The Impact of the Social Media Sentiment Index on S&P500 Returns

El índice de sentimiento en las redes sociales y su impacto en los rendimientos del S&P 500

Lizeth Gordillo Martínez

EGADE Business School, Tecnológico de Monterrey, Mexico

Received: April 1st, 2024.
Approved: May 30, 2024.

Abstract

The study of the development and analysis of sentiment indexes through social media is a recent technique that has captured interest because it can identify stock price tendencies. Also, using artificial intelligence to quickly analyze large volumes of data from various information sources has created a new way of evaluating massive amounts of information from social media. Natural language processing (NLP) is the preferred method for this research. Originating in the 1950s, NLP emerged at the junction of artificial intelligence and linguistics. Initially, it was used to retrieve information in text. It uses methods based on statistics to effectively index and search large sections of text.

Keywords: *Twitter, social media, natural language processing, sentiment analysis, artificial intelligence, news.*
JEL Classification: *G1, G02, C01.*

Resumen

El estudio de la construcción y el análisis de índices de sentimiento en redes sociales es una técnica reciente que ha captado interés por su capacidad para identificar tendencias en los precios de las acciones. Además, la aplicación de inteligencia artificial para analizar rápidamente grandes volúmenes de datos de diversas fuentes de información ha creado una nueva forma de evaluar información masiva de redes sociales. El procesamiento del lenguaje natural (NLP, por sus siglas en inglés) es el método preferido que se sigue en la investigación. Originado en los años cincuenta, el NLP surgió de la intersección entre la inteligencia artificial y la lingüística. En un comienzo se empleó para recuperar información textual, con métodos basados en estadísticas para indexar y buscar de manera eficaz en grandes secciones de texto.

Palabras clave: *Twitter, redes sociales, procesamiento del lenguaje natural, análisis de sentimiento, inteligencia artificial, noticias.*
Clasificación JEL: *G1, G02, C01.*

The Anáhuac Journal, **Vol. 24, núm. 1, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n1.08

1. Introduction

Stock markets operate on sentiment, where investors' herd mentality often results in emotional overreactions and investment choices. Contemporary social media plays a significant role in affecting sentiments and decisions in the stock market. Access to news and information is crucial for successful investment strategies, with social media serving as an auxiliary resource for decision-making, facilitating the rapid dissemination of financial news, sometimes even faster than traditional news tickers (Lugmayr, 2012).

The rapid expansion of digital data creation has ushered in the big data era, made increasingly significant as around 70% of the world's collected data originates from social media, sentiment analysis, identifying sentiments and emotions expressed about societal issues on social media (Lee et al., 2020).

We are quoting several cases where social media spread the word of relevant events that will affect financial markets significantly. Our financial markets are described as the S&P 500 Index and their returns according to these events. Social media and artificial intelligence have made this evaluation process more efficient, automated, and data transparent tracking to process bid data coming from Twitter.

In a recent paper, the Federal Reserve Board of Governors revealed that the Twitter financial sentiment index, which applies natural language processing to social media data, can assist in forecasting stock market returns for the following day. The study observed that sentiment declines following unexpected policy tightening, indicating that the data can also somewhat predict the extent of interest rate increases by the central bank (Smith, 2023).

Throughout 2021 and 2022, the Silicon Valley Bank (SVB) found itself with an abundance of bank deposits and not enough loans to disburse. Consequently, the bank was compelled to invest these customer deposits into long-term securities, primarily bonds. Typically, banks holding such securities aim to retain them until maturity to optimize returns. U.S. Treasury bonds, backed by the U.S. Government, carry a very low risk of default. However, Silicon Valley Bank's inadequate risk management practices left it unprepared for the Federal Reserve's rapid interest rate increases from under 1% to 4.75% within a year. This abrupt surge in interest rates resulted in a sharp decline in bond prices, highlighting the inverse relationship between interest rates and bond values. The hikes in interest rates unveiled mismatches in assets and liabilities at various regional banks across the U.S. (Manda, 2023).

Existing studies claim that investors and portfolio managers insight into how the Silicon Valley Bank collapse influenced various equity sectors, while regulators can gain an understanding of its broader effects. The role of social media in quickly triggering bank runs remains under investigation, as the U.S. has not seen a bank failure since 2008. Research works show that investors did not foresee the swift bank run, and also that the run did not uniformly affect all sectors of the U.S. market. This information is crucial for regulators attempting to address the risks posed by social media-driven "flash mob" reactions to the stability of banks. It is imperative for policymakers to closely watch social media channels to proactively identify and mitigate the risk of contagious financial panics (Yousaf & Goodell, 2023).

On April 14th, 2022, Elon Musk made his well-known bid to acquire Twitter for 43 billion dollars in a private offer. In response to this takeover bid, Twitter announced that it would conduct a thorough review. Through interviews and social media, Elon Musk expressed his intention to transform Twitter into a global free speech platform, emphasizing that the acquisition was not for business purposes or to boost visibility.

On April 25th, Twitter's board unanimously and publicly agreed to the acquisition for 44 billion dollars. Following the completion of this transaction, Twitter would officially transition into a private entity (Jia & Xu, 2022).

Primarily, Twitter aimed to leverage financial backing from Elon Musk to enhance its platform, software development, and IT maintenance, thereby gaining a competitive edge over rivals like Facebook and Instagram. Additionally, Twitter's founders viewed Musk's acquisition as an opportunity to implement an exit strategy, allowing them to divest their shares to Musk and disengage from the company.

Elon Musk, on the other hand, envisioned transforming Twitter by making its algorithms open source, meaning the rationale behind tweet visibility to users would be publicly accessible. This approach contrasted with the current norms in social media, characterized by content restrictions and pre-validation requirements. Musk aspires to create a more open environment that champions free speech (Jia & Xu, 2022).

The background of this study begins with Tetlock (2007) who analyzes the interactions of the media, which in this case is the print media of *The Wall Street Journal* and the influence that these interactions have on the share price. It is identified that a low perception of the company expressed in the news precedes a high trading volume and a high perception of the company precedes a low trading volume. For this article, we will use autoregressive vectors.

For the methodology and the extraction of data and its analysis, we will be using daily data that was extracted with the Twitter API focused on the 2225 American companies corresponding to the SPX (Standard & Poors 500 Index). Once the information has been extracted, we will go on to identify the influence of negative news concerning positive news. An automated process was used for data extraction, calculating sentiment, creating time series, and estimating VAR. To determine a specific period, a rolling widow will be applied on day 1, day 2, day 3... until day 18 taking Elon Musk's event. In the case of 18 days before and after the event, we found a statistically significant impact of the compound index on the returns of the S&P 500.

Given that we want to identify if there is a bidirectional relationship between the sentiment Index and the S&P 500, our hypotheses will be the following:

H0: The composite sentiment has a negative effect and a statistical effect on the S&P 500 returns.

H1: The composite sentiment has a non-negative effect on the S&P 500 returns.

The study concludes by assessing the practical and theoretical implications of incorporating a sentiment index using artificial intelligence into the analysis of financial markets described by their main stock indices, as well as its limitations and potential challenges. This analysis contributes to understanding the complexity and dynamics of the markets through current technology and new media self-creation by society, providing investors, analysts, and regulators with a valuable tool for informed and strategic decision-making which includes a methodology that can be replicated in different media outlets.

2. Theoretical Framework

Before the advent of social media, the question of how the news affected the share price was already appearing in various sources of financial literature. Short-term stock prices could be predicted by financial news (Gidófalvi, 2001). Advances in technology developed a dynamic environment for creating sharing and collaborating among Internet users, allowing investors to access information on the performance of companies quickly (Li et al., 2014). News was available in real-time through various sources of information on the Internet and social media. Thomson Reuters News API was used to construct a daily series for the DJI Index's sentiment scores (Tikkanen, 2021). Unstructured data, such as news, could be analyzed with machine learning to identify causality through entropy transfer. Multiple kernel learning identified distinct types of information for the Korean market (Nam & Seong, 2019). Several disciplines document the impact of news on stock prices—computer science, statistics, economics, and finance. The reliability of Big Data's computational models in price predictions is very sensitive and can lead to profit losses in the industry analyzed (Shah et al., 2018). Various studies have questioned the hypothesis of efficient markets in different sectors and over different periods. All phenomena that have not been explained by this theory can be described by the adaptive markets' hypothesis and the fractal market's hypothesis (Núñez-Mora & Mendoza-Urdiales, 2023).

Text analysis is relatively new in the study of financial market behavior. Tetlock (2007) analyzed political forecasts and the relationship between analyst behavior and market outcomes. Schindler (2013) researched market psychology and the development of the Case-Shiller Index, used to measure house prices. Thaler (2015) developed behavioral economics and behavioral finance theory. O'Shaughnessy (2006) worked on quantitative investment strategies, including the consideration of behavioral factors.

Feuerriegel, Heitzmann, and Neumann (2015) analyzed oil prices to determine the relationship between these prices and terrorist attacks using the methodology of autoregressive vectors. Ruan et al. (2018) set out to determine the level of trust of users and identify whether there is a correlation with the results of financial markets. Derakhshan and Beigy (2019) used econometrics and artificial intelligence to demonstrate the relationship between news and price returns. Other people's opinions are an essential piece of information for decision-making. The use of the Internet has generated the availability of these opinions in large volumes on different topics, creating complexity due to contrasting opinions.

Broadstock and Zhang (2019) evaluated how Twitter influences share price. They assessed the stock's intraday returns and how they reacted to sentiment about the stock. This demonstrated how the share price is susceptible to price factors involving social media sentiment. Mendoza-Urdiales et al. (2022) used a database of more than 50 million tweets using artificial intelligence to obtain the information and analyze the results. In the same study, the polarity of news is highlighted, whether in a positive or a negative tone, giving greater weight to negative news in the reaction of the share price. It was observed that negative comments have a more relevant impact on the price. This study aims to leverage all accessible information sources that mention companies with recognizable tickers which can be automatically extracted to evaluate the sentiment of each text.

In another study**,** the authors indicate that sentiment has an impact on prices and there is an asymmetric effect—for example, negative news tends to have a stronger impact than positive news (Mendoza-Urdiales et al., 2022). One other problem mentioned is the growing volume managing these large databases and the challenge this involves. Investors must analyze the information within an extremely brief time frame. Several market players used Twitter to follow decision-makers and official company channels, and it became a popular source of information (Heston & Sinha, 2016). Steinert and Herff (2018) elaborated on the prediction of stock returns from the news.

There are also documented instances of false information affecting a stock's price. As a historical example, one of the most notorious instances of false information impacting the financial market was the cyberattack on the Associated Press (AP) news agency in 2013. In this incident, hackers compromised the AP's Twitter account, posted a fabricated tweet alleging explosions at the White House, and claimed that President Barack Obama had been injured. Although the tweet was quickly deleted and confirmed to be false, it caused a temporary dip in financial markets. The Dow Jones Industrial Average lost several points in a matter of minutes.

Silicon Valley Bank (SVB) was a pivotal financial institution in the startup and technology ecosystem, particularly in Silicon Valley. Founded in 1983, SVB's specialty was providing banking services to startups, technology companies, and the venture capital industry. Its unique focus on this niche allowed it to better understand the needs of its customers and offer financial products and services tailored to fast-growing, high-risk businesses.

SVB not only provided traditional banking services such as account management and loans, but also offered equity financing, investment services, and advice to companies in various phases of growth, ranging from nascent startups to more mature, publicly traded entities. In addition, his deep involvement with the community of investors and entrepreneurs allowed him to be a key player in the networking and business development network within the technology and innovation sector (Akhtaruzzaman et al., 2023).

2.1 The Bankruptcy of SVB in California, United States, in March 2023

SVB was an American commercial bank that failed due to a lack of deposits to cope with rising interest rates. Its strategy focused on financing technology startups that pose a greater risk because they do not yet have enough income. The authorities took control of them and initiated measures to cover depositors (Van Vo & Le, 2023). It was the biggest U.S. bank failure since 2008. The following was posted on Twitter: "Europe calls for calm in the face of the failure of Silicon Valley Bank. The Commission recalled that its presence in Europe is very limited, but they assured that they will be very vigilant (Radio 5, 2013, March 13)."

The news was not well-received by the media and the share price fell by 60%, the adverse effect on the share price corroborates the hypothesis that the negative effects of a news story have a greater effect than positive news. Pandey et al. (2023) also analyzed this.

3. Methodology

In this paper, we will use the methodology developed by Mendoza-Urdiales et al. (2022). Our sampling plan is as follows: The sentiment index was created by obtaining the historical tweets referring to each stock of 2,557 companies in the S&P 500 index. This index represents 85% of US equity markets. We extracted the English comments mentioning each company ticker in Twitter ($+Ticker) (Núñez-Mora & Mendoza-Urdiales, 2023). A daily time frame was set up. The time interval ran from October 3, 2022, to April 27, 2023. The Twitter API was used through Rstudio.

3.1 Data Extraction

The information was expressed in text, and the next step was to convert the text into variables. We describe the process sequentially:

a) Request Type. The *rtweet* package offers various request options based on the level of Twitter access granted. The most frequently used function is "search_tweets" which allows for the retrieval of tweets from up to seven days before extraction, with a limit of up to 500,000 tweets.

b) The identification of the ticker is done with '$+Ticker'.

c) Identify the maximum number of tweets—that is 500,000 tweets.

d) We determined that English was the language of extraction.

e) The way to check if we were extracting the entire tweet is through the code tweet_mode= 'extended' to make sure that we were getting all the information of the tweet which was up to 280 characters. Refinitiv Eikon API was used to extract the financial information of the companies. Refinitiv software uses its RIC code.

f) Selection of variables. We selected 2,557 American companies representing 85% of the capital markets and extracted all the comments in English identified by a ticker on Twitter ($+Ticker).

To create sentiment analysis, we needed a dependent variable, in this case, stock returns. The independent variable was the composite sentiment index created from tweets. With all the above, we ran several statistical analyses and converted the information into qualitative and quantitative data. The NLP algorithm was executed using Python, utilizing the TextBlob library, which fragments each text into words and assigns a numerical value to each word according to the categories of the libraries, one value for polarity and another for subjectivity. Here we used the composite sentiment index as an independent variable.

3.2 The Econometric Model

Vector autoregressive models (VAR) gained prominence in the field of econometrics through Sims's work in 1980. They serve as an extension of univariate autoregressive models to a broader context. VARs function as multi-equation regression models, incorporating multiple dependent variables, thus positioning themselves as a midpoint between singular time series models and concurrent equations models.

They have been frequently presented as a preferable option to the expansive simultaneous equation models used in structural analysis. In its most basic form, a VAR model can be bivariate, involving just two variables, y1t and y2t.

The present value of each variable is influenced by a unique mix of the past k values of both variables.

$$Y_{1t\,=}\beta_{10+}\beta_{11}\,Y_{1t-1+}\ldots{+}{+}\ldots{+}{+}\,\beta_{1k}\,Y_{1t-k}+\propto_{1t}Y_{2t-1}\,\alpha_{1k}\,Y_{2t-k}\,U_{1t}$$

$Y_{2t} = \beta_{20} + \beta_{21} Y_{2t-1} + \ldots + + \ldots + + \beta_{2k} Y_{2t-k} + \propto_{2t} Y_{2t-1} \alpha_{2k} Y_{1t-k} U_{2t}$

Another advantage of VAR models lies in the compactness of their notation. For instance, take the previously mentioned scenario where k equals 1, meaning each variable's current value relies solely on the past values of y1t and y2t, accompanied by an error term. This could be expressed as:

$Y_{1t} = \beta_{10} + \beta_{11} Y_{1t-1} + \propto_{1t} Y_{2t-1} + U_{1t}$

$Y_{2t} = \beta_{20} + \beta_{21} Y_{2t-1} + \propto_{2t} Y_{1t-1} + U_{2t}$

Autoregressive vectors (VARs) are used to identify relationships between variables and the bias that emerges from simultaneous equations if it is overlooked. We select autoregressive vectors to find the bidirectional relationship between two variables. We want to find the explanation for the possible existence of two-way causality between stock prices and the composite sentiment index.

The Granger representation theorem states that if there exists a dynamic linear model with stationary disturbances and the data are I(1), then the variables must be cointegrated of order (1,1) (Brooks, 2019).

3.3 The Granger Causality Test

This test is a statistical hypothesis test for determining whether a one-time series is useful in forecasting another. While the Granger causality test is not a test of true causality, it checks whether past values of one variable help predict the future value of another variable better than using past values of the target variable alone.

The test is typically applied to a pair of time series X and Y. Here is how the test is mathematically structured using a simple bivariate model.

Step 1: Model only with past values of the target variable.

First, we fit an autoregressive model (AR) to the target variable Y:

$$Y_t = \propto + \sum_{i=1}^{p} \beta_i \gamma_{t-i} + \epsilon_t$$

The Anáhuac Journal, **Vol. 24, núm. 1, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n1.08

Where:

- α is a constant.
- β_i are the coefficients of lagged values of \(Y\).
- ε_t is the error term.
- ρ is the number of lags.

Step 2: Augmented model with past values of both variables.

Next, we fit an augmented model that includes past values of both \(Y\) and the potential causal variable X:

$$Y_t = \propto' + \sum_{i=1}^{p} \beta'_i \gamma_{t-i} + \sum_{j=1}^{q} y_j x_{t-j} + \epsilon'_t$$

Where:

- $\alpha' \beta'_i$ and y_i are coefficients.
- ε_t is the new error term.
- q can be equal to p or different, depending on the model selection.
- ρ is the number of lags.

3.4 Granger Causality Test

To determine if **X** Granger causes **Y**, we performed an F-test to compare these two models. The null hypothesis is that **X** does not Granger-cause **Y**, which implies that the coefficients y_j for all lags **j** are zero:

$$H_0: \gamma 1 = \gamma 2 = \ldots = \gamma q = 0$$

If the null hypothesis is rejected (meaning at least one of the y_i coefficients is significantly different from zero), then we conclude that **X** Granger causes **Y**.

This method checks if the inclusion of the past values of **X** significantly improves the model's ability to predict **Y**, compared to the model that only includes past values of **Y**. It is essential to note that if **X** Granger causes **Y**, it does not imply a causal relationship in a traditional sense, but rather a predictive directional relationship (Brooks, 2019).

4. Results

The Autoregressive Vector results are presented in Table 1, showing that the sentiment variable (named compound) has an almost immediate negative effect on the SPX index performance. Considering a 95% confidence interval with p-value ≤ 0.05, the statistically significant coefficient is only present within a day (Lag 1), and the effect on performance is negative (as expected).

Moreover, neither variable influences the sentiment equation, meaning that there is no feedback from the sentiment itself or from its performance.

As for the result of the influence of the SPX index performance on itself, the results show intermittence (considering a 95% confidence interval with p-value ≤ 0.05), the statistically significant coefficients are present within a day (Lag 1) and again on day 8th (Lag 8), and the effect on the performance is negative (see Table 1).

This result concludes that there is a negative sentiment in the market around the date of Twitter's acquisition. As we know, this acquisition involved highly criticized actions from Elon Musk, generating uncertainty and, consequently, a negative sentiment.

5. The Model's Results

As a key date, we selected October 28, 2022, when Elon Musk bought Twitter. Our results came out positive with the variation of the composite index. When you see a positive return, it already hits the composite index positively. The result we got is that sentiment is positively affected, and lagging positive sentiment is observed. In data analysis, statistics, and financial modeling, a rolling window is a method where computations are conducted on a set number of sequential data points. After each calculation, the window moves forward one or more periods to incorporate new data. This technique is employed to evaluate time series data, allowing analysts to detect trends, assess variability, and conduct various other analyses on a dataset that is continually updated. We use 18 days on a rolling window testing daily, day 1, day 2, day 3, day 4, until day 18 (Brooks, 2019).

The Anáhuac Journal, **Vol. 24, núm. 1, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n1.08

Table 1. Regression Results for SPX Returns with the statistically Significant Variables Highlighted (p-value ≤ 0.05)

```
Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Wed, 06, Mar, 2024
Time:                     13:18:58
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                  -7.52723
Nobs:                     30.0000    HQIC:                 -8.60723
Log likelihood:           85.5925    FPE:               0.000149006
AIC:                     -9.11525    Det(Omega_mle):    6.07088e-05
--------------------------------------------------------------------
```

Results for equation SPX Returns

	coefficient	std. error	t-stat	prob
const	0.003549	0.001573	2.256	0.024
L1.SPX Returns	-0.432398	0.210097	-2.058	0.040
L1.Compound	-0.003140	0.001554	-2.021	0.043
L2.SPX Returns	-0.104791	0.236009	-0.444	0.657
L2.Compound	0.000025	0.001650	0.015	0.988
L3.SPX Returns	-0.084670	0.223500	-0.379	0.705
L3.Compound	-0.000549	0.001663	-0.330	0.741
L4.SPX Returns	0.117132	0.217736	0.538	0.591
L4.Compound	-0.002552	0.001694	-1.507	0.132
L5.SPX Returns	0.244943	0.209261	1.171	0.242
L5.Compound	0.001060	0.001642	0.646	0.519
L6.SPX Returns	-0.140113	0.206786	-0.678	0.498
L6.Compound	0.001319	0.001746	0.756	0.450
L7.SPX Returns	-0.341930	0.210242	-1.626	0.104
L7.Compound	0.001889	0.001757	1.075	0.282
L8.SPX Returns	-0.418846	0.199433	-2.100	0.036
L8.Compound	0.001559	0.001517	1.027	0.304

Results for equation Compound

	coefficient	std. error	t-stat	prob
const	0.054744	0.279006	0.196	0.844
L1.SPX Returns	44.207016	37.263753	1.186	0.235
L1.Compound	0.141436	0.275564	0.513	0.608
L2.SPX Returns	-13.429302	41.859507	-0.321	0.748
L2.Compound	-0.114001	0.292632	-0.390	0.697
L3.SPX Returns	33.155597	39.640905	0.836	0.403
L3.Compound	-0.008319	0.294945	-0.028	0.977
L4.SPX Returns	8.745297	38.618612	0.226	0.821
L4.Compound	0.353053	0.300419	1.175	0.240
L5.SPX Returns	8.416371	37.115352	0.227	0.821
L5.Compound	0.096010	0.291294	0.330	0.742
L6.SPX Returns	8.994378	36.676488	0.245	0.806
L6.Compound	-0.036604	0.309656	-0.118	0.906
L7.SPX Returns	-36.473068	37.289423	-0.978	0.328
L7.Compound	0.300948	0.311576	0.966	0.334
L8.SPX Returns	-24.754087	35.372264	-0.700	0.484
L8.Compound	-0.218066	0.269132	-0.810	0.418

Correlation matrix of residuals

	SPX Returns	Compound
SPX Returns	1.000000	0.059417
Compound	0.059417	1.000000

Source: Prepared by the autor.

Granger causality Wald-test. H_0: Compound does not Granger-cause SPX Returns.

Conclusion: fail to reject H_0 at a 5% significance level.

The composite sentiment index is significant at 5% and negatively affects the performance of the Standard & Poor's Index with parameter -0.003140. What we are saying is that around the event on October 28, 2022, we selected 18 days before and 18 days after. Remember that the composite index is the sum of the positive, the negative, and the neutral sentiment. What dominated this period was negative sentiment and that hit the performance negatively.

In the S&P 500 performance equation, the influential lag is L1 which is the lag of the composite sentiment index at -0.003140.

Performance is negatively influenced by the composite sentiment index which in turn is influenced by lags in the performance of the S&P 500.

Granger causality simply indicates a correlation between the current value of one variable and the past values of others; it does not imply that changes in one variable directly cause changes in another (Brooks, 2019).

There is no Granger causality in the sense of the composite index to the performance of the S&P 500, i.e., the composite does not cause the performance of the S&P 500 in the Granger sense.

5.1 Impact of the Composite Sentiment Index on S&P 500 Returns

The findings validate the distinct connection between the composite sentiment index and the S&P 500 returns index, showing significant influence at the 5% level and a negative impact on the S&P 500 index's performance. The composite index affects the S&P stock index's performance with a -.003 impact.

The test concludes that you cannot reject the null hypothesis. Given that we cannot reject it, then the composite index does not, in Granger's sense, cause the return.

6. Discussion

X (formerly Twitter), as a frontrunner among social media platforms, is a constant presence in people's daily lives and has become a vital component of digital communication, but the emergence of short video-centric platforms, like TikTok, changed the game rules. This global social media boasted a sizeable user base of approximately 319 million people in 2022. It facilitates the widespread distribution of information, underpinning the value of enabling instant and unrestricted creation and sharing of ideas and information by all users. Additionally, it provides access to trending topics and has a considerable influence (Jia & Xu, 2022).

X (formerly Twitter) primarily generates its revenue through advertising, operating as an advertising-centric company. It is likely that Elon Musk recognized the potential in this network advertising sector and aimed at enhancing his presence in the digital advertising realm, seeing an opportunity to profit from acquiring the platform. Additionally, Musk might have viewed the ownership of Twitter as a means to shape public opinion more effectively, furthering his business and political agendas (Jia et al., 2023).

Observing X (formerly Twitter), it is evident that the emergence of short video-centric social media platforms has disrupted the growth trajectory of traditional social media. The rate of increase in this network daily active users has been on a downward trend in recent years, plummeting from a peak of 34% to 11%. Conversely, new social media platforms, with TikTok at the forefront, are steadily chipping away at the finite market share. X (formerly Twitter) has reported a profit decrease since 2019 (Jia et al., 2023).

Following Elon Musk's announcement of his intention to buy Twitter, the move garnered significant interest from various societal sectors. Musk, known for his unconventional approach, is a key opinion leader and, with this acquisition, he places himself at the forefront of the social media sector that moves the financial market. He has political and personal branding motivations behind this acquisition.

The event we determined as relevant was the purchase of Twitter by Elon Musk, the CEO of Tesla and SpaceX. This was one of the most controversial acquisitions in recent memory. In April 2022, Musk revealed that he had acquired a significant stake in Twitter, becoming one of its largest individual shareholders. On April 14, 2022, Musk offered to buy Twitter for $44 billion, proposing a price of $54.20 per share. Musk argued that his interest in buying it was his desire to ensure free speech on the platform to ensure the functioning of a democracy.

According to Musk, Twitter had the potential to be the quintessential free speech platform, but he believed it did not live up to that potential under its current administration. Initially, Twitter's board of directors responded to Musk's offer by implementing a measure to make it harder for the company to a hostile takeover by allowing other shareholders to buy more shares at a reduced price. After further negotiations, the board accepted Musk's offer on April 25, 2022.

The acquisition faced several hurdles and controversies. Over the next few months, Musk expressed concerns about the number of fake and spam accounts on the platform, leading to a temporary pause on the deal as he sought more information.

After a legal battle and the possibility of a trial, Elon Musk proceeded with the purchase of Twitter, closing the deal in October 2022.

The S&P 500 captures the wider market sentiment concerning the leading publicly traded companies sourced from the world's most extensive focus group. The connection between S&P 500 returns and Twitter acquisition can be reflected in our sentiment index: it shows a relevant impact on negative news compared with positive comments. This opens a new discussion: Why does negative news have a greater relevance on price stock than positive comments?

An increasing amount of research highlights the human inclination to favor negative news over positive. But what drives this preference? Stuart Soroka proposes that this bias might stem from neurological or physiological predispositions, as the potential risks associated with negative information greatly exceed the potential advantages of positive information (Soroka, 2015).

However, sharing news content on social media differs fundamentally from choosing and reading articles. A study mentioned by Arianna Huffington, cited in Soroka (2015), acknowledges this distinction. While we might be more inclined to share positive content on social media, our tendencies when it comes to reading news still lean towards prioritizing negative information. Ultimately, there is an expanding pool of evidence that underscores humans' tendency to favor negative over positive news content (Soroka, 2015).

7. Conclusions

This study underscores the innovative use of social media sentiment, particularly X (formerly Twitter) data, to assess its influence on stock prices, highlighting the transformative role of artificial intelligence in analyzing vast volumes of data from diverse sources rapidly. This approach enables investors to identify positive or negative market signals more efficiently. The research pivots around Twitter's role in the current information ecosystem, emphasizing the evolving nature of news analysis in the digital age, influenced by the Internet and social media. By leveraging Twitter data to construct a sentiment index and applying an autoregressive vector (VAR) model, the study explores the bidirectional relationship between this sentiment index and the S&P 500 index returns, through the Granger causality.

The introduction of news sentiment from social media into the analysis is novel in two main respects: showcasing the impact of social commentary on stock performance and employing natural processing language (NPL) for the automated text analysis. The specific case of Elon Musk's acquisition of Twitter and its impact on the S&P 500 is examined, aiming to reveal the relationship between market returns and aggregate sentiment. This research contributes to the understanding of market dynamics. An example is the announcement of Silicon Valley Bank's financial problems. This bank went bankrupt due to poor performance of its risk management strategy. In this sense, the methodology in this paper offers valuable insights for investors and regulators by presenting a replicable methodology for sentiment analysis across different media outlets. This can be used by investors analyzing market sentiment to detect movements in the value of assets that can be generated from a specific event. An application for public policy can be made when an abnormal event appears and the bank needs to be rescued, regulations should be put in place. It can be a part of bankruptcy prediction tools (Kliestik et al., 2018).

Key findings include the identification of a significant, negative correlation between the composite sentiment index and the S&P 500's performance, highlighting the predominance of negative sentiment in influencing market behavior around specific events. The absence of Granger causality between the sentiment index and S&P 500 returns suggests complex dynamics in how sentiment reflects or predicts market movements.

After applying the methodology, we accept H0—the composite sentiment has a negative effect on S&P 500 returns.

The discussion reflects on Twitter's changing role amid the rise of short video platforms like TikTok, noting Twitter's decline in user growth and its implications for the platform's advertising revenue and market position. Musk's acquisition is contextualized within this shifting landscape, suggesting his motives may extend beyond financial gain to include influencing public opinion and leveraging Twitter for political and personal branding purposes.

On the one hand, a limitation of this research is that the autoregressive vector technique remains a linear technique and does not capture complexity when there is no linearity. On the other hand, the absence of the Granger causality, which is a concept that refers to linearity, allows us to consider other types of models for future studies regarding prediction.

The Anáhuac Journal, **Vol. 24, núm. 1, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n1.08

The study also contemplates the broader market sentiment captured by the S&P 500 regarding Twitter's acquisition, opening discussions on the disproportionate impact of negative versus positive news on stock prices. This observation aligns with broader psychological research indicating a human predisposition towards negative information, a trend that persists in social media engagement and news consumption patterns.

References

Akhtaruzzaman, M., Boubaker, S., & Goodell, J. W. (2023). Did the collapse of Silicon Valley Bank catalyze financial contagion? *Finance Research Letters*, 56, 104082. https://doi.org/10.1016/j.frl.2023.104082

Broadstock, D. C., & Zhang, D. (2019). Social-media and intraday stock returns: The pricing power of sentiment. *Finance Research Letters*, *30*, 116-123. https://www.sciencedirect.com/science/article/pii/S1544612318307888

Brooks, C. (2019). *Introductory econometrics for finance*. Cambridge University Press.

Derakhshan, A., & Beigy, H. (2019). Sentiment analysis on stock social media for stock price movement prediction. *Engineering Applications of Artificial Intelligence*, 85, 569-578. https://www.sciencedirect.com/science/article/pii/S0952197619301666

Feuerriegel, S., Heitzmann, S. F., & Neumann, D. (2015, January). Do investors read too much into news? How news sentiment causes price formation. In the *48th Hawaii International Conference on System Sciences*, IEEExplore, 4803-4812. https://ieeexplore.ieee.org/abstract/document/7070391

Gidófalvi, G. (2001). Using news articles to predict stock price movements. Department of Computer Science and Engineering, University of California, San Diego. https://people.kth.se/~gyozo/docs/financial-prediction.pdf

Heston, S. L., & Sinha, N. R. (2016). News versus Sentiment: Predicting Stock Returns from News Stories. (FEDS Working Paper No. 2016-48). *Finance and Economics Discussion Series*, 1–35. https://doi.org/10.17016/feds.2016.048

Jia, J., Pan, H., & Su, J. (2023). Analysis of Motivations, Process, and Implications of Elon Musk's Acquisition of Twitter. *BCP Business & Management*, *47*, 145-153. https://doi.org/10.54691/bcpbm.v47i.5185

Jia, Q., & Xu, S. (2022). An Overall Analysis of Twitter and Elon Musk M&A Deal. *Highlights in Business, Economics and Management*, *2*, 436-441. https://doi.org/10.54097/hbem.v2i.2399

Kliestik, T., Misankova, M., Valaskova, K., & Svabova, L. (2018). Bankruptcy prevention: New effort to reflect on legal and social changes. *Science and Engineering Ethics*, *24*, 791-803. https://doi.org/10.1007/s11948-017-9912-4

Lee, H., Lee, N., Seo, H., & Song, M. (2020). Developing a supervised learning-based social media business sentiment index. *The Journal of Supercomputing*, 76, 3882-3897. https://doi.org/10.1007/s11227-018-02737-x

Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., & Chen, Y. (2014). The effect of news and public mood on stock movements. *Information Sciences*, *278*, 826-840. https://doi.org/10.1016/j.ins.2014.03.096

Lugmayr, A. (2012). Predicting the future of investor sentiment with social media in stock exchange investments: A basic framework for the DAX performance index. In M. Friedrichsen, & W. Mühl-Benninghaus (Eds.), *Handbook of social media management: value chain and business models in changing media markets* (pp. 565-589). Springer. https://doi.org/10.1007/978-3-642-28897-5_33

Manda, V. K. (2023). The Collapse of Silicon Valley Bank. *MAR-Ekonomi: Jurnal Manajemen, Akuntansi Dan Rumpun Ilmu Ekonomi*, *2*(1), 59-70. https://jurnal.seaninstitute.or.id/index.php/marekonomi/article/view/232

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

Nam, K., & Seong, N. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, *117*, 100-112. https://doi.org/10.1016/j.dss.2018.11.004

Núñez-Mora, J. A., & Mendoza-Urdiales, R. A. (2023). Social sentiment and impact in US equity market: an automated approach. *Social Network Analysis and Mining*, *13*(1), 111. https://doi.org/10.1007/s13278-023-01116-6

O'Shaughnessy, J. P. (2006). *Predicting the markets of tomorrow: A contrarian investment strategy for the next twenty years*. Penguin. https://shorturl.at/hxHIV

Pandey, D. K., Hassan, M. K., Kumari, V., & Hasan, R. (2023). Repercussions of the Silicon Valley Bank collapse on global stock markets. *Finance Research Letters*, *55*, 104013. https://doi.org/10.1016/j.frl.2023.104013

Radio 5 [@radio5_rne]. (2023, March 13). "Europa llama a la tranquilidad frente a la quiebra del Silicon Valley Bank. La Comisión recuerda que su presencia en Europa es muy limitada, pero aseguran que estarán muy pendientes." [Tweet] X (formerly Tweeter). https://twitter.com/radio5_rne/status/1635251865624674304

Ruan, Y., Durresi, A., & Alfantoukh, L. (2018). Using Twitter trust network for stock market analysis. *Knowledge-Based Systems*, 145, 207-218. https://www.sciencedirect.com/science/article/pii/S0950705118300248

Schindler, F. (2013). Predictability and persistence of the price movements of the S&P/Case-Shiller house price indices. *The Journal of Real Estate Finance and Economics*, *46*(1), 44-90. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2192128

Shah, D., Isah, H., & Zulkernine, F. (2018, December 10–13). Predicting the effects of news sentiments on the stock market. In *2018 IEEE International Conference on Big Data*, Seattle, WA, United States. https://doi.org/10.1109/BigData.2018.8621884

Smith, Molly (2023; May 23) Fed Tracks Market Sentiment With New Index Built from 4.4 Million Tweets. *Bloomberg*. https://www.bloomberg.com/news/articles/2023-05-23/fed-index-built-from-4-4-million-tweets-predicts-size-of-hikes?embedded-checkout=true

Soroka, S. (2015; May 25). *Why do we pay more attention to negative news than to positive news? British Politics and Policy at LSE* [blog entry]. https://eprints.lse.ac.uk/62222/

Steinert, L., & Herff, C. (2018). Predicting altcoin returns using social media. *Plos one*, 13(12), article 0208119. https://doi.org/10.1371/journal.pone.0208119

Tikkanen, M. (2021). Predicting the FTSE All-Share index daily close-to-close price direction using sentiment analysis on tweets from UK. LUT University. https://lutpub.lut.fi/handle/10024/163481

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139-1168. https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2007.01232.x

Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. W.W. Norton & Company. https://shorturl.at/adC17

Van Vo, L., & Le, H. T. T. (2023). From Hero to Zero: The Case of Silicon Valley Bank. *Journal of Economics and Business*, *127*, article 106138. https://doi.org/10.1016/j.jeconbus.2023.106138

Yousaf, I., & Goodell, J. W. (2023). Responses of US equity market sectors to the Silicon Valley Bank implosion. *Finance Research Letters*, *55*, article 103934. https://doi.org/10.1016/j.frl.2023.103934

About the author

Lizeth Gordillo Martínez began her professional career as coordinator for the Center for Technology and Financial Innovation while a BA student in international trade at the Tecnológico de Monterrey, Mexico City Campus. Gordillo Martínez conducted training workshops on databases specializing in finance and technology in various companies such as Bloomberg, Reuters, Economática, SaS, Eviews, and Numerix, and taught financial administration to 14 generations of students at the undergraduate level. Starting in 2008, she worked full-time leading negotiations with clients from the financial sector in the Latin America region at multinational companies with a technological and financial profile like Bloomberg, Thomson Reuters, Numerix, and Identy. Among her professional achievements are the creation of financial labs for private universities with Bloomberg and Reuters terminals, and the implementation of a Thomson Reuters ticker with financial indicators at the Mexican Stock Exchange (BMV) building. Gordillo Martínez secured Grupo Bancolombia as the first client—and first international bank—for Numerix in Latin America. There, she implemented a CVA calculation module for the bank's treasury. At the same time, she continued with her graduate studies, earning a master's degree in Finance, and in 2021 she started a PhD in financial sciences.

gordillo.lizeth@gmail.com
https://orcid.org/0009-0000-9210-4955