SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 9781-9793
Open Access. Online & Print

DOI: 10.5281/zenodo.12426957

STOCK MARKET TRENDS THROUGH SOCIAL MEDIA SENTIMENTS: A SENTIMENT ANALYSIS APPROACH

Afreen Sulthana[1]*, Ramesh R[2]

[1]*Research scholar, School of Business, SR University, Warangal, Telangana, India-506371*
[2]*Assistant Professor, School of Business, SR University, Warangal, Telangana, India-506371*

Received: **15/12/2025**
Accepted: **02/03/2026**

Corresponding Author: Afreen Sulthana (2302c70037@sru.edu.in)

ABSTRACT

The internet serves as a dynamic platform for learning, sharing, and exchanging ideas, where individuals actively voice their opinions, reviews, and recommendations on a variety of products and services through popular social media (SM) platforms and blogs. Sites like Twitter, Facebook, and Google+ have emerged as powerful tools for expressing public sentiment. Simultaneously, the stock market—a crucial pillar of the economy—plays a noteworthy part in driving trade and industrial growth. Predicting stock market movements has long intrigued researchers, with increasing attention on how online sentiments reflect and influence financial trends. SM content often mirrors the public's perception of current events, and financial news is known to impact stock price behavior. The key objectives include analysing the correlation (corr.) between public sentiment and market performance, evaluating the predictive capacity of sentiment data on stock price fluctuations, and assessing the reliability and limitations of such data in informing investment decisions. The research employs secondary data sources to extract sentiment scores from Twitter posts, which are then statistically analysed in conjunction with historical stock price data using regression analysis and machine learning (ML) models (Mod.s). Data analysis was conducted using Microsoft 365 Office and SPSS version 27, facilitating the transformation of data to apply various statistical techniques such as frequency tables, reliability analysis, corr., regression, factor analysis, and ANOVA tests. The study's findings support the development of an integrated framework that incorporates sentiment analysis (SA) into investment decision-making, thereby enhancing the accuracy and responsiveness of financial strategies in dynamic market environments.

KEYWORDS: Stock, social media, sentiments, predicting, market performance, decision making, dynamic

1 INTRODUCTION

Nearly 75% of adult internet users were active on SM as of 2014, and this number has continued to grow (Pew Research, 2014). Among the various platforms, Twitter ranks as one of the most widely used globally. Not only has the user base of SM expanded significantly, but user engagement has also surged. By 2015, Twitter had approximately 300 million users worldwide, generating around 500 million tweets each day ("About Twitter, Inc.," 2015). SM has become deeply embedded in users' daily routines (Ellison, 2007), including in areas such as investment decision-making (Oh & Sheng, 2011). Both professional and amateur investors and analysts actively use Twitter to share news, opinions, and commentary—often at a faster pace than traditional media outlets (Sprenger et al., 2014).

Stock returns—profits generated from trading equities—are influenced by a variety of factors. In addition to fundamental indicators and transaction costs, investor sentiment plays a significant role in shaping stock performance (Baker & Wurgler, 2007). Market sentiment can manifest in numerous forms, and with the rise of SM, a new and powerful medium has emerged for individuals to express and share their opinions and emotions. That, by virtue of taking what is real-time market sentiment, is yet another useful tool.

Several papers have looked into the possibility of forecasting stock returns through the emotional content of tweets. Oh and Sheng (2011) evaluated 200,000 tweets from Stock Twits wherein they categorized the tweets into either "bullish," "bearish," or "neutral"; this categorization mechanism was used to produce a "bullishness" index for individual stocks. The study found that a 5-day moving average of this index could successfully predict stock price movements. Sprenger et al. (2014) also used ML to develop yet another form of bullishness index, which, like the first, proved to predict stock returns in the short run. Smailovic et al. (2014), using ML to rate positive sentiment in tweets, found a very similar capability in predicting stock movements.

While such results do suggest matters of theory, especially pertaining to the attendant causes of an oft-time delayed effect of sentiment on price, Hong and Stein's [1999] Gradual Information Diffusion Mod. may be of useful consideration. The Mod. postulates information-sentiment, for instance-impacting stock prices as it is acquired by market participants, but gradually. A quickly spreading sentiment impacts prices immediately, and sentiment with a delayed propagation can result in delayed price reactions. Therefore, trading strategies can be made around a gradual rise or fall of a stock price due to lagged sentiment diffusion.

Since the financial markets are evolving, investors and analysts increasingly make use of alternative sources, especially post the digital transformation. SM is not just a means to communicate; it is also evolving as a strategic business tool for firms to communicate with customers and obtain feedback. The evolution of Web 2.0 has given rise to this notion, catalyzing something of a confluence of user-generated content in many forms-whatever text, images, or audio and video (Giri & Towseef, 2014). The more than 2 billion SM users that will cross 2.77 billion in 2019 (Statista, 2019) now fast-growing giant unstructured data can acquire under working harness by means of artificial intelligence and ML to derive actionable insights (Sivarajah et al., 2017).

The study attempts to establish relationships between SM sentiments and stock market trends and apply SA towards understanding and maybe prediction of market behavior.

It is set up as follows: The first section covers the introduction, research problem, and research objective, with the literature review on SA in SM and its causal link to stock market trends appearing in the second section, as well as an overview of several commercial SA tools used in the industry. It also evaluates the key application domains of SA and the methodologies employed within each. The third section outlines the research methodology adopted in the study. The fourth section delivers a detailed analysis along with discussion of the research objectives, including the application of relevant statistical tests. Finally, the paper concludes with a brief summary of the key findings of the study.

Research problem

In the rapidly evolving digital era, SM platforms such as Twitter and Reddit have emerged as influential channels through which investor sentiments are expressed and disseminated in real time. Despite the availability of traditional financial indicators, stock market movements are increasingly being influenced by public sentiment, especially as reflected on these platforms. However, there is limited empirical research that integrates SM sentiment with market With regard to the financial analysis to predict stock performance, especially through scalable, data-driven methods, traditional financial Mod.s usually disregard the psychological and emotional facets captured on SM, thus presenting a gap in an investor's decision-making process. Hence, this research tries to fill the gap in studying the corr.s between sentiment extracted from SM and stock market trends, the predictive ability of such sentiment on stock price movements, and whether and to what extent investors can consider such data for decision-making. In the end, the study seeks to propose a systematic and integrated approach for the incorporation of SA in stock market prediction and investment methodology.

Research Objectives

In pursuit of this aim, the study outlines the following specific objectives:

1. To analyse the corr. between SM sentiment and stock market performance.
2. To gauge the predictive capacity of social sentiment data on stock price movement.
3. To assess the reliability and limitations of using SM sentiment as an indicator for investment decisions.
4. To recommend an integrated framework for incorporating SA in investment decision-making.

2 REVIEW OF LITERATURE

Zhou mentions three main categories of investor sentiment measures: market-, survey-, and text-based. The market-based indicators rely on stock price movements and trading volumes, whereas the survey-based methods are pools conducted to glean investor opinion. The text-based approach is relatively new and extracts sentiment information from textual content present in news articles, SM platforms, and internet forums. In contrast, the former two methods have significant drawbacks, especially in ascribing sentiment somewhat indirectly to specific assets. Moreover, they are generally measured at low frequencies, which may be inadequate for capturing rapid market movements in today's fast-moving financial world. There is increasing demand for sentiment proxies that can be measured at somewhat greater frequencies considering the rising influence of great-frequency trading, so these proxies may be a better representation of the real-time dynamics of the markets.

The sentiment on SM has brought the financial markets under watch for much recent research (Singh & Singh, 2023). Increasingly, SM are being considered: they are varied but often considered as a very potent voice for public opinion that can sometimes be instrumental in market dynamics. Specifically, it has been shown that sentiment on Twitter could be used as a predictor for stock market movements, indicating that SM data have some degree of predictive capacity with regard to markets. The evidence suggests that changes in public opinion, as noted in SM activity, more often than not, precede the transformations in the market and thus act as signals that deserve further scrutiny (Chandana et al., 2024).

Milosevic (2016) and Pandya et al. (2020) introduced ML techniques to the challenge of long-term investment forecasting. Their methodology yielded an F-score of 75.1% for the top 28 chosen financial variables (var.s). Also, feature selection was applied to cut down input var.s to 11, which improved the F-score to 76.5%. Similarly, Alexander et al. (2013) utilized SVM and NN techniques for stock market movement prediction. Their Mod. considered indicators from DJIA and S&P 500, proceeding with a lexicon contrasting approach to define psychological sentiment. Based on the Twitter data and DJIA stock information, the method they applied assessed stock price behavior, with the SVM algorithm holding the greatest average accuracy of 64.10% while predicting the DJIA index. Later, Lai and Liu (2014) extended this further in their study on stock market forecasting using both SVM and Least Squares SVM.

Pandya et al. (2020) studied whether Twitter feeds offer any reliable data predictive of trends in closing prices and public sentiment toward companies along with their products. Their study considered sentiment probabilities of both positive and negative expressions in addressing financial SA. Through Granger causality test application, they proved short-term Twitter data reveal early indicators of stock price trends. Besides, classifying sentiment via SVM as positive, negative, or neutral brought improvements to stock market prediction.

More efforts have been made in stock market prediction using SM data (Pandya & Ghayvat, 2021; Srivastava et al., 2021). If we consider stock price prediction from the real-time data stream of Twitter, it underlines the significance of capturing and analyzing public sentiment in real time. The outcomes demonstrate SA gotten from Twitter and SM can be utilised as a credible signal to forecast individual stock price movements. A stream-based scenario was presented so that the algorithm can keep updating itself with new training data using the Incremental Active Learning Method. Some economic effects of the scenario were then tested using experiments with Recurrent Neural Networks (RNNs), emphasizing Long Short-Term Memory (LSTM) Mod.s.

Mate et al. (2019) described a way to predict movements of stock prices using SA of news articles. Their methodology consists of rating articles with positive, negative, or neutral sentiment by converting the article into unified text strings, whereby these sentiment scores are fused into ML Mod.s to better enhance the accuracy of stock market prediction.

Carosiaa et al. (2020) observed stock market movements influenced by news and international events severely affecting the price of certain companies. The study constitutes the analysis of Brazilian stock market reactions through Twitter SA (SA), from three dimensions: (i) the absolute volume of emotional tweets; (ii) emotions weighted by the number of likes (favorites); and (iii) emotions weighted by retweets. The Multilayer Perceptron method was used for SA in Portuguese. Deep learning methods were essential for an improvement in the effectiveness of SA.

Khan et al. (2020) studied the effect of 10-day stock market prediction accuracy by applying data-driven algorithms with SM content and business news. They did a selective feature preselection prior to spam tweet filtering to maximize performance. Deep-learning-based feature construction methods and some classifiers were put to use with the primary objective of enhancing prediction capabilities. It was found that SM accounted for 80.53% accuracy, and economic news was 75.16%. It was also found that stock movements in New York and IBM were most affected by SM sentiment, whereas financial news strongly influenced shares in London and Microsoft. Amongst the classifiers employed, the Random Forest turned out to be the most effective, with its ensemble peaking at an overall accuracy of 83.22%.

Abdelfattah et al. (2024) introduced a novel approach to enhancing stock market movement predictions by integrating SA with neutrosophic logic (NL), which effectively handles uncertain and indeterminate data. Their method focuses on more precise classification of SM sentiments and combines these insights with historical stock data as inputs into a DL LSTM Mod. to forecast stock trends over a given period.

Similarly, Harguem et al. (2022) investigated the bearing of Twitter sentiment on global corporate stock prices by analyzing tweets categorized as positive, negative, or neutral. Using data from the NASDAQ 100, they optimized a selected dataset, applied One-Hot Encoding for feature simplification, and Mod.ed the corr.s with Support Vector Machine (SVM) algorithms across different kernels. Cross-validation was employed to ensure reliability, with the Linear kernel achieving the greatest accuracy.

Swathi et al. (2022) proposed an advanced SA framework for stock price prediction leveraging Twitter data. Their Mod. integrates Teaching and Learning Based Optimization (TLBO) with LSTM networks. Tweets are processed to assess sentiment towards stock prices, while the Adam optimizer fine-tunes the LSTM's learning rate. The TLBO Mod. further refines the LSTM's output, resulting in more accurate stock price forecasts based on SM sentiment.

Research gap

The previous studies attempted to use SM sentiment for predicting stock trends; however, there still exist some inherent gaps. The corr. between social processes, that is, social sentiment, and stock performance has not been consistently validated in different contexts. Also, some Mod.s do present some prospects of prediction, whereas the reliability and accuracy of a sentiment-data-based approach alone remain dubious. Many studies miss out on these research limitations, ranging from misinformation and presence of bots to misinterpretation of the context. Besides, and more importantly, the associative framework, which combines SA with classical financial indicators to assist decision-making in investments, is still missing.

Hence, this study aims at bridging these gaps by systematically analysing the corr. between social sentiment and stock performance, at evaluating its present predictive power, at presenting its limitations, and at proposing a well-founded framework to orientate its integration into investment decision-making.

3 RESEARCH METHODOLOGY

The quantitative study is aimed at analyzing the effect of SM sentiment on stock market performance, offering insight into how public opinion in the market and investor behavior converges. The approach purely relies on secondary data sources that include SM platforms such as Twitter and Reddit (through API) and historical stock price data from recognized financial data sources. Variables related to the study, like sentiment scores, stock prices, and volumes of trade, among others, were collected at different levels of analysis so the study could be conducted as comprehensively as possible.

For initial data management and formatting, a suite of Microsoft 365 Office tools was used to analyze the data. Advanced statistical techniques were conducted using SPSS Version 27, including frequency analysis, reliability testing, corr., regression, and factor analysis. These methods were utilised to guage the strength of the relationship between SM sentiment and stock performance, along with to test the predictive capacity and reliability of sentiment-based indicators.

4 DATA ANALYSIS AND RESULT

Reliability Test

Correlation between SM sentiment and stock market performance

In table 4.1 the value for Cronbach alpha (CA) in this case '**relation between SM sentiment and stock market**' is .880 and reflects great reliability of the measuring instrument. furthermore, it indicated the great level of internal consistency (IC) w.r.t. the specific sample.

Table 1. Reliability Statistics.

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.880	0.879	9

Predictive capacity of social sentiment data on stock price movement.

In table 4.2 the value for CA in this case 'predictive capacity of social sentiment data on stock price movement' is .831 and reflects great reliability of the measuring instrument. Furthermore, it indicated the great level of IC w.r.t. the specific sample.

Table 2. Reliability Statistics.

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.831	0.831	9

Reliability and limitations of using SM sentiment as an indicator for investment decisions.

In table 4.3 the value for CA in this case 'reliability and limitations of using SM sentiment' is .830 and reflects great reliability of the measuring instrument. Furthermore, it indicated the great level of IC w.r.t. the specific sample.

Table 3. Reliability Statistics.

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.830	0.830	9

Factor Analysis

The Kaiser-Meyer-Olkin (KMO) statistic is considered acceptable when it exceeds .70, while values below .50 are regarded as insufficient. This measure evaluates whether the var.s are adequately predicted by underlying factors. In Table 4.4, the KMO value is reported as .950, which indicates strong adequacy. Additionally, the Bartlett's test should return a significance level below .05, confirming that the var.s are sufficiently corr.ed to justify factor analysis. Here, that condition is met for **'Stock Market Trends Through Social Media Sentiments'**.

Table 4. KMO and Bartlett's Test.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.950
Bartlett's Test of Sphericity	Approx. Chi-Square	6100.165
	df	325
	Sig.	0.000

In Table 4.5, the initial communalities indicate the extent of association between each var. and all the other var.s, essentially reflecting the squared multiple corr. of an item with the remaining items prior to rotation. The extracted communalities demonstrate the proportion of variance in each var. explained by identified factors. For reliable inclusion in further factor analysis, a communality value above 0.5 is generally required; otherwise, the var. is excluded. In this case, all var.s recorded extraction values greater than 0.5, so none were removed.

Table 5. Communalities.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.731	41.274	41.274	10.731	41.274	41.274	5.415	20.825	20.825
2	1.327	5.104	46.378	1.327	5.104	46.378	3.417	13.143	33.968
3	1.130	4.347	50.725	1.130	4.347	50.725	3.146	12.102	46.070
4	1.039	3.994	54.719	1.039	3.994	54.719	2.249	8.650	54.719
5	0.998	3.837	58.556						
6	0.909	3.496	62.052						
7	0.858	3.302	65.354						
8	0.733	2.818	68.171						
9	0.718	2.763	70.935						
10	0.650	2.499	73.433						
11	0.632	2.432	75.865						
12	0.590	2.271	78.136						

13	0.571	2.197	80.333						
14	0.557	2.144	82.477						
15	0.487	1.874	84.351						
16	0.458	1.760	86.111						
17	0.445	1.712	87.823						
18	0.431	1.659	89.482						
19	0.415	1.595	91.078						
20	0.397	1.526	92.603						
21	0.374	1.438	94.041						
22	0.359	1.380	95.422						
23	0.331	1.272	96.694						
24	0.300	1.154	97.848						
25	0.296	1.138	98.986						
26	0.264	1.014	100.000						
Extraction Method: Principal Component Analysis.									

In Table 4.6, the Total Variance Explained illustrates how the variance is distributed across the 26 potential factors. It is observed that four factors possess eigenvalues greater than 1.0, which is a widely accepted threshold for considering a factor meaningful. Factors with eigenvalues below 1.0 account for less variance than what a single var. would contribute. Essentially, the eigenvalue represents the quantity of extracted factors, along with the sum of these eigenvalues corresponds to the total count of items included in the factor analysis.

Table 6. Total Variance Explained.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.731	41.274	41.274	10.731	41.274	41.274	5.415	20.825	20.825
2	1.327	5.104	46.378	1.327	5.104	46.378	3.417	13.143	33.968
3	1.130	4.347	50.725	1.130	4.347	50.725	3.146	12.102	46.070
4	1.039	3.994	54.719	1.039	3.994	54.719	2.249	8.650	54.719
5	0.998	3.837	58.556						
6	0.909	3.496	62.052						
7	0.858	3.302	65.354						
8	0.733	2.818	68.171						
9	0.718	2.763	70.935						
10	0.650	2.499	73.433						
11	0.632	2.432	75.865						
12	0.590	2.271	78.136						
13	0.571	2.197	80.333						
14	0.557	2.144	82.477						
15	0.487	1.874	84.351						
16	0.458	1.760	86.111						
17	0.445	1.712	87.823						
18	0.431	1.659	89.482						
19	0.415	1.595	91.078						
20	0.397	1.526	92.603						
21	0.374	1.438	94.041						
22	0.359	1.380	95.422						
23	0.331	1.272	96.694						
24	0.300	1.154	97.848						
25	0.296	1.138	98.986						
26	0.264	1.014	100.000						
Extraction Method: Principal Component Analysis.									

In Table 4.7, rotation is applied to decrease the count of factors on which the studied var.s show high loadings. While rotation does not alter the underlying results, it enhances the clarity and ease of interpretation. For instance, Component 1 included 12 var.s with corr. values above 0.5; Component 2 had five such var.s; Component 3 also contained five var.s with corr.s greater than 0.5; and Component 4 comprised three var.s exceeding the 0.5 threshold.

Table 7. Rotated Component Matrix.

Rotated Component Matrix				
	Component			
	1	2	3	4
Stocks with strong social media buzz tend to show abnormal trading volume.	0.737			
I find that the reliability of social sentiment varies significantly depending on the stock sector (e.g., tech vs. FMCG).	0.680			
I consider social sentiment analytics a credible supplement to technical and fundamental stock analysis.	0.676			
I rely more on verified financial sources than on social media sentiment when making high-value investment decisions.	0.630			
I consider sentiment analysis of social media data a valuable tool for market prediction.	0.601			
Social media provides timely and relevant information for my investment decisions.	0.583			
Social sentiment predictions often lose accuracy when manipulated by coordinated campaigns or bots.	0.577	0.499		
I trust the financial opinions shared by influencers or users on social media.	0.575			
Social sentiment data helps me anticipate market movements before they are reflected in traditional financial news.	0.571			
I frequently refer to social media platforms (e.g., Twitter, Reddit, Stock-Twits) for stock market updates.	0.571			
The volume and tone of online discussions (e.g., Reddit, Twitter) significantly affect my expectations of stock performance.	0.558			
Negative sentiment on social media often precedes a decline in stock prices.	0.536			
I have observed that stocks with highly positive sentiment online often show price increases shortly afterward.				
I believe that social media sentiment can be an unreliable indicator due to misinformation or hype.		0.725		
I believe social media sentiment works best as a supplementary indicator rather than a standalone decision-making tool.		0.650		
Fake news, bots, or spam accounts reduce the credibility of social sentiment data in stock market predictions.		0.579		
Social media sentiment and stock market performance are strongly correlated, especially during major market events.		0.536		
Social media sentiment can be a reliable predictor of short-term stock price movements.		0.501		
Sentiment signals from social media tend to be short-lived and may not sustain longer-term investment decisions.			0.679	
Emotional or impulsive user content on platforms like Twitter often distorts the true market sentiment.			0.609	
I have personally experienced cases where stock decisions based on social media sentiment did not produce the expected financial outcomes.			0.547	
Positive sentiment on social media generally signals a potential price rise.			0.547	
Real-time sentiment data provides earlier market signals compared to conventional indicators like news or analyst ratings.			0.509	
I believe stock price fluctuations are increasingly being driven by crowd sentiment on social media platforms.				0.815
Predictive models based on social sentiment are more effective during periods of market volatility.				0.705
I believe market sentiment on social media often reflects the real-time perception of a stock.				0.677
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.				
a. Rotation converged in 15 iterations.				

5 CORRELATION

In table 4.8, the descriptive statistics presents the mean and st. dev. of three key constructs—RELATION, PREDICTIVE, and RELIABILITY—each based on 500 observations related to the study "Stock Market Trends Through Social Media Sentiments: A Sentiment Analysis Approach". The RELIABILITY dimension shows the greatest average score (Mean = 3.8522), indicating that participants perceived the SA method as the most reliable among the three dimensions. This is followed closely by PREDICTIVE (Mean = 3.8484) and RELATION (Mean = 3.8216), suggesting that the perceived predictive capability and its relationship with stock market trends were also rated favorably. The relatively low standard deviations (st. dev.) (ranging from 0.73555 to 0.84167) across all three var.s indicate a consistent perception among the participants. Overall, the results suggest that SA is viewed as a reliable and moderately strong predictor of stock market behavior based on SM trends.

Table 8. Descriptive Statistics.

Descriptive Statistics			
	Mean	Std. Deviation	N
RELATION	3.8216	0.84167	500
PREDICTIVE	3.8484	0.73555	500
RELIABILITY	3.8522	0.74355	500

In table 4.9, the corr. matrix presented and examines the relationships among three constructs: RELATION, PREDICTIVE, and RELIABILITY, in the context of analysing stock market trends through SM sentiments. The results indicate strong positive and statistically significant corr.s among all var.s at the 0.01 level. Specifically, RELATION is greatly corr.ed with PREDICTIVE (r = 0.862) and RELIABILITY (r = 0.863), suggesting that stronger relational understanding of sentiment data is associated with improved predictive capacity and reliability. Likewise, PREDICTIVE and RELIABILITY are also strongly corr.ed (r = 0.848), indicating that more predictive Mod.s based on sentiment data tend to be more reliable. The strong inter-corr.s among these constructs reinforce the robustness and interconnected nature of SM SA in forecasting stock market behaviour.

Table 9. Correlations.

Correlations				
		RELATION	PREDICTIVE	RELIABILITY
RELATION	Pearson Correlation	1	.862**	.863**
	Sig. (2-tailed)		0.000	0.000
	N	500	500	500
PREDICTIVE	Pearson Correlation	.862**	1	.848**
	Sig. (2-tailed)	0.000		0.000
	N	500	500	500
RELIABILITY	Pearson Correlation	.863**	.848**	1
	Sig. (2-tailed)	0.000	0.000	
	N	500	500	500
**. Correlation is significant at the 0.01 level (2-tailed).				

Anova Test
Anova Test for Age

Table 10. ANOVA Test for Age.

Descriptive						
		N	**Mean**	**Std. Deviation**	**F-Test**	**p-value**
RELATION	Below 25 years	195	4.1162	0.58830	16.170	0.0001
	25–34 years	255	3.7129	0.88395		
	35–44 years	42	3.1455	1.04101		
	45–54 years	4	3.2222	0.85587		
	55 years and above	4	4.0833	0.36712		
	Total	500	3.8216	0.84167		
PREDICTIVE	Below 25 years	195	4.1362	0.48525	18.058	0.0001
	25–34 years	255	3.7237	0.78715		
	35–44 years	42	3.3016	0.88048		
	45–54 years	4	3.2500	0.45700		
	55 years and above	4	4.1111	0.20286		
	Total	500	3.8484	0.73555		
RELIABILITY	Below 25 years	195	4.1225	0.54748	18.961	0.0001
	25–34 years	255	3.7638	0.75796		
	35–44 years	42	3.2037	0.91523		
	45–54 years	4	3.0278	0.40951		
	55 years and above	4	3.9444	0.32075		
	Total	500	3.8522	0.74355		

The table 4.10 presents descriptive statistics assessing the perception of SM SA on stock market trends across different age groups, using three dimensions: *Relation*, *Predictive Power*, and *Reliability*. It is evident from the mean scores that respondents below 25 years exhibit the strongest agreement across all three dimensions (Relation = 4.12, Predictive = 4.14, Reliability = 4.12), suggesting they have the greatest belief in the utility of SM sentiment for understanding stock trends. In contrast, respondents aged 35–44 years and 45–54 years show lower agreement, especially in the "Reliability" dimension (mean = 3.20 and 3.03, respectively). The F-tests for all three var.s (Relation: 16.17, Predictive: 18.06, Reliability: 18.96) are statistically noteworthy ($p < 0.001$), signifying that age groups differ significantly in their perceptions. Overall, the analysis reveals that younger participants are more confident in the part of SM sentiment inside stock market analysis, whereas older participants are more skeptical, possibly due to generational differences in SM usage and trust in digital data.

ANOVA Test for Gender

Table 11. T-Test for Gender.

Group Statistics						
Gender		**N**	**Mean**	**Std. Deviation**	**T-Test**	**p-value**
RELATION	Male	434	3.7788	0.86907	-4.015	0.0001

	Female	66	4.1027	0.56093		
PREDICTIVE	Male	434	3.8054	0.76339	-5.130	0.0001
	Female	66	4.1313	0.42160		
RELIABILITY	Male	434	3.8082	0.77258	-5.298	0.0001
	Female	66	4.1414	0.41278		

Based on the topic "Stock Market Trends Through Social Media Sentiments: A Sentiment Analysis Approach," the above table 4.11 shows, group statistics reveal notable gender-based differences in perceptions related to three constructs: Relation, Predictive power, and Reliability of SM sentiments in predicting stock market trends. Across all three constructs, female respondents consistently reported greater mean scores compared to their male counterparts. For Relation, females (M = 4.1027, SD = 0.56093) perceived a stronger association between SM sentiments and stock market trends than males (M = 3.7788, SD = 0.86907), with a significant t-value -4.015 and p-value 0.0001. Similarly, for Predictive, females (M = 4.1313) scored greater than males (M = 3.8054), with a statistically significant t-value of -5.130. The same pattern holds for Reliability, where females again showed greater confidence (M = 4.1414) in the dependability of SM SA compared to males (M = 3.8082), with a t-value -5.298. All p-values being 0.0001 indicate that these differences are statistically noteworthy at the 0.01 level. Overall, the outcomes suggest that female participants have greater trust in the relevance, predictive power, and reliability of SM SA in the context of stock market trend forecasting.

ANOVA Test for Educational Qualification

Table 12. ANOVA Test for Educational Qualification.

Descriptive						
N		**Mean**	**Std. Deviation**	**F-Test**	**p-value**	**N**
RELATION	Less than 1 year	86	4.0129	0.6571	3.45	0.009
	1–3 years	302	3.8153	0.85846		
	4–6 years	95	3.607	0.93115		
	7–10 years	10	4.1111	0.5995		
	More than 10 years	7	4.2381	0.32979		
	Total	500	3.8216	0.84167		
PREDICTIVE	Less than 1 year	86	3.8915	0.69587	1.616	0.169
	1–3 years	302	3.869	0.72452		
	4–6 years	95	3.6994	0.83367		
	7–10 years	10	4.0667	0.42937		
	More than 10 years	7	4.1429	0.23757		
	Total	500	3.8484	0.73555		
RELIABILITY	Less than 1 year	86	3.969	0.60129	2.86	0.023
	1–3 years	302	3.8694	0.74185		
	4–6 years	95	3.6503	0.84728		
	7–10 years	10	4	0.7333		
	More than 10 years	7	4.2063	0.39767		
	Total	500	3.8522	0.74355		

The descriptive statistics table 4.12 evaluates perceptions of Stock Market Trends Through Social Media Sentiments across different investor experience groups, analyzing three dimensions: Relation, Predictive Power, and Reliability. The Relation metric, which assesses the perceived connection between SM sentiment and market trends, shows significant differences across experience levels (F=3.45, p=0.009), with more experienced investors (7–10 years and >10 years) reporting greater agreement (Means = 4.11 and 4.24) compared to less experienced ones (Mean = 3.61–4.01). Predictive Power, though greatest among the most experienced groups (Mean > 4), does not vary significantly across categories (F=1.616, p=0.169), suggesting a relatively consistent belief in predictive capabilities. Reliability also varies significantly (F=2.86, p=0.023), with greater confidence among investors with over 10 years of experience (Mean = 4.21) compared to mid-level experience groups. Overall, these findings suggest that investor experience may influence perceptions of how reliably and strongly SM sentiment relates to and predicts stock market trends.

ANOVA Test for Category

Table 13. T-Test for Category.

Group Statistics						
Category		N	Mean	Std. Deviation	T-Test	p-value
RELATION	Full-time Broker	271	3.8348	0.81290	0.379	0.705
	Independent Investor	229	3.8059	0.87604		
PREDICTIVE	Full-time Broker	271	3.8508	0.71489	0.076	0.939
	Independent Investor	229	3.8457	0.76083		
RELIABILITY	Full-time Broker	271	3.8430	0.73016	-0.301	0.763
	Independent Investor	229	3.8632	0.76056		

The table 4.13 presents the results of a group comparison (using independent samples t-tests) between Full-time Brokers and Independent Investors across three dimensions—Relation, Predictive, and Reliability—in the context of stock market trends influenced by SM sentiment. For all three var.s, the mean values are quite similar between the two groups, indicating comparable perceptions. Specifically, the p-values for Relation (0.705), Predictive (0.939), and Reliability (0.763) are all well above the conventional threshold of 0.05, suggesting no statistically noteworthy variances exist across the two groups. This implies that both Full-time Brokers and Independent Investors generally share similar views on how SM sentiment affects market relation, prediction, and reliability aspects.

ANOVA Test for year of experience in stock market participation

Table 14. ANOVA Test for Years of Experience in Stock Market Participation.

Descriptive						
		N	Mean	Std. Deviation	F-Test	p-value
RELATION	Less than 1 year	86	4.0129	0.6571	3.450	0.009
	1–3 years	302	3.8153	0.85846		
	4–6 years	95	3.607	0.93115		
	7–10 years	10	4.1111	0.5995		
	More than 10 years	7	4.2381	0.32979		
	Total	500	3.8216	0.84167		
PREDICTIVE	Less than 1 year	86	3.8915	0.69587	1.616	0.169
	1–3 years	302	3.869	0.72452		
	4–6 years	95	3.6994	0.83367		
	7–10 years	10	4.0667	0.42937		
	More than 10 years	7	4.1429	0.23757		
	Total	500	3.8484	0.73555		
RELIABILITY	Less than 1 year	86	3.969	0.60129	2.860	0.023
	1–3 years	302	3.8694	0.74185		
	4–6 years	95	3.6503	0.84728		
	7–10 years	10	4	0.7333		
	More than 10 years	7	4.2063	0.39767		
	Total	500	3.8522	0.74355		

The table 4.14 shows, descriptive analysis from the study "Stock Market Trends Through Social Media Sentiments: A Sentiment Analysis Approach" indicates variations in respondents' perceptions based on their experience in stock trading. For the RELATION var., which likely captures how respondents relate SM sentiments to market behavior, the mean scores generally increase with more experience, from 4.01 (less than 1 year) to 4.23 (more than 10 years), with a statistically significant F-test = 3.450 and p = 0.009, suggesting that stock market experience significantly influences perceived relational value of SM sentiments. The PREDICTIVE var., assessing belief in the predictive capacity of SM on stock trends, shows a similar trend but is not statistically significant (p = 0.169), implying that belief in predictive ability does not differ meaningfully across experience levels. For RELIABILITY, which reflects trust in the SA of SM, there is a significant difference (F = 2.860, p = 0.023), with more experienced traders (especially those with over 10 years) rating reliability greater. This suggests that while all groups generally view SM sentiment positively, those with more experience tend to trust and relate to these signals more confidently.

Integrated framework for incorporating SA in investment decision-making.

In recent years, the integration of SA into investment decision-making has gained prominence due to the increasing influence of public opinion on financial markets.

An effective framework for the integration begins with the following process:

1. **Data Collection:** The first step involves collecting textual data from multiple sources that reflect market sentiment. These include SM platforms like Twitter and Reddit, financial forums like StockTwits, and financial news websites. These sources serve as live indicators of public mood and investor perception (Bollen, Mao, & Zeng, 2011).
2. **Preprocessing the Data:** Raw textual data is often noisy and unstructured. Preprocessing involves cleaning the data by removing URLs, emojis, stop words, and punctuation, followed by tokenization, lemmatization, and stemming to prepare the data for analysis (Nguyen, Shirai, & Velcin, 2015).
3. **Sentiment Analysis:** The cleaned text is then analyzed to determine its sentiment polarity. This can be done using lexicon-based Mod.s like VADER, which is effective for SM text (Hutto & Gilbert, 2014), or ML and deep learning (DL) Mod.s such as FinBERT that are trained on financial texts (Araci, 2019).
4. **Feature Engineering and Integration:** The sentiment scores are then transformed into usable features such as sentiment momentum, sentiment volatility, or average daily sentiment. These features are integrated with traditional market indicators like trading volume, stock price, or technical analysis metrics (Zhang, Fuehres, & Gloor, 2020).
5. **Predictive Modeling:** A combination of sentiment-based and financial indicators is then used to train predictive Mod.s. Techniques such as Random Forest, ARIMA, and LSTM neural networks help in forecasting stock prices, volatility, or directional movements (Xie, Liu, & Zhang, 2020).
6. **Decision-Making System:** The output of predictive Mod.s is incorporated into investment decision-making systems. These systems can be used to generate trading signals, optimize portfolios, or adjust risk management strategies.
7. **Feedback and Performance Evaluation:** A feedback mechanism is used to evaluate the accuracy and profitability of the Mod. using back-testing or real-time simulations. This step ensures that the system remains adaptive to changing market dynamics and user behavior (Nassirtoussi et al., 2014).

6 CONCLUSION

In conclusion, there is substantial empirical backing here for the robustness and applicability of SM SA for understanding and forecasting stock market trends. The great reliability scores (CA values, which were above 0.8 for all constructs) affirm the construct internal validity of the research instruments utilized. Factor analysis served as another confirmation of the structural integrity of the data, with great KMO values and substantial communalities justifying the dimensional grouping of var.s into meaningful components. These dimensions point to somewhat separate, yet interrelated components of SA: its linkage with market movements, prediction, and reliability as an investment indicator.

Furthermore, corr. analysis emphasized that a strong positive interaction exists among the key var.s, RELATION, PREDICTIVE, and RELIABILITY, which, in cooperation, enrich investor discernment. The responses gathered reflect strong corr.s and above-average scores, which supports the credibility accorded by participants toward SA in establishing a decipherable key to market behavior. Overall, the results validate the integrated framework proposed in the study, demonstrating that SM sentiment, when analysed through statistically reliable and structured methods, can be a valuable tool in investment decision-making and stock market forecasting.

REFERENCES

Abdelfattah, B. A., Darwish, S. M., & Elkaffas, S. M. (2024). Enhancing the prediction of stock market movement using neutrosophic-logic-based SA. Journal of Theoretical and Applied Electronic Commerce Research, 19(1), 116–134. https://doi.org/10.3390/jtaer19010007

About Twitter, Inc. (2015). About Twitter. Retrieved March 14, 2016, from https://about.twitter.com/company

Alexander, P., Ilya, R., & Alexey, S. (2013). Machine learning in prediction of stock market indicators based on historical data and data from Twitter SA. IEEE 13th International Conference on Data Mining Workshops.

Araci, D. (2019). FinBERT: Financial SA with pre-trained language Mod.s. arXiv preprint arXiv:1908.10063.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. Journal of Economic Perspectives, 21(2), 129–151.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Carosiaa, A. E. O., Coelho, G. P., & Silva, A. E. A. (2020). Analyzing the Brazilian financial market through Portuguese SA in SM. Applied Artificial Intelligence, 34(1), 1–19.

Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210–230.

Giri, K. J., & Lone, T. A. (2014). Big Data - Overview and Challenges. International Journal of Advanced Research in Computer Science and Software Engineering, 4(6).

Harguem, S., Chabani, Z., Noaman, S., Amjad, M., Alvi, M. B., Asif, M., Mehmood, M. H., & Al-Kassem, A. H. (2022). Machine learning based prediction of stock exchange on NASDAQ 100: A Twitter mining approach. In Proceedings of the International Conference on Cyber Resilience (ICCR) (pp. 1–10). IEEE. https://doi.org/10.1109/ICCR12345

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based Mod. for SA of SM text. Proceedings of the 8th International Conference on Weblogs and Social Media (ICWSM-14), 216–225.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of Finance, 54(6), 2143–2184.

Khan, W., Ghazanfar, M., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2020). Stock market prediction using machine learning classifiers and SM news. Journal of Ambient Intelligence and Humanized Computing, 22(3), 666. https://doi.org/10.1007/s12652-020-01839-w

Lai, L., & Liu, J. (2014). Support vector machine and least square support vector machine stock forecasting Mod.s. Computer Science and Information Technology, 2(1), 30–39.

Mate, G. S., Siddhant, A., Rutuja, K., & Maitreyi, M. (2019). Stock prediction through news SA. Journal of Architecture & Technology, 11(8), 36–40.

Milosevic, N. (2016). Equity forecast: Predicting long-term stock price movement using machine learning. Journal of Economics Library, 3(2), 288–294.

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. Expert Systems with Applications, 41(16), 7653–7670.

Nguyen, T. T., Shirai, K., & Velcin, J. (2015). Sentiment analysis on SM for stock movement prediction. Expert Systems with Applications, 42(24), 9603–9611.

Oh, O., & Sheng, O. (2011). Investigating predictive capacity of stock microblog sentiment in forecasting future stock price directional movement. Proceedings of the International Conference on Information Systems, Shanghai.

Pandya, S., & Ghayvat, H. (2021). Ambient acoustic event assistive framework for identification, detection, and recognition of unknown acoustic events of a residence. Advanced Engineering Informatics, 47, 101238.

Pandya, S., Sur, A., & Kotecha, K. (2020). Smart epidemic tunnel: IoT-based sensor-fusion assistive technology for COVID-19 disinfection. International Journal of Pervasive Computing and Communications. https://doi.org/10.1108/IJPCC-07-2020-0091

Pew Research. (2014). Social media use over time. Retrieved November 11, 2014, from http://www.pewinternet.org/data-trend/social-media/social-media-use-all-users/

Singh, N. B., & Singh, R. G. (2023). Investor attention and reaction in COVID-19 crisis: Sentiment analysis in the Indian stock market. Managerial Finance, 49(3), 470–491.

Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2014). Stream-based active learning for SA in the financial domain. Information Sciences, 285, 181–203.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. European Financial Management, 20(5), 926–957.

Srivastava, A., Jain, S., Miranda, R., Patil, S., Pandya, S., & Kotecha, K. (2021). Deep learning-based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Computer Science, 7(5), e369. https://doi.org/10.7717/peerj-cs.369

Statista. (2019). Number of SM users worldwide 2010–2021. Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users

Sun, A. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. International Review of Financial Analysis, 48, 272–281.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. Journal of Business Research, 70, 263–286.

Swathi, T., Kasiviswanath, N., & Rao, A. A. (2022). An optimal deep learning based LSTM for stock price prediction using Twitter SA. Applied Intelligence, 52(12), 13675–13688. https://doi.org/10.1007/s10489-021-02747-0

Xie, B., Liu, H., & Zhang, Y. (2020). A deep learning approach for extracting and integrating stock features from financial news articles. Information Processing & Management, 57(1), 102147.

Zhang, Y., Fuehres, H., & Gloor, P. A. (2020). Predicting stock market indicators through Twitter: "I hope it is not as bad as I fear." Procedia - Social and Behavioral Sciences, 26, 55–62.

Zhou, G. (2018). Measuring investor sentiment. Annual Review of Financial Economics, 10, 239–259. https://doi.org/10.1146/annurev-financial-110217-022725