

Analyzing the Interplay Between Social Media Sentiment and Traditional Public Opinion in Politics

Seo jun

Korea University Business School, Seongbuk District, Seoul, South Korea.
seojunwoo@hotmail.com

Correspondence should be addressed to Seo jun : seojunwoo@hotmail.com

Article Info

Journal of Enterprise and Business Intelligence (<https://anapub.co.ke/journals/jebi/jebi.html>)

Doi: <https://doi.org/10.53759/5181/JEBI202505019>

Received 06 January 2025; Revised from 12 March 2025; Accepted 16 May 2025.

Available online 05 October 2025.

©2025 The Authors. Published by AnaPub Publications.

This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Abstract – In the present day, due to the growing number of Web 2.0 tools, users are producing vast quantities of data in a massive and constantly changing manner. Opinion mining or Sentiment analysis (SA) is an informative technique to help obtain useful information from users' data automatically. Over the years, a number of SA challenges have been solved using deep learning techniques which has led to the achievement of state-of-art performances. Thus, it is necessary to solicit help to help the researchers to learn the present progress and outstanding issues to be solved quickly. This paper discusses the multidimensional nature of public opinion in South Korea through analyzing the entire Korean tweets from January 1, 2023 to December 31, 2023. Using Twitter (X) API, we collected over 5 million tweets with emphasis on phrases related to South Korean president and significant events in the country. Through the application of set filters, the authors were able to arrive at a dataset of approximately 4 million tweets. These developed tweets were then taken through rigorous preprocessing in order to make them ready for SA. In the current work, the BERT model was chosen to act as the major focus of the study. This model was particularly trained on a given tagged dataset of Korean text for the purpose of classifying the text's sentiment. The findings of the present study reveal complex patterns of the processes that define the nature of online sentiment, thus revealing that several forces work in concert to form public opinion. Despite certain concordances between the results of the online SA and the offline public opinion polls on presidential job performance approval, the relationship is rather weak.

Keywords – Sentiment Analysis, Sentence-Level Sentiment Analysis, Supervised Deep Learning Methods, Bidirectional Encoder Representations from Transformers Model, Word Opinion Mining.

I. INTRODUCTION

Currently, social media is a widely used technology that utilizes micro-blogging platforms to establish connections among millions of individuals [1]. People are free to express ideas, opinions, and concepts in the form of short messages commonly referred to as “tweets” in several microblogging sites accessible in social networks (like Twitter (X)), in web forums and even company websites [2]. Scientists collect these unorganized tweets and employ various techniques to extract information from them. This analysis of tweets or comments offers forecasts or assessments in various application fields, including business, telecommunication, government, biomedicine, education, tourism, and sports, services [3]. Opinion mining or Sentiment Analysis (SA) is the branch of study that is used to analyze and predict sentiments. Sentiment analysis (SA) [4] is a text mining technique that uses Natural Language Processing (NLP) techniques to classify binary data. SA can be performed at four different levels based on the degree of the text analysis, and these are sentence-level, word level, document-level, and aspect-level [5]. Document level SA [6] categorizes the overall opinion of a document regarding a single entity as either positive or negative. Sentence-level sentiment analysis [7] categorizes the opinion presented in a sentence as either negative or positive. Aspect-level sentiment analysis involves categorizing attitudes about entities based on distinct elements of those items.

Popularly utilized approaches for analyzing public sentiment include CNN and LSTM, BiLSTM, GRU, BiGRU, Capsule, and Capsule-based attention [8][9]. Using deep learning for sentiment analysis faces several challenges that require meticulous attention to enhance and ensure the dependability of the method. An important obstacle is the acquisition of a sufficient quantity of high-quality labeled data, which is a vital necessity for efficiently training strong models. Overfitting, a prevalent problem, necessitates the utilization of techniques like as data augmentation and regularization to avoid models

from achieving incomparable presentation on the training database but performing poorly on unseen data [10]. The process of modifying the models to suit different domains presents difficulties as a result of variations in language, utterances, and indicators of sentiment. Understanding the setting and nuances of language, particularly when it comes to negations or modifiers, can be complex. Performing sentiment analysis on many languages presents a significant challenge, as models must understand various languages and their distinct linguistic features. The issues that need to be addressed include the uneven distribution of data, the ability to understand frameworks, actual-time dispensation while maintaining accuracy, ethical considerations to minimize biases, and assuring continual flexibility.

This research work therefore seeks to establish the possibility of using deep learning (DL) to extract and analyze sentiments from Twitter (X). This research work will propose these online sentiments against the traditional offline polling data especially Presidential approval ratings in South Korea. The project employs the BERT (Bidirectional Encoder Representations from Transformers) model [11] for sentiment analysis of the Korean texts posted by users on Twitter (X). This requires extensive data pre-processing where features like the punctuation marks, special characters etc. are removed to ensure that the data is clean and relevant. The research then transforms the daily aggregated sentiment data weekly and monthly to minimize the daily volatility and identify larger trends, thereby offering a clearer picture of public sentiment over time. Also, the study analyses the correlation between online sentiment data and the conventional polling data from Realmeter and Gallup. This comparison entails comparing the demographic variables that include the political affiliation, gender, and age to determine the effects of such factors on the correlation between the offline and online public opinion. The study additionally explores the impact of temporal changes by analyzing correlations with data that has been delayed in time.

Section II presents an appraisal of related works on SA of social media users towards political agendas using deep learning algorithms. In Section III, data collection methodology, sentiment analysis using BERT, and comparison with offline poll data have been presented. A critical discussion of the results has been provided in Section IV. Lastly, Section V summarizes the findings, and proposes future works in the field of sentiment analysis.

II. RELATED WORKS

Yue et al. [12] and Ravi and Ravi [13] defined Sentiment Analysis (SA) as a prevalent technique that is progressively employed to evaluate the emotions of social media users towards a particular topic. Data mining is the most often used method for conducting sentiment analysis. The main concept from Hagenau, Liebmann, and Neumann [14] is to utilize DL to ascertain investors' predictions regarding the company price and the general market by analyzing their messages. Lan et al. [15] choose the Deep Learning approach over data mining because in data mining, the most difficult challenge is to find and pick the most optimal characteristics, particularly when dealing with Big Data. Unlike data mining, a DL model [16] acquires features through the learning process. DL methods generate abstract representations, which allows them to remain unaffected by local variations in the input data.

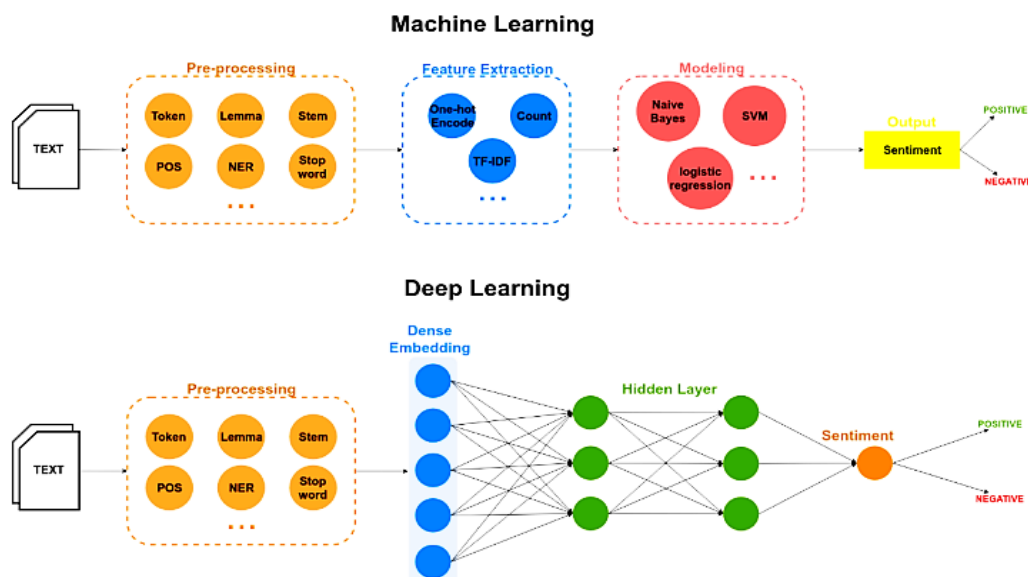


Fig 1. The Differences Between Machine Learning (ML) And DL Methods for Categorizing Sentiment Polarity

The top part of **Fig 1** represents ML, while the bottom part represents DL. Named Entity Recognition (NER) [17], Term Frequency-Inverse Document Frequency (TF-IDF) [18], and Part of Speech (POS) [19] are three techniques used in natural language processing.

Adnan and Akbar [20] argue that Deep Learning can provide more effective solutions for Big Data challenges such as semantic indexing, data tagging, and quick information retrieval. Deep Learning has the potential to employ a less complex model in order to achieve complex Artificial Intelligence (AI) tasks. According to Dargan et al. [21], while DL algorithms

have been effectively applied in domains like speech recognition and computer vision, their application in the setting of Big Data analysis remains limited. Khalil et al. [22] assesses the utilization of DL for SA of financial data. DL methods enable the extraction of intricate data by defining high-level properties with more abstraction in terms of lower-level properties with less abstraction. According to Litjens et al. [23], Deep Learning can effectively differentiate several sources of variance in data, such as light conditions, object forms, and object materials in an image. The concept of hierarchical learning in DL is derived from the key sensory areas of the neocortex in the human brain.

Sivaram, Das, and Venkatasubramanian [24] and Brahma, Wu, and She [25] argue that deep learning employs a multilayer technique when it comes to the hidden layers of a neural network. Conventional machine (ML) systems mostly rely on feature selection methods or human feature definition. Conversely, deep learning models has the ability to autonomously acquire and extract information, resulting in enhanced accuracy and performance. Typically, the hyperparameters of classifier models are automatically monitored. Many scholars have described the distinction between deep learning and conventional ML methods like Support Vector Machine (SVM) [26], Bayesian networks [27], or decision trees [28] in the context of sentiment polarity classification. Currently, deep learning is the most effective approach for tackling many issues in image and audio detection, natural language processing, and artificial neural networks.

According to Fitzgerald and Howcroft [29], there has been much research focused on categorizing polarity. A sentiment may be categorized as good, moderate, or negative [30]. SA may be used to identify and model topics, ideas, emotions, and social or political leanings, as well as polarity. The words opinion mining (OM) [31] and sentiment analysis are often used interchangeably in [32]. Generally, feelings may be analyzed at three levels: document, phrase, and aspect (thing). An analysis of the text's overall tone is conducted at the document level. The polarity of each phrase is assessed and categorized at the sentence level. Furthermore, the determination of an object or entity's polarity may be done at the aspect level [33]. Twitter (X) SA, also known as SA on Twitter (X), is a specific branch of sentiment analysis [34]. Due to the limited character count of tweets, there is little difference between document and sentence levels. There are two levels of analysis that may be used to Twitter (X) sentiment analysis: the sentence level, which is utilized for Tweets or messages, and the entity level [35]. Identifying sentiments on Twitter (X) poses several challenges that need to be solved. Twitter (X) tweets exhibit certain language idiosyncrasies, such as a more informal writing style and limitations on length, which distinguish them from other modes of communication, such as blog entries or forum postings.

According to Khurana et al. [36], NLP is a branch of AI that emphasizes on the study and advancement of models and algorithms for understanding and processing human language. Linguistics allows us to develop software that can examine and imitate comprehension of natural languages [37][38]. The primary domains of natural language processing, including machine scanning of texts, speech discrimination, automated text or voice production, machine translation, and question understanding and answering, need a model. Instead of constructing a model from the ground up to address a comparable issue, you use a pre-trained model from a different problem as a starting point. There are many categories of pre-trained models. Some models, like ELMO [39] and GPT [40], analyze text in a left-to-right way. On the other hand, BERT is a bidirectional model that examines text from both the left and right sides. BERT is a pre-inculcated framework that is used for feature-based tasks and fine-tuning. According to Santos, Marcacini, and Rezende [41], BERT relies on the transformer architecture, which is made of two components: a decoder and an encoder. BERT only utilizes the encoder component while disregarding the decoder component. The encoder has two components: self-recognition and a feed-forward neural system. BERT comes in two sizes: BERT basic and BERT big. The tokens provided as input to BERT undergo a conversion process into numerical vectors, which is achieved by the use of embedding methods.

This research aims to fill the gap by conducting a comparative analysis of popular opinion attained from social media webs like Twitter (X) and conventional offline polling methodologies. Although conventional polls have been widely used to gauge public opinion, their effectiveness is sometimes hindered by factors such as small sample sizes, infrequent data collection, and inherent biases in survey methods. On the other hand, social media provides a large, up-to-date, and unrequested pool of public opinion, but its accuracy and dependability as a gauge of wider popular attitude have not been well investigated. This work addresses the aforementioned research gap by using sophisticated deep learning methodologies, particularly the BERT model, to examine Korean tweets and juxtapose these results with well-established polling data from Realmeter and Gallup. The study seeks to gain an inclusive comprehension of the relationship between two data sources by analyzing their correlations and differences across different demographic factors and temporal adjustments. This analysis aims to determine the extent to which online sentiments accurately reflect traditional measures of public opinion. The findings will provide valuable insights into the potential integration of social media analysis into public opinion research.

III. DATA AND METHODS

Data Collection

In the first phase of this research, we aimed to gather a vast number of Korean tweets during the year 2023. This involved using the extended functionalities of the Twitter (X) API to search for tweets that contain specific words that were selected very carefully. These keywords were chosen to cover almost all possible topics, such as the South Korean president and his activities, the major political processes, important events in the country, and other topics. By using this particular strategy, we wanted to be as inclusive as possible of tweets that were directly relevant to the research goals, thereby providing a solid foundation for the subsequent analysis of the trends in public sentiment within the context of South Korea. The process of data collection provided a large sample of tweets, which at the beginning of the analysis included more than 5 million

individual tweets. However, aware of the fact that data quality is more valuable than the quantity, we used strict filters to clean the dataset and make it more appropriate for sentiment analysis. There was one important step in this process of filtration, and that was the removal of retweets from the analysis, as we were interested in the first-hand user opinions only. Retweets as opposed to being original tweets are more of a rehash of existing tweets, they do not offer the same level of detail or variety as original tweets, and as such are less useful for sentiment analysis. Also, all the tweets that contain non-Korean text were excluded systematically from the dataset in order to keep the analysis clean and consistent for the following processes.

Moreover, to reduce the interference of noise, only those tweets that have more than five characters have been considered for analysis. Even when such tweets may contain some useful information at times, most of the time they are likely to be spam or irrelevant messages or noise, which may distort the sentiment analysis results. By using these strict filtering conditions, we were able to obtain a highly selective sample of around 4 million tweets, which would be used in the subsequent analysis, and it was reasonable to assume that the rest of the tweets were genuine user-generated content within the Korean Twitter (X)sphere. Having the improved dataset, the next step was to filter the gathered data to make it suitable for sentiment analysis. This preprocessing phase entailed a process of a set of systematic steps that sought to bring the text data to a common format and structure that would allow for a sound sentiment analysis. The initial operation in this process was the tokenization process in which the text data was divided into meaningful subparts like words and phrases using a Korean language tokenizer. It was crucial for the process of the text analysis since it allowed the further analysis at the level of the lexical units.

After the tokenization process, the dataset was pre-processed to eliminate Korean keywords which are words that are prevalent in the Korean language but are not very useful in sentiment analysis. These are words that do not add much meaning to the text and are often used in large numbers; by removing them we hoped to minimize noise and enhance the sentiment analysis. Moreover, all sorts of punctuations, numerals, and special characters were also excluded from the analysis systematically to eliminate noise and maintain the quality of text data. Last of all, the text preprocessing was conducted with the help of the text normalization to eliminate the differences in spelling. This entailed the process of capitalizing all letters to remove inconsistency in capitalization and correct the spelling of words to reduce variation in the use of similar words. Through performing the above-mentioned preprocessing steps on the text data, the setup and edifice of the data were normalized and made more homogenous to allow for accurate sentiment analysis, which forms the basis of understanding trends in public sentiment in the South Korean Twitter (X)sphere during the specified time period.

Sentiment Analysis Using BERT

The foundation of our sentiment analysis was based on the BERT (Bidirectional Encoder Representations from Transformers) model which was trained on the Korean labelled corpus for the sentiment classification. The BERT model then takes the cleaned tweets and gives an evaluation of how likely the tweet is to be negative or positive. The sentiment score $S(t)$ for each tweet t is calculated using Equation (1).

$$S(t) = P_{positive}(t) - P_{negative}(t) \quad (1)$$

where $P_{positive}(t)$ and $P_{negative}(t)$ are the predicted probabilities of the tweet t being positive or negative, respectively. To quantify the overall sentiment for a given time period, we used Equation (2).

$$W \frac{1}{N_d} \sum_{i=1}^{N_d} S(t_i), \frac{1}{N_w} \sum_{i=1}^{N_w} S(t_i), \frac{1}{N_m} \sum_{i=1}^{N_m} S(t_i) \quad (2)$$

where N_d , N_w , and N_m are the number of tweets in a day, week, and month, correspondingly.

Since the study seeks to capture the sentiment across various periods, we combined the daily sentiment scores into weekly and monthly averages. The formula used to aggregate the sentiment scores is as provided in Equation (3) and (4).

$$Aggregated\ Sentiment = \frac{1}{N} \sum_{i=1}^N S(t_i) \quad (3)$$

$$S_d = \frac{1}{N_d} \sum_{i=1}^{N_d} P_{positive}(t_i) - P_{negative}(t_i) \quad (4)$$

where N represents the total amount of tweets in the given period, N_d is the number of tweets on day d , and $S(t_i)$ is the sentiment score of tweet t_i .

Comparison with Offline Poll Data

The offline polling data of presidential approval ratings were collected from Realmeter and Gallup. These polls give weekly and monthly approval ratings of the President of South Korea. In order to carry out correlations, we synchronized the online sentiment data with the offline polling dates and then computed the correlation coefficients between the two sets. So as to find out the connection among online sentiment and offline polling data, Pearson correlation coefficients were calculated on

daily, weekly and monthly data. Furthermore, we conducted a time-lag analysis to determine the extent to which the online sentiments can predict the offline polls in the future. The coefficient of correlation r is defined as in Equation (5).

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (5)$$

where X and Y represent the online sentiment scores and offline approval ratings, respectively. We also tested the reliability of the results for different age, gender, and political orientation subgroups. This analysis entailed categorizing the offline poll data and then comparing it to the online sentiment in the same categories. **Table 1** shows the Pearson correlation coefficients, standard deviations, and means of the variables.

Table 1. Statistics and Correlation Coefficients of Tweets

Statistic/Correlation		Value	Realmeter (Positive)	Realmeter (Negative)	Gallup (Positive)	Gallup (Negative)
Total Tweets		4,000,000				
Tweets	Positive	2,100,000				
	Negative	1,900,000				
Date Range		Jan 1 - Dec 31, 2023				
Sentiment	Daily	Mean: 0.02	0.16	0.13	-0.10	0.07
	Weekly	Mean: 0.03	0.20	0.18	-0.08	0.10
	Monthly	Mean: 0.04	0.25	0.22	-0.05	0.12
Age Group	18-29		0.12	0.10	-0.05	0.08
	30-39		0.185	0.233	-0.03	0.11
	40-49		0.14	0.17	-0.02	0.09
	50-59		0.15	0.19	-0.04	0.10
	60+		0.21	0.176	0.282	0.369
Gender	Male		0.302	0.411	-0.08	0.15
	Female		0.18	0.14	0.32	0.283
Political Ideology	Conservative		0.11	0.09	0.27	0.22
	Progressive		0.09	0.07	-0.06	0.05
	Neutral		0.310	0.327	-0.04	0.08

IV. RESULTS AND DISCUSSION

South Korea reached a crucial point on April 10, 2024, when it conducted its 22nd general election [42]. The opposition Democratic Party (DP) won the most recent general election for the 300-seat National Assembly, obtaining 175 seats and thereby controlling 58.33% of the legislature [43]. Meanwhile, the ruling People Power Party (PPP) successfully obtained 108 seats, which corresponds to 36% of the legislature. This election, widely seen as a pivotal midterm assessment of President Yoon Suk-yeol's administration, has been construed by many media sources as a distinct demand from voters for more discernment in governance [44]. Populism has evolved in South Korea, including many political ideologies, and is being strategically used to leverage popular social media opinion for electoral gain [45]. This work uses supervised DL to abstract public sentiment from the gathered tweets.

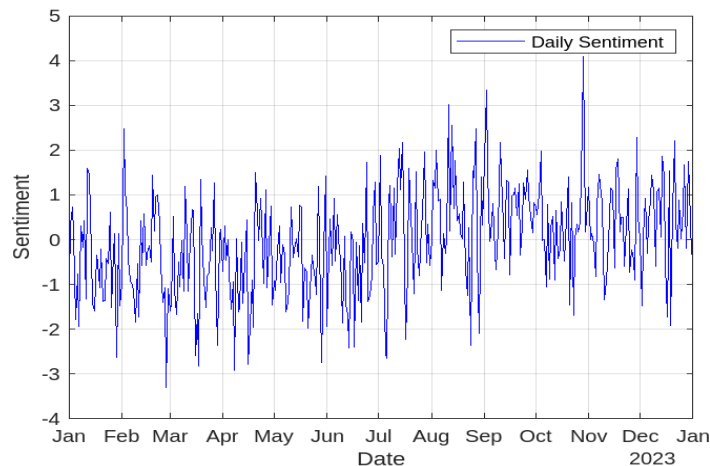


Fig 2. Daily Aggregated Sentiment in 2023

The neural system, which has been trained, analyzes Twitter (X) messages and produces predictions about the sentiment expressed in them [46]. The BERT model does SA on user-produced Korean texts, as described in [47]. Prior to being included into the network, the data need a process of purification. This preprocessing phase involves eliminating redundant words, punctuation marks, and special characters that do not provide any relevant information in determining the user's attitude. The adapted BERT framework computes the likelihood of a tweet's sensibility being inclined towards either a negative or positive emotion. **Fig 2** displays the time-sequences chart of daily aggregated sentiment in 2023. It reveals that there were 190 days with mostly adverse feelings and 175 days with predominantly positive thoughts. This tendency has a very high level of instability, which makes it difficult to identify any probable pattern.

Fig 3 and **4** demonstrate the weekly and monthly changes in the projected feelings, respectively. The graphs demonstrate a substantial reduction in volatility when measured on a daily basis. Although the online public initially had a favorable mood in the first half of 2023, there was a shift in the overall opinion throughout the second half. The first half of the year had some good occurrences, notably the third summit conference among North and South Korea on 18 September 2018. The optimistic atmosphere created by the ongoing interaction between the two countries persisted until 2023.

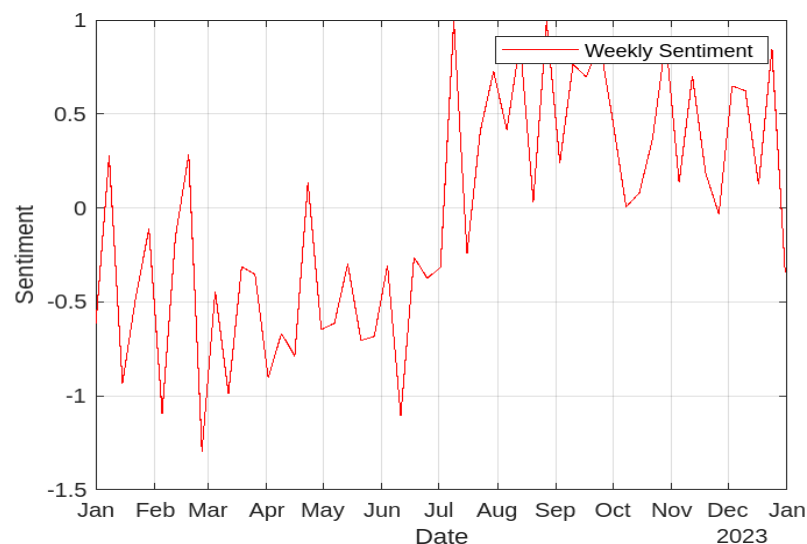


Fig 3. Weekly Transformation of Online Sentiment In 2023

Subsequently, the heads of state of the North Korea and United States unexpectedly convened in the Joint Safety Area of Pannumjeom shortly after the G20 summit in 2019. This succession of international events generated a favorable mood amidst the internet populace. In addition, 2019 saw the celebration of Independence Movement Day, Samiljeol's 100th anniversary, the Oscar win for the film *Parasite*, and a second-place finish by one of South Korea's national soccer teams in the first part of the U20 World Cup. Nevertheless, the latter part of 2023 started with Japan implementing export restrictions on semiconductor resources, which are sold by Japanese corporations to South Korea. August marked the appointment of Guk Cho, a university professor, as the minister of the Justice Department by South Korean President Moon Jae In, causing a great deal of controversy because of worries that his family would abuse his position of power. In December, the administration implemented rigorous rules on the real estate sector.

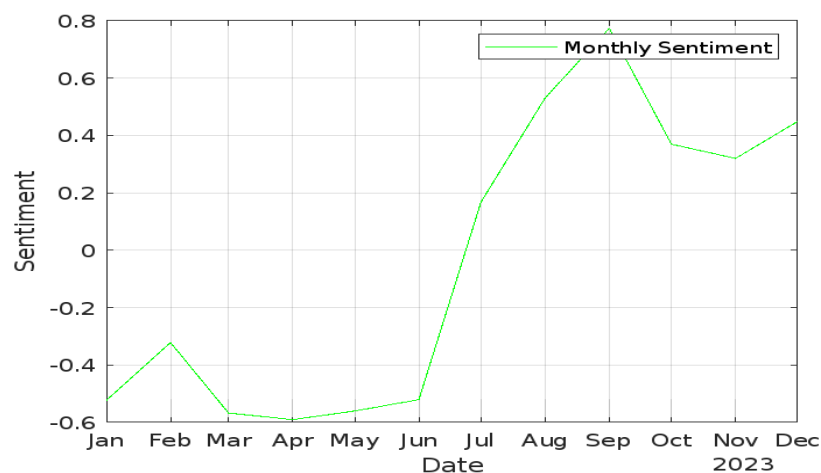


Fig 4. Monthly Transformation of Online Sentiment In 2023

Furthermore, the National Assembly enacted legislation to establish an autonomous investigative body with a focus on senior government officials. However, the occurrences during the latter half of 2023 are more contentious than those in the first half [48]. The internet public's perception of the president was greatly influenced by the disputes surrounding the former minister of justice and the establishment of a committee to investigate high-ranking government officials [49]. The research investigates the associations among offline and online indicators of presidential approval. Gallup and Realmeter, two polling firms, consistently provide the public with updates on the level of support that the South Korean president receives.

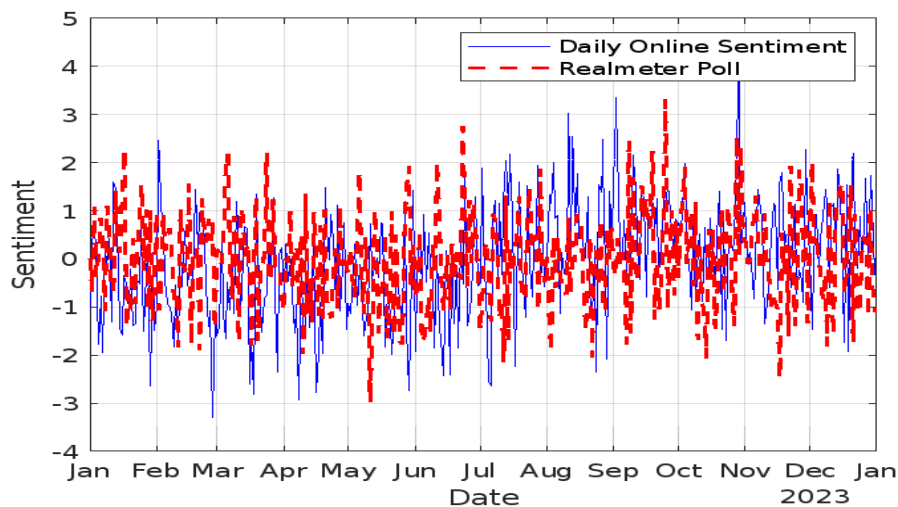


Fig 5. Comparison of Daily Online Sentiments and Realmeter's Poll

Fig 5 displays the online and offline measurements of presidential approval ratings. Neither polling firm has regular data about presidential approval ratings. Realmeter conducts an ample number of daily polls in comparison to the daily internet sentiment standards. **Fig 5** illustrates the correlation between the online sentiment and the daily survey conducted by Realmeter. Both shown trends exhibit significant volatility. Cotten and Gupta [50] describes the relationships between the daily data obtained online and offline. The correlation coefficients among the Realmeter's and online trend survey are 0.16 for adverse and 0.13 for positive sentiment. While the overall correlations for the whole year of 2023 are near to zero, certain time periods, like January and February, show a similar pattern where public views in both contexts follow similar pathways. The sentiment experiences an upward trend towards the conclusion of January and February.

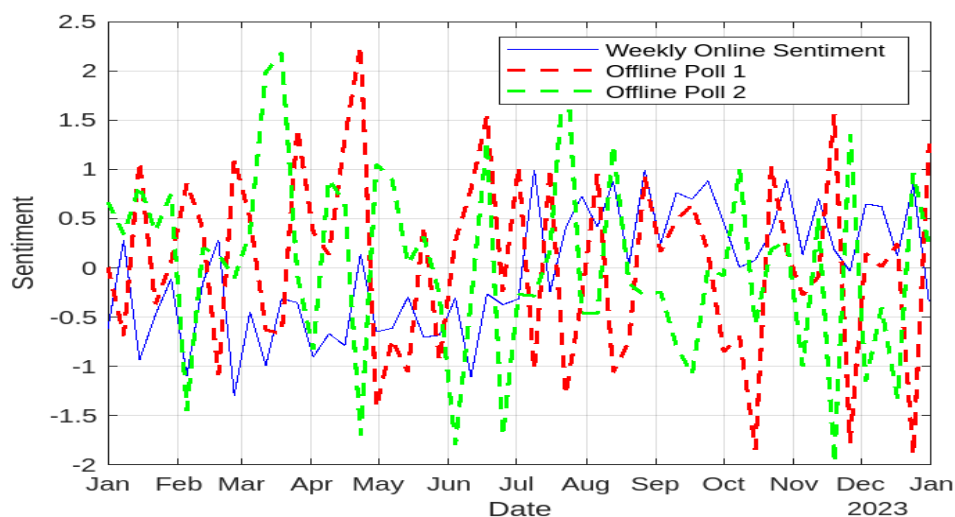


Fig 6. Weekly Fluctuations in Popular Sentiments Towards the President

The level of positive attitude experiences a big decline at the start of March, but undergoes a substantial rebound by the conclusion of June. Based on the survey, there is a notable surge in favorable mood seen on online platforms. Online and offline data exhibit basic disparities and polls contain a limited number of individuals in a sample, but internet assortments do not execute any limitations on the number of tweets a computer may gather [51]. The overall number of messages varies daily, leading to an inherent rise in volatility. Polls conduct polls across numerous days in order to obtain more reliable and consistent data. For instance, Realmeter and Gallup use a 3-day moving average to calculate figures (Gallup sometimes resorts to a 2-day moving average contrary to a 3-day statistic when the review period contains a vacation). In order to contrast sentiment levels across offline and online data, the research employs a method of converting daily online feelings

into weekly ones. This is achieved by calculating the average of the online standards on days when offline polls are accessible. **Fig 6** depicts the weekly fluctuations in popular sentiment towards the president. The polls conducted by the two organizations provide a nearly comparable graphical representation, however, the online trend diverges from the offline trend. The coefficients obtained from both Gallup and online data are lower, with a value of -0.1 for diverse sentiment and 0.07 for affirmative sentiment. Furthermore, these coefficients do not show any relationship. Polls are surveys that provide direct inquiries to individuals regarding a specific matter.

Table 2. Connection Among Online Data and The Polls with Time Adjustments

	Realmeter Negative	Realmeter Positive	Gallup Negative	Gallup Positive
t - 1	0.298 (0.477)	0.196 (0.655)	0.366 (0.005)	0.304 (0.006)
t - 2	0.307 (0.494)	0.194 (0.679)	0.386 (0.005)	0.310 (0.006)
t - 3	0.273 (0.484)	0.154 (0.664)	0.110 (0.005)	0.050 (0.007)
t + 1	0.288 (0.420)	0.180 (0.579)	0.103 (0.004)	0.040 (0.005)
t + 2	0.236 (0.429)	0.135 (0.593)	0.143 (0.006)	0.127 (0.006)
t + 3	0.162 (0.484)	0.026 (0.643)	0.254 (0.005)	0.212 (0.006)

Conversely, online data lacks any form of intentional regulation; a machine merely gathers accessible facts from the Internet that align with the parameters of a given topic. The inherent disparity between these two factors can result in a diminished degree of connection among opinions expressed online and those expressed offline. The investigation explores two more comparisons to ascertain the disparities between offline and online public sentiments. Online attitudes can accurately reflect the public view in the offline world, albeit with some time lag. In essence, the cognitive processes observed in the physical world can eventually become evident in the virtual realm. The study quantifies the associations between offline and online data, taking into account time modifications. **Table 2** depicts the correlation among the two forms of popular sentiment across different time intervals.

The majority of enhancements are observed in the Realimeter data and online emotions over the two days preceding the time-limit, specifically $t-1$ and $t-2$. For each of the two sentiments (negative and positive) the correlation rises by 0.02 and 0.03 points and 0.003 and 0.001 points. The connection between Gallup and time alterations grows. The highest disparity in correlations is observed two days earlier to the time-limit, denoted as $t-2$. The findings from Realimeter and Gallup demonstrate that internet mood has an impact on the development of offline public opinion, albeit the correlations between the two are not very high. The correlation among offline polls and time-adjusted online attitudes further verifies that the two public sentiments are mutually exclusive, irrespective of temporal factors.

Twitter (X) can serve as a platform for public opinion, which can accurately reflect the perspectives of distinct demographic groups based on age, gender, and political philosophy. This study examines the feasibility by utilizing subgroup data from the polls and establishing correlations between these subgroups and online opinions. **Table 3** displays the association constants pertaining to the various age groups. When examining the connection among online attitudes and Realimeter, it is shown that individuals in their 30s have the strongest correlations, with positive feelings having a correlation coefficient of 0.185 and negative sentiments having a correlation coefficient of 0.233. Furthermore, those aged 60 and above exhibit statistically significant coefficients of 0.21 and 0.176, respectively, when compared to other age groups. The age group of 60 and above exhibits significant correlations with the Gallup poll, with correlation coefficients of 0.282 and 0.369, respectively.

Table 3. Correlation Between Polls and Dataset on Various Age Groups

Age Group	Realmeter Negative	Realmeter Positive	Gallup Negative	Gallup Positive
10-19	0.212 (0.679)	0.238 (0.652)	0.386 (0.005)	0.310 (0.006)
20-29	0.196 (0.589)	0.230 (0.609)	0.354 (0.006)	0.279 (0.007)
30-39	0.132 (0.528)	0.176 (0.588)	0.310 (0.006)	0.212 (0.005)
40-49	0.120 (0.515)	0.154 (0.573)	0.278 (0.005)	0.186 (0.006)
50-59	0.105 (0.489)	0.138 (0.551)	0.258 (0.005)	0.168 (0.006)
60+	0.176 (0.258)	0.210 (0.488)	0.369 (0.256)	0.282 (0.340)

Contrary to the common belief that only the younger generation actively uses social networking sites (SNS), our finding suggests that older individuals also engage fervently in expressing their political opinions through SNS. The concept of gender also yields fascinating findings. The online sentiments exhibit a strong correlation with Gallup for females and Realimeter for males. **Table 4** displays the comprehensive outcomes for various gender categories. Regarding Realimeter, the online sentiment shows the strongest link with the male group, with a correlation coefficient of 0.302 for constructive sentiment and 0.411 for adverse sentiment. These figures are considerably more than the others, except for Gallup's online emotions between females, which are 0.32 and 0.283. The connection coefficients for these gender clusters exhibit significantly higher values compared to the other potential combinations.

Table 4. Correlation Between the Polls and Datasets on Gender

Gender	Realmeter Negative	Realmeter Positive	Gallup Negative	Gallup Positive
Male	0.411 (0.484)	0.302 (0.562)	0.195 (0.499)	0.191 (0.606)
Female	0.156 (0.343)	0.063 (0.479)	0.283 (0.455)	0.320 (0.509)

The data indicates that there is no correlation between internet sentiments and offline polls. In recent times, the utilization of online sentiment analysis on information created by users has been implemented to forecast public sentiment and human actions. Nevertheless, the overall efficacy of it remains questionable. Twitter (X)-derived public sentiment offers a distinct and autonomous indicator of presidential approval, diverging from conventional metrics. Over short durations, online and offline sentiments may exhibit similarities, but their correlation weakens over extended periods. By mitigating fluctuations, the link between online attitudes and election outcomes can be enhanced to some extent [52]. Furthermore, altering online opinions prior to and following the specified polling date does not substantially enhance the correlation. The time difference produces varied outcomes among Realimeter and Gallup. Online executive sanction does not fall within the category of offline ratings. There is just a minor association between gender, age, and political ideology when considering both online and offline public opinion. The analysis consistently shows that there is no significant correlation between presidential approval ratings obtained from Twitter (X) and those obtained from offline polls. The findings suggest that online public attitudes can accurately reflect the broader population, in contrast to traditional offline surveys.

V. CONCLUSION AND FUTURE WORKS

This research provides significant contributions to the understanding of the intricate nature of the public sentiment in South Korea, especially in relation to the interactions between the online and offline opinions of the citizens. By analyzing a large dataset of sentiments from the entire year 2023, we have found subtle patterns in the online sentiments, which occasionally correlate with offline polling data in the approval ratings of presidents. Nevertheless, the overall trend of the correlation between the online sentiment and traditional measures remains moderate, which means that the online discourse can be considered as an independent aspect of people's sentiment. Thus, it is necessary to carry out more research to investigate the factors that lead to the changes in online sentiment. This entails analyzing the effect of web algorithms, the presence of opinion leaders, and the development of echo chambers. It would be more beneficial to identify the peculiarities of sentiment evolution during longer periods and compare the results obtained for different groups of users and platforms. In addition, more qualitative approaches like analyzing the sentiment of the textual content and the classification of sentiments could provide a better perspective on the causes of the observed tendencies in the online sentiment. By adopting a broader framework of sentiment analysis and by building on the methodological developments that are expected to occur in the future, it will be possible to investigate the relations between online and offline processes, which contribute to the formation of public sentiment in the context of the emerging digital society that will subsequently help to make better decisions and design adequate policies.

CRedit Author Statement

The author reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author declares that they have no conflicts of interest.

Funding

This research was supported by the Korea Institute for Advancement of Technology (KIAT) funded by the Ministry of Trade, Industry and Energy (MOTIE), Republic of Korea.

Competing Interests

There are no competing interests.

References

- [1]. W. X. Zhao, S. Li, Y. He, E. Y. Chang, J.-R. Wen, and X. Li, "Connecting Social Media to E-Commerce: Cold-Start Product Recommendation Using Microblogging Information," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 5, pp. 1147–1159, May 2016, doi: 10.1109/tkde.2015.2508816.
- [2]. M. R. Auer, "The Policy Sciences of Social Media," *Policy Studies Journal*, vol. 39, no. 4, pp. 709–736, Nov. 2011, doi: 10.1111/j.1541-0072.2011.00428.x.
- [3]. R. R. Pai and S. Alathur, "Assessing mobile health applications with twitter analytics," *International Journal of Medical Informatics*, vol. 113, pp. 72–84, May 2018, doi: 10.1016/j.ijmedinf.2018.02.016.
- [4]. M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," *Procedia Computer Science*, vol. 87, pp. 44–49, 2016, doi: 10.1016/j.procs.2016.05.124.

- [5]. K. Schouten and F. Frasinicar, "Survey on Aspect-Level Sentiment Analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813–830, Mar. 2016, doi: 10.1109/tkde.2015.2485209.
- [6]. M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM Model for Document-Level Sentiment Analysis," *Machine Learning and Knowledge Extraction*, vol. 1, no. 3, pp. 832–847, Jul. 2019, doi: 10.3390/make1030048.
- [7]. M. Araújo, A. Pereira, and F. Benevenuto, "A comparative study of machine translation for multilingual sentence-level sentiment analysis," *Information Sciences*, vol. 512, pp. 1078–1102, Feb. 2020, doi: 10.1016/j.ins.2019.10.031.
- [8]. Md. S. Islam et al., "Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach," *Artificial Intelligence Review*, vol. 57, no. 3, Mar. 2024, doi: 10.1007/s10462-023-10651-9.
- [9]. J. Zhang, F. Liu, W. Xu, and H. Yu, "Feature Fusion Text Classification Model Combining CNN and BiGRU with Multi-Attention Mechanism," *Future Internet*, vol. 11, no. 11, p. 237, Nov. 2019, doi: 10.3390/fi11110237.
- [10]. R. Moradi, R. Berangi, and B. Minaei, "A survey of regularization strategies for deep models," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 3947–3986, Dec. 2019, doi: 10.1007/s10462-019-09784-7.
- [11]. A. Bello, S.-C. Ng, and M.-F. Leung, "A BERT Framework to Sentiment Analysis of Tweets," *Sensors*, vol. 23, no. 1, p. 506, Jan. 2023, doi: 10.3390/s23010506.
- [12]. L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, Jul. 2018, doi: 10.1007/s10115-018-1236-4.
- [13]. K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14–46, Nov. 2015, doi: 10.1016/j.knosys.2015.06.015.
- [14]. M. Hagenau, M. Liebmann, and D. Neumann, "Automated news reading: Stock price prediction based on financial news using context-capturing features," *Decision Support Systems*, vol. 55, no. 3, pp. 685–697, Jun. 2013, doi: 10.1016/j.dss.2013.02.006.
- [15]. K. Lan, D.-T. Wang, S. Fong, L.-S. Liu, K. K. L. Wong, and N. Dey, "A survey of data mining and deep learning in bioinformatics," *Journal of Medical Systems*, vol. 42, no. 8, Jun. 2018, doi: 10.1007/s10916-018-1003-9.
- [16]. A. Hernández-Blanco, B. Herrera-Flores, D. Tomás, and B. Navarro-Colorado, "A Systematic Review of Deep Learning Approaches to Educational Data Mining," *Complexity*, vol. 2019, no. 1, Jan. 2019, doi: 10.1155/2019/1306039.
- [17]. J. Li, A. Sun, J. Han, and C. Li, "A Survey on Deep Learning for Named Entity Recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 50–70, Jan. 2022, doi: 10.1109/tkde.2020.2981314.
- [18]. Z. Lu, W. Kim, and W. J. Wilbur, "Evaluating Relevance Ranking Strategies for MEDLINE Retrieval," *Journal of the American Medical Informatics Association*, vol. 16, no. 1, pp. 32–36, Jan. 2009, doi: 10.1197/jamia.m2935.
- [19]. S. Maqsood, A. Shahid, F. Nazar, M. Asif, M. Ahmad, and M. Mazzara, "C-POS: A Context-Aware Adaptive Part-of-Speech Language Learning Framework," *IEEE Access*, vol. 8, pp. 30720–30733, 2020, doi: 10.1109/access.2020.2972591.
- [20]. K. Adnan and R. Akbar, "An analytical study of information extraction from unstructured and multidimensional big data," *Journal of Big Data*, vol. 6, no. 1, Oct. 2019, doi: 10.1186/s40537-019-0254-8.
- [21]. S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning," *Archives of Computational Methods in Engineering*, vol. 27, no. 4, pp. 1071–1092, Jun. 2019, doi: 10.1007/s11831-019-09344-w.
- [22]. M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption — A systematic review," *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105287, Oct. 2022, doi: 10.1016/j.engappai.2022.105287.
- [23]. G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [24]. A. Sivaram, L. Das, and V. Venkatasubramanian, "Hidden representations in deep neural networks: Part 1. Classification problems," *Computers & Chemical Engineering*, vol. 134, p. 106669, Mar. 2020, doi: 10.1016/j.compchemeng.2019.106669.
- [25]. P. P. Brahma, D. Wu, and Y. She, "Why Deep Learning Works: A Manifold Disentanglement Perspective," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 10, pp. 1997–2008, Oct. 2016, doi: 10.1109/tnnls.2015.2496947.
- [26]. M. E. Mavroforakis and S. Theodoridis, "A geometric approach to Support Vector Machine (SVM) classification," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 671–682, May 2006, doi: 10.1109/tnn.2006.873281.
- [27]. N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using Bayesian Networks to Analyze Expression Data," *Journal of Computational Biology*, vol. 7, no. 3–4, pp. 601–620, Aug. 2000, doi: 10.1089/106652700750050961.
- [28]. J. R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1023/a:1022643204877.
- [29]. B. Fitzgerald and D. Howcroft, "Towards Dissolution of the is Research Debate: From Polarization to Polarity," *Journal of Information Technology*, vol. 13, no. 4, pp. 313–326, Dec. 1998, doi: 10.1177/026839629801300409.
- [30]. M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, Jun. 2011, doi: 10.1162/coli_a_00049.
- [31]. J. A. Balazs and J. D. Velásquez, "Opinion Mining and Information Fusion: A survey," *Information Fusion*, vol. 27, pp. 95–110, Jan. 2016, doi: 10.1016/j.inffus.2015.06.002.
- [32]. M. Ebrahimi, A. H. Yazdavar, and A. Sheth, "Challenges of Sentiment Analysis for Dynamic Events," *IEEE Intelligent Systems*, vol. 32, no. 5, pp. 70–75, Sep. 2017, doi: 10.1109/mis.2017.3711649.
- [33]. E. Hajičová, "A Natural history of negation," *Lingua*, vol. 82, no. 2–3, pp. 253–261, Nov. 1990, doi: 10.1016/0024-3841(90)90064-r.
- [34]. D. Antonakaki, P. Fragopoulou, and S. Ioannidis, "A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks," *Expert Systems with Applications*, vol. 164, p. 114006, Feb. 2021, doi: 10.1016/j.eswa.2020.114006.
- [35]. C. Gerlitz and B. Rieder, "Mining One Percent of Twitter: Collections, Baselines, Sampling," *M/C Journal*, vol. 16, no. 2, Mar. 2013, doi: 10.5204/mcj.620.
- [36]. D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimedia Tools and Applications*, vol. 82, no. 3, pp. 3713–3744, Jul. 2022, doi: 10.1007/s11042-022-13428-4.
- [37]. T. Winograd, "Understanding natural language," *Cognitive Psychology*, vol. 3, no. 1, pp. 1–191, Jan. 1972, doi: 10.1016/0010-0285(72)90002-3.
- [38]. M. A. K. Halliday, "Towards a language-based theory of learning," *Linguistics and Education*, vol. 5, no. 2, pp. 93–116, Jan. 1993, doi: 10.1016/0898-5898(93)90026-7.
- [39]. Z. Huang and W. Zhao, "Combination of ELMo Representation and CNN Approaches to Enhance Service Discovery," *IEEE Access*, vol. 8, pp. 130782–130796, 2020, doi: 10.1109/access.2020.3009393.
- [40]. M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, Feb. 2022, doi: 10.1007/s10462-022-10144-1.

- [41]. B. N. D. Santos, R. M. Marcacini, and S. O. Rezende, “Multi-Domain Aspect Extraction Using Bidirectional Encoder Representations From Transformers,” *IEEE Access*, vol. 9, pp. 91604–91613, 2021, doi: 10.1109/access.2021.3089099.
- [42]. R. Adams, S. Jeanrenaud, J. Bessant, D. Denyer, and P. Overy, “Sustainability-oriented Innovation: A Systematic Review,” *International Journal of Management Reviews*, vol. 18, no. 2, pp. 180–205, May 2015, doi: 10.1111/ijmr.12068.
- [43]. D. Boyd, S. Golder, and G. Lotan, “Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter,” 2010 43rd Hawaii International Conference on System Sciences, pp. 1–10, Jan. 2010, doi: 10.1109/hicss.2010.412.
- [44]. C. Trapnell et al., “Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks,” *Nature Protocols*, vol. 7, no. 3, pp. 562–578, Mar. 2012, doi: 10.1038/nprot.2012.016.
- [45]. J. Lee and K. E. Brown, “‘Make Korea with America Great Again’: An Articulation and Assemblage of South Korean Extreme Right Practices,” *Communication, Culture and Critique*, vol. 11, no. 1, pp. 53–66, Mar. 2018, doi: 10.1093/ccc/tcx004.
- [46]. Z. Jianqiang, G. Xiaolin, and Z. Xuejun, “Deep Convolution Neural Networks for Twitter Sentiment Analysis,” *IEEE Access*, vol. 6, pp. 23253–23260, 2018, doi: 10.1109/access.2017.2776930.
- [47]. T. Vaiyapuri et al., “Sustainable Artificial Intelligence-Based Twitter Sentiment Analysis on COVID-19 Pandemic,” *Sustainability*, vol. 15, no. 8, p. 6404, Apr. 2023, doi: 10.3390/su15086404.
- [48]. Y. M. González, “Reforming to Avoid Reform: Strategic Policy Substitution and the Reform Gap in Policing,” *Perspectives on Politics*, vol. 21, no. 1, pp. 59–77, Mar. 2022, doi: 10.1017/s1537592722000135.
- [49]. D. J. Hemel and E. A. Posner, “Presidential Obstruction of Justice,” *SSRN Electronic Journal*, 2017, doi: 10.2139/ssrn.3004876.
- [50]. S. R. Cotten and S. S. Gupta, “Characteristics of online and offline health information seekers and factors that discriminate between them,” *Social Science & Medicine*, vol. 59, no. 9, pp. 1795–1806, Nov. 2004, doi: 10.1016/j.socscimed.2004.02.020.
- [51]. S. Valenzuela, “Unpacking the Use of Social Media for Protest Behavior,” *American Behavioral Scientist*, vol. 57, no. 7, pp. 920–942, Mar. 2013, doi: 10.1177/0002764213479375.
- [52]. I. Dylko, I. Dolgov, W. Hoffman, N. Eckhart, M. Molina, and O. Aaziz, “The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure,” *Computers in Human Behavior*, vol. 73, pp. 181–190, Aug. 2017, doi: 10.1016/j.chb.2017.03.031.

Publisher’s note: The publisher remains neutral regarding jurisdictional claims in published maps and institutional affiliations. The content is solely the responsibility of the authors and does not necessarily reflect the views of the publisher.