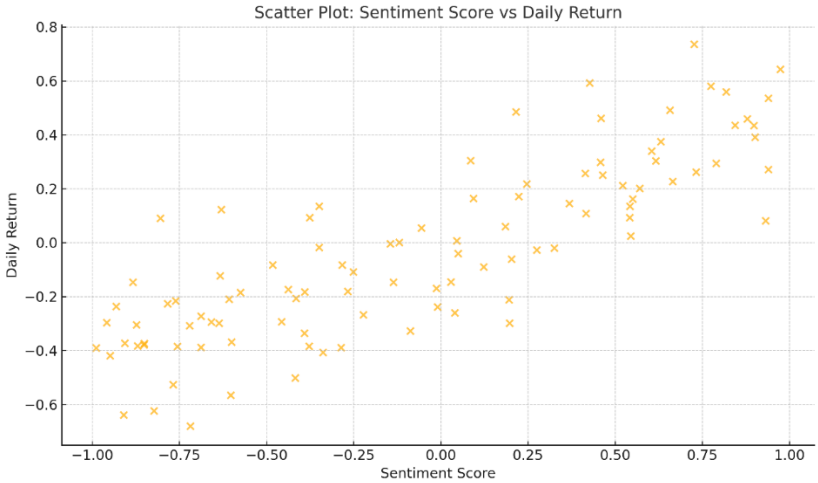

Figure 2. Scatter Plot: Sentiment Score vs Daily Return

Figure 2 illustrates the distribution between sentiment scores and daily returns, where the spread of data points exhibits a tendency toward a negative relationship. This pattern indicates that lower sentiment scores are generally associated with a greater likelihood of negative daily returns. The clustering of data in the quadrant representing low sentiment values and negative returns further reinforces the indication of a systematic association between market perception and daily stock performance. Although several outliers deviate from the general pattern, the majority of data supports the hypothesis that pessimistic sentiment tends to correlate with declining market conditions. This observation underscores the notion that investor sentiment captured on social media platforms may reflect market expectations that directly influence daily trading decisions. This visualization provides an initial foundation for quantitatively testing a more robust statistical relationship between the two variables.

Table 2. Correlation between Negative Sentiment and Stock Index Movements

Stock Index	Correlation (r) with Negative Sentiment	Remarks
IDX Composite	-0.61	Significant, $p < 0.05$
SET Thailand	-0.43	Moderately significant
FTSE Malaysia	-0.27	Weak correlation

These findings indicate that negative sentiment is meaningfully associated with market declines in several countries, although the strength of this relationship varies. In the case of Indonesia, the IDX Composite demonstrated the highest level of sensitivity to fluctuations in retail investor sentiment, as evidenced by the strongest correlation value. This is likely influenced by the significant proportion of retail investors participating in daily trading activities, rendering the market more reactive to shifts in sentiment. On the other hand, the Thai and Malaysian markets appear to exhibit more moderate responses, though they still indicate a similar directional relationship. These results highlight the importance of considering psychological market factors in financial analysis, particularly during periods of heightened uncertainty. The findings support the relevance of sentiment monitoring as part of a quantitative approach to understanding market behavior, especially in times of economic instability.

V. DISCUSSION

The findings of this study indicate that negative sentiment expressed by retail investors on social media, particularly on Twitter, has a significantly negative correlation with stock index movements in Southeast Asia. This correlation is strongest with the IDX Composite index in Indonesia, suggesting that the domestic market is highly responsive to fluctuations in retail investor sentiment. This finding reinforces the results of (Sinaga et al., 2022), who demonstrated that negative sentiment on Twitter can trigger large-scale sell-offs in the Indonesian stock market during periods of crisis. The present study is also consistent with (Dumiter et al., 2023), who noted that emotional conditions such as anxiety influence the direction of global stock indices,

although their research focused on developed markets. The successful application of AI-based models in capturing daily sentiment dynamics lends support to the study by (Rezaei et al., 2025), which emphasizes the importance of integrating textual analysis in mapping market behavior based on public narratives. Data obtained from social media underscore the crucial role of public opinion as a non-traditional indicator capable of providing early signals of market direction changes, particularly in times of high volatility.

These findings also extend the scope of prior research that focused on lexicon-based or conventional NLP approaches by demonstrating that LLMs such as BERT and GPT exhibit higher accuracy in capturing complex emotional expressions in investors' tweets. This is supported by the study of (Liapis & Kotsiantis, 2023), which highlights BERT's capability to identify emotional sentiment closely associated with sharp market shifts. The analysis shows that heightened negative sentiment often precedes market index declines, a pattern also observed by (Liu et al., 2023) in the context of the U.S. technology market. In emerging markets, where retail investors constitute the majority of daily trading volume, public sentiment on social media becomes a highly relevant element for analysis. This study also highlights that the Indonesian market, which is predominantly retail-driven, is more susceptible to emotionally and collectively driven sentiment fluctuations. The approach adopted in this research demonstrates that unstructured data such as tweets can offer valuable insights not captured by conventional market indicators.

VI. CONCLUSION AND RECOMMENDATION

Based on the results of this study, it can be concluded that negative sentiment expressed by retail investors on social media, particularly through the Twitter platform, has a significant correlation with declines in stock market indices across Southeast Asia. This correlation is most pronounced in the IDX Composite Index in Indonesia, underscoring the substantial influence of collective perceptions among individual investors on the domestic market. These findings reinforce the role of digital sentiment as an alternative indicator for understanding market dynamics that cannot be fully explained by conventional economic indicators. Furthermore, the use of AI models such as BERT and GPT has proven effective in capturing complex emotional expressions in unstructured text, offering an advantage in the real-time assessment of investor sentiment. Therefore, social media sentiment not only reflects emotional responses but also has the potential to function as an early warning system for potential shifts in market direction, especially amid turbulent conditions. Consequently, the integration of market data with social data becomes crucial in building a more holistic analytical framework for understanding market behavior.

For future research development, it is recommended that studies be expanded to include additional variables, such as the frequency and intensity of social media user interactions, which may enrich the understanding of collective sentiment dynamics. Moreover, the development of AI-powered monitoring dashboards that integrate market and digital sentiment data in real time could serve as strategic tools for regulators and industry stakeholders. Such approaches should also be complemented by efforts to enhance digital financial literacy among retail investors in order to reduce excessive emotional reactivity in investment decision-making. Future research may also explore the effectiveness of public policy interventions in mitigating the impact of widespread negative sentiment on social media. The application of advanced NLP techniques, including multimodal models that incorporate text, images, and audio, could further expand the scope of analysis to encompass various forms of digital investor expression. As such, future studies are expected not only to improve the predictive accuracy of sentiment models but also to contribute to the creation of a more stable and emotionally adaptive market ecosystem.

REFERENCES

Abimbola, B., de La Cal Marin, E., & Tan, Q. (2024). Enhancing Legal Sentiment Analysis: A Convolutional Neural Network–Long Short-Term Memory Document-Level Model. *Machine Learning and Knowledge Extraction*, *6*(2), 877–897. https://doi.org/10.3390/make6020041

Addo, J. O., Cúg, J., Keelson, S. A., Amoah, J., & Petráková, Z. (2025). Behavioral Risk Management in Investment Strategies: Analyzing Investor Psychology. *International Journal of Financial Studies*, *13*(2), 53. https://doi.org/10.3390/ijfs13020053

Albahli, S., Irtaza, A., Nazir, T., Mehmood, A., Alkhalifah, A., & Albattah, W. (2022). A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data. *Electronics*, *11*(20), 3414. https://doi.org/10.3390/electronics11203414

Alshattnawi, S., Shatnawi, A., AlSobeh, A. M. R., & Magableh, A. A. (2024). Beyond Word-Based Model Embeddings: Contextualized Representations for Enhanced Social Media Spam Detection. *Applied Sciences*, *14*(6), 2254. https://doi.org/10.3390/app14062254

Aziz, M. I. A., Ahmad, N., Zichu, J., & Nor, S. M. (2022). The Impact of COVID-19 on the Connectedness of Stock Index in ASEAN+3 Economies. *Mathematics*, *10*(9), 1417. https://doi.org/10.3390/math10091417

Bello, A., Ng, S. C., & Leung, M. F. (2023). A BERT Framework to Sentiment Analysis of Tweets. *Sensors*, *23*(1), 506. https://doi.org/10.3390/s23010506

Bhattacharya, A., & Sardashti, H. (2022). The Differential Effect of New Product Preannouncements in Driving Institutional and Individual Investor Ownership. *Journal of Business Research*, *149*, 811–823. https://doi.org/10.1016/j.jbusres.2022.05.080

Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. *Electronics*, *11*(3), 374. https://doi.org/10.3390/electronics11030374

Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The Applications of Artificial Neural Networks, Support Vector Machines, and Long–Short Term Memory for Stock Market Prediction. *Decision Analytics Journal*, *2*, 100015. https://doi.org/10.1016/j.dajour.2021.100015

Dumiter, F. C., Turcaș, F., Nicoară, Ștefania A., Bențe, C., & Boiță, M. (2023). The Impact of Sentiment Indices on the Stock Exchange—The Connections between Quantitative Sentiment Indicators, Technical Analysis, and Stock Market. *Mathematics*, *11*(14), 3128. https://doi.org/10.3390/math11143128

Figà-Talamanca, G., & Patacca, M. (2022). An Explorative Analysis of Sentiment Impact on S&P 500 Components Returns, Volatility and Downside Risk. *Annals of Operations Research*, *342*(3), 2095–2117. https://doi.org/10.1007/s10479-022-05129-w

Kirtac, K., & Germano, G. (2024). Sentiment Trading with Large Language Models. *Finance Research Letters*, *62*(PB), 105227. https://doi.org/10.1016/j.frl.2024.105227

Koukaras, P., Nousi, C., & Tjortjis, C. (2022). Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning. *Telecom*, *3*(2), 358–378. https://doi.org/10.3390/telecom3020019

Li, T., Chen, H., Liu, W., Yu, G., & Yu, Y. (2023). Understanding the Role of Social Media Sentiment in Identifying Irrational Herding Behavior in the Stock Market. *International Review of Economics & Finance*, *87*, 163–179. https://doi.org/10.1016/j.iref.2023.04.016

Liapis, C. M., & Kotsiantis, S. (2023). Temporal Convolutional Networks and BERT-Based Multi-Label Emotion Analysis for Financial Forecasting. *Information*, *14*(11), 596. https://doi.org/10.3390/info14110596

Liu, Q., Lee, W. S., Huang, M., & Wu, Q. (2023). Synergy Between Stock Prices and Investor Sentiment in Social Media. *Borsa Istanbul Review*, *23*(1), 76–92. https://doi.org/10.1016/j.bir.2022.09.006

Maqsood, H., Maqsood, M., Yasmin, S., Mehmood, I., Moon, J., & Rho, S. (2022). Analyzing the Stock Exchange Markets of EU Nations: A Case Study of Brexit Social Media Sentiment. *Systems*, *10*(2), 24. https://doi.org/10.3390/systems10020024

Mendoza-Urdiales, R. A., Núñez-Mora, J. A., Santillán-Salgado, R. J., & Valencia-Herrera, H. (2022). Twitter Sentiment Analysis and Influence on Stock Performance Using Transfer Entropy and EGARCH Methods. *Entropy*, *24*(7), 874. https://doi.org/10.3390/e24070874

Oralbekova, D., Mamyrbayev, O., Othman, M., Kassymova, D., & Mukhsina, K. (2023). Contemporary Approaches in Evolving Language Models. *Applied Sciences*, *13*(23), 12901. https://doi.org/10.3390/app132312901

Rezaei, A., Abdellatif, I., & Umar, A. (2025). Towards Economic Sustainability: A Comprehensive Review of Artificial Intelligence and Machine Learning Techniques in Improving the Accuracy of Stock Market Movements. *International Journal of Financial Studies*, *13*(1), 28. https://doi.org/10.3390/ijfs13010028

Sin Huei, N., Zhuang, Z., Toh, M. Y., Ong, T. S., & Teh, B. H. (2022). Exploring Herding Behavior in an Innovative-Oriented Stock Market: Evidence from ChiNext. *Journal of Applied Economics*, *25*(1), 523–542. https://doi.org/10.1080/15140326.2022.2050992

Sinaga, J., Wu, T., & Chen, Y. (2022). Impact of Government Interventions on the Stock Market During COVID-19: A Case Study in Indonesia. In *SN Business & Economics* (Vol. 2, Issue 9). Springer

International Publishing. https://doi.org/10.1007/s43546-022-00312-4

Tran, K. L., Le, H. A., Lieu, C. P., & Nguyen, D. T. (2023). Machine Learning to Forecast Financial Bubbles in Stock Markets: Evidence from Vietnam. *International Journal of Financial Studies*, *11*(4), 133. https://doi.org/10.3390/ijfs11040133

Velu, S. R., Ravi, V., & Tabianan, K. (2023). Multi-Lexicon Classification and Valence-Based Sentiment Analysis as Features for Deep Neural Stock Price Prediction. *Sci*, *5*(1), 8. https://doi.org/10.3390/sci5010008

Yang, S. (2024). Pandemic, Policy, and markets: Insights and Learning from COVID-19's Impact on Global Stock Behavior. *Empirical Economics*, *68*(2), 555–583. https://doi.org/10.1007/s00181-024-02648-2