

Received: 28-07-2023 | Revised in Revision: 24-10-2023 | Accepted: 23-11-2023



Tweeting the economy: Analyzing social media sentiments and macroeconomic indicators

Insi Fitriani¹, Anna Amalyah Agus²
*Management Department, Faculty of Economics and Business,
Universitas Indonesia, Indonesia^{1,2}*

DOI: <https://doi.org/10.22219/jibe.v7i02.28227>

Abstract

This research aims to examine the correlation between social media sentiment and the Consumer Confidence Index (CCI) as well as Gross Domestic Product (GDP) in Indonesia. Data were collected through web scraping from Twitter (now also known as X) spanning from 2019 to 2022 on a monthly basis. Using Pearson and Kendall's Tau correlation tests, the study found that the correlation between Twitter sentiment and the CCI is not significant. However, there is a significant correlation between Twitter sentiment from news accounts and GDP. The findings indicate that the views and perceptions expressed in social media sentiment, particularly from news accounts on Twitter, could serve as an initial indicator of Indonesia's GDP.

Keywords: Social media; sentiment; Consumer Confidence Index; Gross Domestic Product

Introduction

In recent years, social media has emerged as one of the most powerful and influential tools for conveying messages, facilitating communication, and sharing information (Chu & Chen, 2019; Chwialkowska, 2019; Walsh, 2020). Initially utilized as a means to interact with friends and family, social media has evolved into a significant platform for collecting the latest views, perceptions, and societal experiences (Gan et al., 2020; Jain et al., 2021; Lou et al., 2019). Notably, platforms like Twitter, now also recognized as X, enable individuals to swiftly and easily express their thoughts, views, and opinions (Thompson, 2018; Yarış & Aykol, 2022), fostering an environment for self-expression, sharing perspectives on issues, and engaging in online discussions. The inclusivity of social media allows diverse groups and individuals to participate, creating a level playing field for the aggregation of diverse views and opinions. Moreover, social media serves as a significant alternative indicator for measuring consumer confidence (Ilhamalimy et al., 2021; Tjep et al., 2021; Yu et al., 2021). Through these platforms, consumers can articulate their views on various aspects of the economy that impact their daily lives. Complaints and comments regarding goods prices, inflation, or other economic issues, often shared on social media, offer real-time insights into consumer satisfaction or dissatisfaction with economic conditions (Domalewska, 2021; Hirata & Matsuda, 2023).

Conrad et al. (2019) discovered a relatively high correlation between the sentiment of tweets containing the word "jobs" and the Consumer Sentiment Index (CSI) during 2008-2009. However, the relationship began to decline after 2011, suggesting that such tweets may not have been as effective as expected as a surrogate for survey responses. Meanwhile, Shayaa et al. (2018) found a significant relationship between the Consumer Confidence Index (CCI) and social media sentiment on consumer

¹E-mail: insifitriani@outlook.com

²E-mail: anna@ui.ac.id

Copyright (c) 2023 Insi Fitriani, Anna Amalyah Agus



Creative Commons License

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

purchasing behavior. Furthermore, [Shayaa et al. \(2018\)](#) concluded that social media can provide large volumes of data about consumer beliefs, which can be analyzed more quickly and integrated with existing methods. In a different approach, [Ortega-Bastida et al. \(2021\)](#) proposed a multimodal approach to predict regional Gross Domestic Product (GDP). This approach successfully predicting regional GDP in Spain.

Considering Indonesia's status as a developing country with an archipelago base, measuring the CCI poses inherent challenges. Following the methodologies of [Conrad et al. \(2019\)](#) and [Shayaa et al. \(2018\)](#), this study explores the correlation of Twitter sentiment with CCI. We also argue that variations in results among previous researchers were influenced by the use of non-comprehensive keywords. In contrast to [Conrad et al. \(2019\)](#), who exclusively measured the correlation of social media sentiment with the CSI using the keyword "Job" and [Shayaa et al. \(2018\)](#) who used keywords "Car" and "Holiday" to measure CCI, our research selected 28 keywords to collect Twitter data based on a questionnaire related to the Bank Indonesia consumer survey. Assuming Twitter sentiment can predict CCI, as found by [Conrad et al. \(2019\)](#) and [Shayaa et al. \(2018\)](#), it should predict GDP, given that consumption is an integral part of GDP calculation. Therefore, following [Ortega-Bastida et al. \(2021\)](#), this research not only explores the correlation of Twitter sentiment with CCI but also its correlation with GDP.

Literature Review

Behavior Economics

Behavioral Economics is a branch of economics that examines the impact of psychological factors on economic decisions made by individuals or groups ([de Bruijn & Antonides, 2022](#); [Vlaev et al., 2019](#)). Notably, [Kahneman & Tversky \(1979\)](#), pivotal figures in Behavioral Economics, have proposed theories that offer valuable insights into consumer behavior. One of their key theories is Prospect Theory, which integrates two fundamental concepts: the nature of utility and rationality. Unlike normative economic theory, Prospect Theory elucidates how society behaves in the face of uncertainty ([Zielonka & Szymanek, 2023](#)). According to this theory, an individual's utility function follows an S-shaped pattern, also known as a concave utility function. This implies that consumers tend to be risk-averse when facing profits but become risk-seeking when facing losses. This pattern underscores that consumers are more responsive to changes in the value of their reference position than to changes in absolute values ([Ayaa et al., 2022](#); [Bougherara et al., 2021](#)).

Furthermore, [Kahneman & Tversky \(1979\)](#) introduced the concept of framing, which asserts that the presentation of information can influence consumer decisions. This concept holds particular relevance in marketing and promotional strategies, where pricing and promotional framing can impact consumer perceptions of value and purchasing decisions. The theory highlights that individuals often make irrational economic decisions due to cognitive and emotional biases. For instance, individuals are more averse to losing than they are inclined to enjoy equivalent gains, a phenomenon known as loss aversion. Referring to [Kahneman & Tversky \(1979\)](#), it is plausible that the noise on social media, especially Twitter, can serve as a proxy for understanding consumer behavior, and this concept is referred to as sentiment analysis.

Sentiment Analysis

Sentiment analysis, an advanced computational technique extensively discussed by scholars such as [Guenich et al. \(2022\)](#) and [Guo et al. \(2023\)](#), intricately traverses a vast expanse of individual opinions, evaluations, attitudes, and emotions. Its application extends beyond the confines of diverse entities, encompassing products, services, and organizations. Within this expansive scope, sentiment analysis encompasses multifaceted concepts, including opinion gathering, opinion analysis, and influence analysis, each contributing to a nuanced understanding of subjective expression in digital discourse ([López-Cabarcos et al., 2020](#); [Xiang, 2022](#)). Rooted in a combination of linguistic and computational approaches, this analytical method carefully dissects and categorizes sentiments, with a primary emphasis on discerning the polarity of these sentiments—both positive and negative. This approach, as delineated by [López-Cabarcos et al. \(2020\)](#) and [Xiang \(2022\)](#), yields invaluable insights into the intricacies of human sentiment within the digital communications landscape.

[Conrad et al. \(2019\)](#) investigated the viability of utilizing social media as an alternative to survey data. They observed a notable correlation between the sentiment in tweets containing the word "job" and survey-based measures of consumer confidence. Another study by [Shayaa et al. \(2018\)](#) delved into the correlation between the Consumer Confidence Index (CCI) and social media sentiment regarding consumer purchases of two product types during the two-year period from 2015 to 2016. [Shayaa et al. \(2018\)](#) identified a significant relationship between CCI and social media sentiment analysis, demonstrating that social media can yield substantial data on consumer trust. In a different approach, [Ortega-Bastida et al.](#)

(2021) applied a multimodal approach to predict regional GDP using Twitter data. This method successfully delivers GDP estimates at a higher frequency than official statistics and provides robust quarterly predictions.

Sentiment Analysis and Economic Indicators in Indonesia

Indonesia stands to derive substantial benefits from leveraging sentiment analysis as a tool for measuring key economic indicators such as the Consumer Confidence Index (CCI) and Gross Domestic Product (GDP). This proposition is rooted in consumer behavior theory, supported by empirical evidence demonstrating that consumer sentiment—a critical element of economic activity—is effectively captured through sentiment analysis of social media. Consumer behavior theory posits that sentiments, perceptions, and attitudes significantly influence purchasing decisions, thereby impacting economic outcomes. Social media, as a reflection of public discourse, serves as a rich, real-time data source summarizing these sentiments. Empirical evidence further supports the idea that fluctuations in consumer sentiment, as expressed on platforms like Twitter, are correlated with changes in economic indicators (Conrad et al., 2019; Shayaa et al., 2018).

In the distinctive context of Indonesia—a vast archipelagic nation—traditional methods of obtaining economic data may present challenges, resulting in data gaps and reduced accuracy. The decentralized nature of the country poses logistical challenges for timely and comprehensive data collection (Andiojaya et al., 2022). However, sentiment analysis provides a solution by offering a rapid and diverse understanding of public sentiment regarding economic conditions. This approach not only circumvents the limitations of traditional data collection methods but also enables a more dynamic and responsive assessment of economic indicators. By embracing sentiment analysis, researchers and authorities in Indonesia can overcome obstacles associated with obtaining accurate and timely economic data. The availability of social media data allows for quicker responses to emerging trends and sentiments, fostering a more agile and informed approach to economic analysis and policymaking. In essence, sentiment analysis is emerging as a valuable tool for enhancing the accuracy, speed, and comprehensiveness of measuring economic indicators in a geographically diverse and expansive country like Indonesia.

Research Method

The initial phase of the research procedure involves data collection. CCI and GDP data were sourced from the Bank Indonesia website. For tweet data, we utilized the "sncrape" and "tweepy" libraries in Python to scrape Twitter data, covering the period from January 2019 to December 2022. The selection of keywords for scraping aligns with the Indonesian bank consumer survey (refer to Table 1). Due to limited computing resources, the authors constrained data collection to 1,000 tweets per month per keyword. This limitation resulted in the collection of a total of 1,220,992 tweets in this first step. A comprehensive review of the collected tweet data was conducted to ascertain whether the tweet text contains the specified keywords. The data cleansing process aimed to ensure that the harvested keyword data is associated with legitimate users. Following data cleaning, 1,131,401 clean tweets were retained, meeting the specified criteria.

Data Labelling

In this stage, the labeled data is categorized into two groups: positive and negative. Positively labeled tweets convey positive thoughts or impressions towards specific terms, while negatively labeled tweets express negative ideas or criticism. The Lexicon-Based method, implemented with the "Literature" library in Python, is employed for this labeling process. After assigning a sentiment label to each tweet, researchers further categorized the tweets based on the account owner into three groups: All Accounts, News Accounts, and Non-News Accounts. The following is an explanation of these categories: a) All Accounts, category This is the original category that before tweets were grouped based on their account owner. All tweets in the dataset are included in this category automatically; b) News Accounts, this category includes tweets whose account owners are news media companies that have been registered and verified by the Indonesian press council. Using the "Lookup" function in Excel, researchers carried out a matching process with the database from the Indonesian press council; c) Non-News Accounts, this category includes tweets from accounts that do not correspond to news media companies from the Indonesian press council database (refer to Figure 1 for detail). The aggregation of tweet sentiment is measured using three methods, as referred to by O'Connor et al., (2010), with details following equations (1)-(3). Sentiment calculations

based on daily tweet data are then transformed into monthly and quarterly sentiment data through arithmetic addition. The determination of this period is adjusted to CCI and GDP data published monthly and quarterly.

$$\frac{P}{N} \tag{1}$$

$$\frac{P-N}{P+N} \tag{2}$$

$$\frac{P}{P+N} \tag{3}$$

Notes: In where P is positive sentiment and N is the negative sentiment

Table 1. Indicators and keywords

Indicator(s)	Keyword(s)	Number of Tweets Collected
Consumer Confidence	<i>Ekonomi</i> (Economics)	47,833
	<i>Gaji</i> (Wages)	47,950
	<i>Pekerjaan</i> (Job)	47,955
	<i>Pengangguran</i> (Unemployment)	47,934
	<i>Omset</i> (Turnover)	47,985
	<i>Bahagia</i> (Happy)	47,804
Economic Expectations	<i>Harga</i> (Price)	47,903
	<i>Cuan</i> (Profit)	32,581
	<i>Untung</i> (Profit)	47,945
	<i>Rugi</i> (Loss)	47,966
	<i>Boncos</i> (Loss)	32,702
	<i>Inflasi</i> (Inflation)	47,985
	<i>Pajak</i> (tax)	43,725
	<i>Tarif</i> (Rates)	47,920
Spending	<i>Tabungan</i> (Saving)	47,940
	<i>Belanja</i> (Spending)	47,915
	<i>Kredit</i> (Credit)	47,314
	<i>Diskon</i> (Discount)	47,503
	<i>Hiburan</i> (Entertainment)	47,873
	<i>Kendaraan</i> (Vehicle)	47,983
	<i>Konsumsi</i> (Consumption)	47,953
Financial Stability	<i>Perbankan</i> (Banking)	47,985
	<i>Bangkrut</i> (Bankrupt)	47,801
	<i>Penipuan</i> (Fraud)	47,957
	<i>Kredit Macet</i> (Credit default)	11,961
	<i>Resesi</i> (Recession)	32,123
	<i>Kurs Rupiah</i> (Rupiah Exchange Rate)	24,324
	<i>Moneter</i> (Monetary)	38,172
Total		1,220,992

Normality test

Before conducting the correlation test, a normality test is imperative to ensure that the data follows a normal distribution, serving as a prerequisite for subsequent parametric statistical tests. The Kolmogorov-Smirnov (K-S) method, executed through R-Studio software, is employed for this normality test. The K-S test, a non-parametric test, compares a sample distribution with a theoretical distribution or compares the distribution of two samples. The formula for the K-S test is as follows (equation 4), where $F_0(x_i)$ represents the theoretical cumulative distribution function, n is the number of observations, and x_i is the i th observation data that has been sorted. The null hypothesis of this test posits that the data conforms to a normal distribution. If the p-value resulting from the K-S test is smaller than the specified significance level of 0.05, then the null hypothesis is rejected, signifying that the data does not follow a normal distribution.

$$D = \max_{1 \leq i \leq n} (F_0(x_i) - \frac{i-1}{n}, \frac{i}{n} - F_0(x_i)) \tag{4}$$

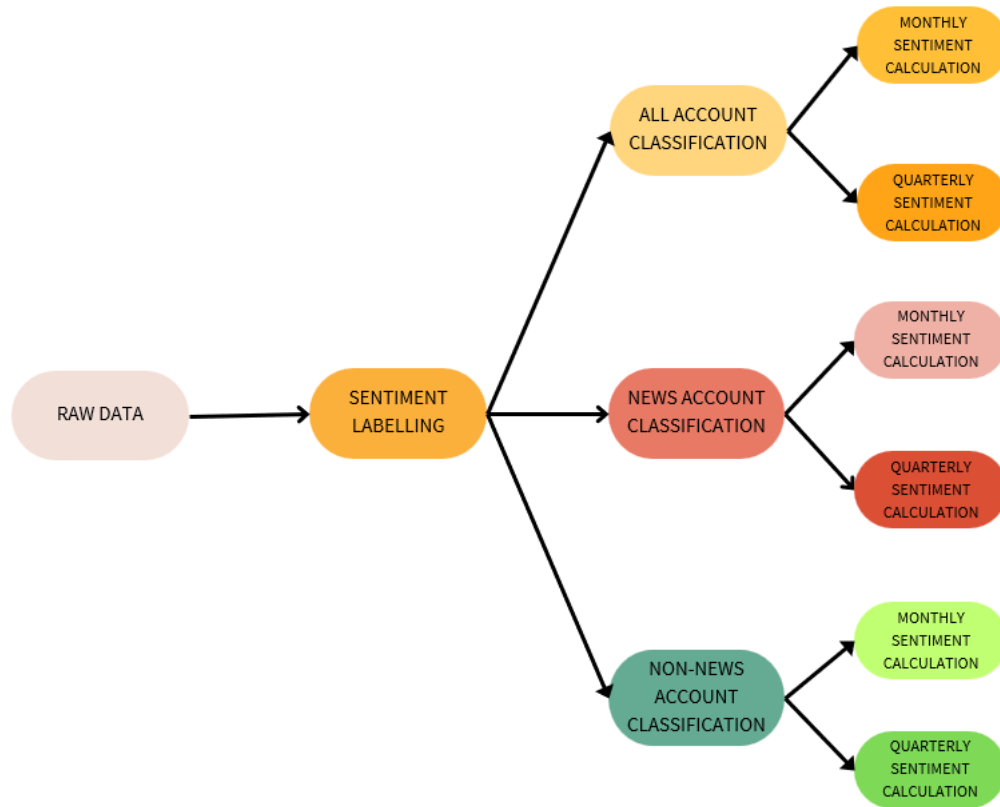


Figure 1. Sentiment calculation flow

Correlation Measurement

The Pearson product-moment correlation coefficient, introduced by Pearson in 1896, is utilized to quantify the degree to which two variables exhibit a linear relationship. It can be calculated using the following formula (equation 5), where \bar{x} and \bar{y} denote the sample means of x_i and y_i . The resulting r value falls within the range of -1 to 1. If $r > 0$, the variables are positively correlated; if $r < 0$, the variables are negatively correlated. An r value close to zero indicates a weak or no linear relationship between the variables.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \tag{5}$$

In addition to the Pearson product-moment correlation coefficient, researchers employ Kendall’s Tau (Kendall, 1938) to conduct a robustness test (refer to equation 6). If $\tau > 0$, the variables are positively correlated, and if $\tau < 0$, the variables are negatively correlated. This method provides an alternative approach for assessing the correlation between variables and is particularly robust in capturing monotonic relationships, even in the presence of outliers or non-normally distributed data.

$$\tau = \frac{\text{Number of concordant pairs} - \text{Number of discordant pairs}}{\text{Number of all possible pairs}} \tag{6}$$

Result and Discussion

The analysis of Table 2 reveals a prevalence of positive sentiment tweets over negative tweets based on the sentiment analysis process conducted on the dataset. Given that the Consumer Confidence Index (CCI) demonstrates a non-normal distribution according to the Kolmogorov-Smirnov test (see Table 3), the researchers opted to perform a normality transformation of the data using natural logarithms,

following the methodology outlined by Lee (2020). In this research, despite the primary focus being on measuring correlation, the normality test remains crucial. This is because data normality is a fundamental assumption that must be satisfied when applying Pearson's correlation coefficient, as emphasized by Khosravi et al. (2023) and van den Heuvel & Zhan (2022).

Table 2. Results of sentiment analysis

Period	Total Tweets per sentiment		Period	Total Tweets per sentiment	
	Negative	Positive		Negative	Positive
Jan-19	7,029	16,475	Jan-21	8,008	17,323
Feb-19	6,331	16,040	Feb-21	7,582	17,196
Mar-19	6,782	15,360	Mar-21	6,856	17,298
Apr-19	6,454	15,183	Apr-21	6,959	17,368
May-19	7,053	15,688	May-21	7,234	16,496
Jun-19	6,474	15,326	Jun-21	7,241	16,559
Jul-19	6,478	15,799	Jul-21	6,813	16,396
Aug-19	7,165	15,838	Aug-21	6,805	16,373
Sep-19	7,288	15,906	Sep-21	6,544	16,037
Oct-19	7,163	16,079	Oct-21	6,786	15,828
Nov-19	7,104	15,965	Nov-21	6,747	15,524
Des-19	6,648	16,433	Des-21	6,199	16,410
Jan-20	7,129	16,786	Jan-22	6,841	16,165
Febi-20	7,415	16,199	Febi-22	7,096	15,658
Mar-20	8,460	16,500	Mar-22	6,811	16,298
Apr-20	8,211	17,059	Apr-22	7,205	16,412
May-20	7,914	16,868	May-22	7,052	16,538
Jun-20	7,551	17,140	Jun-22	7,284	16,574
Jul-20	7,165	16,829	Jul-22	7,684	17,852
Aug-20	7,033	16,933	Aug-22	7,178	16,500
Sept-20	7,547	16,728	Sept-22	7,231	16,280
Oct-20	7,235	17,208	Oct-22	7,037	16,497
Nov-20	7,215	17,511	Nov-22	6,984	16,597
Des-20	6,968	17,653	Des-22	6,453	17,279

Table 3. Result of normality test

Variable(s)	D	p-Value	Distribution
Consumer Confidence Index	0.20095	0.0356	Not normally distributed
Gross Domestic Product	0.2109	0.4174	Normal
Consumer confidence index (Monthly)			
Sentiment All Accounts Method 1	0.076722	0.9195	Normal
Sentiment All Accounts Method 2	0.068471	0.9666	Normal
Sentiment All Accounts Method 3	0.068471	0.9666	Normal
Non-News Account Sentiment Method 1	0.094413	0.7501	Normal
Non-News Account Sentiment Method 2	0.080525	0.89	Normal
Non-News Account Sentiment Method 3	0.080525	0.89	Normal
News Account Sentiment Method 1	0.098389	0.7415	Normal
News Account Sentiment Method 2	0.1361	0.3362	Normal
News Account Sentiment Method 3	0.1361	0.3362	Normal
Gross Domesctic Product (Quarterly)			
Sentiment All Accounts Method 1	0.15972	0.7521	Normal
Sentiment All Accounts Method 2	0.16954	0.6861	Normal
Sentiment All Accounts Method 3	0.16954	0.6861	Normal
Non-News Account Sentiment Method 1	0.17501	0.6487	Normal
Non-News Account Sentiment Method 2	0.18371	0.5897	Normal
Non-News Account Sentiment Method 3	0.18371	0.5897	Normal
News Account Sentiment Method 1	0.094565	0.996	Normal
News Account Sentiment Method 2	0,10741	0.9829	Normal
News Account Sentiment Method 3	0,10741	0.9829	Normal

Correlation analyses from Tables 4 and 5 reveal that Twitter sentiment fails to adequately capture the Consumer Confidence Index (CCI) in Indonesia. This observation aligns with several studies demonstrating that while social media is pivotal in representing public opinion, it does not consistently exhibit a directly proportional impact on consumer confidence (Abbas et al., 2022; Li et al., 2022; Liu et

al., 2023; Pramananda et al., 2022). It is noteworthy, however, that sentiment originating from news accounts demonstrates a higher correlation compared to sentiment from other accounts. This underscores the notion that opinions expressed through news accounts tend to be more reliable (Ortega-Bastida et al., 2021; Sampietro & Salmerón, 2021).

Table 4. Results of Pearson correlation test

Correlation	Correlation coefficient	p-Value	Interpretation
Consumer confidence index (Monthly)			
Sentiment All Accounts Method 1	0.06	0.67	No significant correlation
Sentiment All Accounts Method 2	0.07	0.66	
Sentiment All Accounts Method 3	0.07	0.66	
Non-News Account Sentiment Method 1	0.06	0.67	
Non-News Account Sentiment Method 2	0.06	0.67	
Non-News Account Sentiment Method 3	0.06	0.67	
News Account Sentiment Method 1	0.13	0.39	
News Account Sentiment Method 2	0.13	0.40	
News Account Sentiment Method 3	0.13	0.40	
Gross Domestic Product (Quarterly)			
Sentiment All Accounts Method 1	0.35	0.19	No significant correlation
Sentiment All Accounts Method 2	0.35	0.18	
Sentiment All Accounts Method 3	0.35	0.18	
Non-News Account Sentiment Method 1	0.32	0.23	
Non-News Account Sentiment Method 2	0.32	0.22	
Non-News Account Sentiment Method 3	0.32	0.22	
News Account Sentiment Method 1	0.55	0.03	Significant
News Account Sentiment Method 2	0.55	0.03	
News Account Sentiment Method 3	0.55	0.03	

Notes: Pearson Correlation Test with 95% confidence interval

Table 5. Result of Kendall's Tau correlation test

Correlation	Correlation coefficient	p-Value	Interpretation
Consumer confidence index (Monthly)			
Sentiment All Accounts Method 1	0.04	0.70	No significant correlation
Sentiment All Accounts Method 2	0.04	0.70	
Sentiment All Accounts Method 3	0.04	0.70	
Non-News Account Sentiment Method 1	0.03	0.77	
Non-News Account Sentiment Method 2	0.03	0.77	
Non-News Account Sentiment Method 3	0.03	0.77	
News Account Sentiment Method 1	0.05	0.59	
News Account Sentiment Method 2	0.05	0.59	
News Account Sentiment Method 3	0.05	0.59	
Gross Domestic Product (Quarterly)			
Sentiment All Accounts Method 1	0.13	0.51	No significant correlation
Sentiment All Accounts Method 2	0.13	0.51	
Sentiment All Accounts Method 3	0.13	0.51	
Non-News Account Sentiment Method 1	0.13	0.51	
Non-News Account Sentiment Method 2	0.13	0.51	
Non-News Account Sentiment Method 3	0.13	0.51	
News Account Sentiment Method 1	0.53	0.00	Significant
News Account Sentiment Method 2	0.53	0.00	
News Account Sentiment Method 3	0.53	0.00	

Notes: Kendall's Tau Test with 95% confidence interval

Furthermore, our results indicate a significant and positive relationship between sentiment on Twitter, particularly from news accounts, and Gross Domestic Product (GDP). Conversely, sentiment from non-news accounts does not exhibit a significant relationship with GDP. This finding introduces a novel perspective and evidence compared to studies by Abbas et al. (2022), Jabeen et al. (2022), and Nia et al. (2022), suggesting that a positive relationship between Twitter sentiment and GDP is more likely when sentiment from news accounts is predominant. These findings underscore the pivotal role of social media,

particularly Twitter, as a platform for capturing public opinion that can be linked to macroeconomic indicators. This opens avenues for economic and business policymakers to leverage social media as an analytical tool in decision-making processes.

The results described here present several advantages and significant deviations from prior studies. In this comparative analysis, we emphasize several key points. This research is specifically tailored to Indonesia, a context characterized by unique social, economic, and social media features, thereby contributing valuable insights into the impact of social media sentiment on economic indicators, particularly from the perspective of developing countries. While Conrad et al. (2019) focused their research on the United States with limited keywords, we identified sentiment by involving 28 keywords reflective of consumer survey topics employed by the central bank, ensuring the collected data remains pertinent to critical issues within the Indonesian economic context. Additionally, this study distinguishes between news accounts and non-news accounts on Twitter, providing deeper insights into sentiment analysis—a differentiation absent in the works of Conrad et al. (2019), Shayaa et al. (2018), and Ortega-Bastida et al. (2021).

Moreover, these findings underscore the intricate interactions between social media, news, CCI, and GDP. Positive or negative sentiments posted via news accounts may set off a domino effect in shaping consumer perceptions and actions. Twitter news accounts, recognized as reliable sources of information, can significantly influence consumer perceptions of the economy. Consequently, sentiment emanating from news sources may impact consumer behavior, thereby influencing overall economic activity. Thus, these findings not only affirm the potential of social media in capturing public opinion but also highlight the likelihood of social media influencing consumer behavior. As a result, the development of a lag model becomes imperative for future research, particularly to substantiate the hypothesis that social media can reshape consumer behavior.

Conclusions, suggestions and limitations

Our examination of Twitter sentiment, the Consumer Confidence Index (CCI), and Gross Domestic Product (GDP) in Indonesia provides noteworthy findings. Twitter sentiment, particularly from news accounts, exhibits a robust predictive relationship with GDP, underscoring its potential as an analytical tool for early economic indicators. However, its effectiveness in discerning variations in the CCI is limited, aligning with previous research highlighting the nuanced role of social media in reflecting public opinion and its inconsistent influence on consumer confidence. The outcomes of this study offer valuable insights for policymakers, providing a basis for informed economic analysis and decision-making processes. The study suggests that understanding the distinct impact of social media sentiment from news accounts can contribute to more accurate predictions of economic trends.

Looking ahead, we advocate for future research to explore lag models, offering a more comprehensive assessment of social media's potential in reshaping consumer behavior over time. This could unveil deeper patterns and dynamics in the relationship between social media sentiment and economic indicators. In summary, our findings carry implications for both academia and practical policymaking, contributing to a richer understanding and application of social media in economic analyses.

Competing Interests

The author(s) declare that there are no competing interests relevant to the content of this article.

References

- Abbas, A. F., Jusoh, A., Mas'od, A., Alsharif, A. H., & Ali, J. (2022). Bibliometrix analysis of information sharing in social media. *Cogent Business and Management*, 9(1), 1–23. <https://doi.org/10.1080/23311975.2021.2016556>
- Andiojaya, A., Nefriana, R. R., Rizka, R., & Lestari, T. K. (2022). GRDP forecasting through night-time light data: Evidence from Indonesia. *Jurnal Inovasi Ekonomi*, 8(01). <https://doi.org/10.22219/jiko.v8i01.20239>
- Ayaa, M. M., Peprah, W. K., Mensah, M. O., Owusu-Sekyere, A. B., & Daniel, B. (2022). Influence of Heuristic Techniques and Biases in Investment Decision-Making: A Conceptual Analysis and Directions for Future Research. *International Journal of Academic Research in Business and Social Sciences*, 12(5), 1–17. <https://doi.org/10.6007/ijarbss/v12-i5/13339>
- Bougherara, D., Friesen, L., & Nauges, C. (2021). Risk Taking with Left- and Right-Skewed Lotteries. *Journal of Risk and Uncertainty*, 62(1), 89–112. <https://doi.org/10.1007/s11166-021-09345-w>

- Chu, S. C., & Chen, H. T. (2019). Impact of consumers' corporate social responsibility-related activities in social media on brand attitude, electronic word-of-mouth intention, and purchase intention: A study of Chinese consumer behavior. *Journal of Consumer Behaviour*, 18(6), 453–462. <https://doi.org/10.1002/cb.1784>
- Chwialkowska, A. (2019). How sustainability influencers drive green lifestyle adoption on social media: The process of green lifestyle explained through the lenses of the minority influence model and social learning theory. *Management of Sustainable Development Sibiu*, 11(1), 33–42. <https://doi.org/10.2478/msd-2019-0019>
- Conrad, F. G., Gagnon-Bartsch, J. A., Ferg, R. A., Schober, M. F., Pasek, J., & Hou, E. (2019). Social Media as an Alternative to Surveys of Opinions About the Economy. *Social Science Computer Review*, 39(4), 489–508. <https://doi.org/10.1177/0894439319875692>
- de Bruijn, E. J., & Antonides, G. (2022). Poverty and economic decision making: a review of scarcity theory. In *Theory and Decision* (Vol. 92, Issue 1). Springer US. <https://doi.org/10.1007/s11238-021-09802-7>
- Domalewska, D. (2021). An analysis of COVID-19 economic measures and attitudes: evidence from social media mining. *Journal of Big Data*, 8(1), 1–14. <https://doi.org/10.1186/s40537-021-00431-z>
- Gan, B., Alexeev, V., Bird, R., & Yeung, D. (2020). Sensitivity to sentiment: News vs social media. *International Review of Financial Analysis*, 67(1), 1–72. <https://doi.org/10.1016/j.irfa.2019.101390>
- Guenich, H., Hamdi, K., & Chouaibi, N. (2022). Asymmetric response of Investor sentiment to Economic Policy Uncertainty, interest rates and oil price uncertainty: Evidence from OECD countries. *Cogent Economics & Finance*, 10(1), 2151113. <https://doi.org/10.1080/23322039.2022.2151113>
- Guo, D., Zheng, Y., Wang, W., Hu, P.-N., Yang, Z., & Chen, Z. (2023). The impact of different sentiment in investment decisions: evidence from China's stock markets IPOs. *Economic Research-Ekonomska Istraživanja*, 36(1), 2113739. <https://doi.org/10.1080/1331677X.2022.2113739>
- Hirata, E., & Matsuda, T. (2023). Examining logistics developments in post-pandemic Japan through sentiment analysis of Twitter data. *Asian Transport Studies*, 9(1), 1–7. <https://doi.org/10.1016/j.eastsj.2023.100110>
- Ilhamalimy, R., Ridho, & Ali, H. (2021). Model Perceived Risk and Trust: E-Wom and Purchase Intention (the Role of Trust Mediating in Online Shopping in Shopee Indonesia). *Dinasti International Journal of Digital Business Management*, 2(2), 204–221. <https://doi.org/10.31933/dijdbm.v2i2.651>
- Jabeen, A., Yasir, M., Ansari, Y., Yasmin, S., Moon, J., & Rho, S. (2022). An Empirical Study of Macroeconomic Factors and Stock Returns in the Context of Economic Uncertainty News Sentiment Using Machine Learning. *Complexity*, 1(1), 1–18. <https://doi.org/10.1155/2022/4646733>
- Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review*, 41(1), 1–17. <https://doi.org/10.1016/j.cosrev.2021.100413>
- Kahneman, D., & Tversky, A. (2018). Prospect theory: An analysis of decision under risk. In *Experiments in Environmental Economics* (Vol. 1). <https://doi.org/10.2307/1914185>
- Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, 30(1/2), 81–93. <https://doi.org/10.2307/2332226>
- Khosravi, M., Duti, B. M., Yazdan, M. M. S., Ghoochani, S., Nazemi, N., & Shabanian, H. (2023). Multivariate Multi-Step Long Short-Term Memory Neural Network for Simultaneous Stream-Water Variable Prediction. *MDPI: Eng*, 4(3), 1933–1950. <https://doi.org/10.3390/eng4030109>
- Lee, D. K. (2020). Data transformation: A focus on the interpretation. *Korean Journal of Anesthesiology*, 73(6), 503–508. <https://doi.org/10.4097/kja.20137>
- Li, J., Wang, Y., & Liu, C. (2022). Spatial Effect of Market Sentiment on Housing Price: Evidence From Social Media Data in China. *International Journal of Strategic Property Management*, 26(1), 72–85. <https://doi.org/10.3846/ijspm.2022.16255>
- Liu, Q., Lee, W. S., Huang, M., & Wu, Q. (2023). Synergy between stock prices and investor sentiment in social media. *Borsa Istanbul Review*, 23(1), 76–92. <https://doi.org/10.1016/j.bir.2022.09.006>

- López-Cabarcos, M. Á., Pérez-Pico, A. M., Vázquez-Rodríguez, P., & López-Pérez, M. L. (2020). Investor sentiment in the theoretical field of behavioural finance. *Economic Research-Ekonomiska Istraživanja*, 33(1), 2101–2119. <https://doi.org/10.1080/1331677X.2018.1559748>
- Lou, C., Tan, S. S., & Chen, X. (2019). Investigating Consumer Engagement with Influencer- vs. Brand-Promoted Ads: The Roles of Source and Disclosure. *Journal of Interactive Advertising*, 19(3), 169–186. <https://doi.org/10.1080/15252019.2019.1667928>
- Nia, Z. M., Ahmadi, A., Bragazzi, N. L., Woldegerima, W. A., Mellado, B., Wu, J., Orbinski, J., Asgary, A., & Kong, J. D. (2022). A cross-country analysis of macroeconomic responses to COVID-19 pandemic using Twitter sentiments. *PLoS ONE*, 17(8 August), 1–16. <https://doi.org/10.1371/journal.pone.0272208>
- O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceedings of Fourth International AAAI Conference on Weblogs and Social Media*, 122–129.
- Ortega-Bastida, J., Gallego, A. J., Rico-Juan, J. R., & Albarrán, P. (2021). A multimodal approach for regional GDP prediction using social media activity and historical information. *Applied Soft Computing*, 111(1), 107693. <https://doi.org/10.1016/j.asoc.2021.107693>
- Pearson, K. (1896). VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 187, 253–318. <https://doi.org/10.1098/rsta.1896.0007>
- Primananda, P. G. B. N., Yasa, N. N. K., Sukaatmadja, I. P. G., & Setiawan, P. Y. (2022). Trust as a mediating effect of social media marketing, experience, destination image on revisit intention in the COVID-19 era. *International Journal of Data and Network Science*, 6(2), 517–526. <https://doi.org/10.5267/j.ijdns.2021.12.002>
- Sampietro, A., & Salmerón, L. (2021). Incivility in online news and Twitter: effects on attitudes toward scientific topics when reading in a second language. *Language Sciences*, 85, 101385. <https://doi.org/10.1016/j.langsci.2021.101385>
- Shayaa, S., Ainin, S., Jaafar, N. I., Zakaria, S. B., Phoong, S. W., Yeong, W. C., Al-Garadi, M. A., Muhammad, A., & Zahid Piprani, A. (2018). Linking consumer confidence index and social media sentiment analysis. *Cogent Business and Management*, 5(1), 1–12. <https://doi.org/10.1080/23311975.2018.1509424>
- Thompson, S. (2018). The Role of Social Media in Shaping Public Opinion. *Social Media and Society*, 4(2), 1–10. <https://doi.org/10.61100/tacit.v1i1.37>
- Tiep, N. C., Wang, M., Mohsin, M., Kamran, H. W., & Yazdi, F. A. (2021). An assessment of power sector reforms and utility performance to strengthen consumer self-confidence towards private investment. *Economic Analysis and Policy*, 69(1), 676–689. <https://doi.org/10.1016/j.eap.2021.01.005>
- van den Heuvel, E., & Zhan, Z. (2022). Myths About Linear and Monotonic Associations: Pearson's r , Spearman's ρ , and Kendall's τ . *American Statistician*, 76(1), 44–52. <https://doi.org/10.1080/00031305.2021.2004922>
- Vlaev, I., King, D., Darzi, A., & Dolan, P. (2019). Changing health behaviors using financial incentives: A review from behavioral economics. *BMC Public Health*, 19(1), 1–9. <https://doi.org/10.1186/s12889-019-7407-8>
- Walsh, J. P. (2020). Social media and border security: Twitter use by migration policing agencies. *Policing and Society*, 30(10), 1138–1156. <https://doi.org/10.1080/10439463.2019.1666846>
- Xiang, X. (2022). Investor sentiment, R&D spending and firm performance. *Economic Research-Ekonomiska Istraživanja*, 35(1), 6257–6278. <https://doi.org/10.1080/1331677X.2022.2048193>
- Yarış, A., & Aykol, Ş. (2022). The impact of social media use on restaurant choice. *Anatolia*, 33(3), 310–322. <https://doi.org/10.1080/13032917.2021.1931379>
- Yu, W., Han, X., Ding, L., & He, M. (2021). Organic food corporate image and customer co-developing behavior: The mediating role of consumer trust and purchase intention. *Journal of Retailing and Consumer Services*, 59(1), 1–10. <https://doi.org/10.1016/j.jretconser.2020.102377>
- Zielonka, P., & Szymanek, K. (2023). Unraveling Economic Choices. a Historical Perspective on the Intersection of Decision Science and Behavioral Economics. *Orbis Idearum*, 11(1), 33–56. https://ruj.uj.edu.pl/xmlui/bitstream/handle/item/317808/zielonka_szymanek_unraveling_economic_choices_2023.pdf?sequence=1&isAllowed=y