Scientific Journal of Informatics
Vol. 13, No. 1, February 2026

p-ISSN 2407-7658 https://journal.unnes.ac.id/journals/sji/index e-ISSN 2460-0040

Combination Of VADER Sentiment Analysis and SEQ Scale For Evaluating the Usability of The Gojek Application

Zakiyah Apriliya Budiarti[1], Tenia Wahyuningrum[2*], Adnan Purwanto[3], Muhammad Akbar Setiawan[4], Singgih Setia Andiko[5], Abdul Karim[6]
[1, 3, 4, 5]Department of of Informatics Engineering, STMIK Widya Utama Purwokerto, Indonesia
[2]Study Program of Informatics, Telkom University, Banyumas, Indonesia
[6]Department of Artificial Intelligence Convergence, Hallym University, Republic of Korea

Abstract.
Purpose: This study aims to analyze user sentiment toward the Gojek mobile application using the Valence Aware Dictionary for Sentiment Reasoning (VADER) method and to evaluate the perceived ease of use of the application using the Single Ease Question (SEQ) instrument.
Methods: The research data were obtained by scraping 30,000 user reviews of the Gojek application from the Google Play Store. The reviews were processed through text preprocessing, sentiment classification using the VADER method, and subsequent mapping of sentiment polarity scores to a 1–7 SEQ usability scale. A Spearman rank correlation analysis was conducted to examine the relationship between sentiment scores and derived SEQ values.
Result: The results indicate that user sentiment toward the Gojek application is predominantly positive, followed by neutral and negative sentiments. The overall average SEQ score is 4.11, suggesting that the application is generally perceived as fairly easy to use. Furthermore, a strong and statistically significant positive association was found between VADER sentiment scores and SEQ usability scores, indicating that more positive sentiment tends to be associated with higher perceived ease of use.
Novelty: This study contributes to the literature by empirically integrating sentiment analysis and usability evaluation using VADER and SEQ within the context of an Indonesian super-app. The findings provide practical insights for digital application developers to identify usability strengths and areas for improvement based on large-scale user feedback.

Keywords: Sentiment analysis, Ride-hailing, Single ease question, usability, Valence aware dictionary for sentiment reasoning
Received November 2025 / **Revised** January 2026 / **Accepted** February 2026

INTRODUCTION

Technological developments in Indonesia have facilitated various community activities, including transportation, delivery services, and digital payments. One service that continues to grow rapidly in Indonesia is Gojek, a decacorn that offers a wide range of app-based services and has significantly transformed urban mobility and digital transactions. Understanding user sentiment toward the Gojek application through Google Play Store reviews is essential for evaluating service quality and identifying areas for improvement [1]. User-generated reviews provide direct insights into user experiences and expectations, making them a valuable data source for application evaluation [2].

However, the large volume of review data and the diversity of language styles used by users often make manual analysis inefficient and impractical [3], [4]. As a result, automated sentiment analysis techniques have become increasingly important for extracting meaningful patterns from large-scale textual data [5].

Among various approaches, VADER (Valence Aware Dictionary for Sentiment Reasoning) is a lexicon-based sentiment analysis method specifically designed to process informal and expressive text commonly found in social media and app store reviews [6].

*Corresponding author.
Email addresses: zakiyah@swu.ac.id (Budiarti), teniaw@telkomuniversity.ac.id (Wahyuningrum)*, adnan@swu.ac.id (Purwanto), akbar@swu.ac.id (Setiawan), singgihs.andiko@students.amikom.ac.id (Andiko), khiengchanna@npic.edu.kh (Karim)
DOI: 10.15294/sji.v13i1.35956

Previous studies have demonstrated that VADER performs effectively in analyzing short, non-standard, and emotionally expressive user-generated content, including reviews and comments containing slang, emoticons, and informal language structures [7], [8], [9]. Research applying VADER to digital platforms such as mobile applications, social media, and educational systems has shown that the method can efficiently identify sentiment polarity with reasonable accuracy while maintaining computational simplicity [10], [11], [12]. These characteristics make VADER a suitable choice for sentiment analysis in the context of Google Play Store reviews, where textual data are often brief and unstructured.

In addition to sentiment analysis, evaluating application usability is also critical for understanding user experience and system effectiveness. The Single Ease Question (SEQ) is a widely used single-item usability metric administered after task completion to assess perceived ease of use on a 1–7 scale ranging from "very difficult" to "very easy" [10]. SEQ has been applied extensively in usability studies due to its simplicity, reliability, and ability to capture users' immediate perceptions of task difficulty [13].

Several prior studies have employed usability metrics such as SEQ and SUS to assess user experience in digital applications, demonstrating that mid-scale SEQ values often indicate acceptable usability while highlighting opportunities for further improvement [14], [15], [16]. Other studies have also explored the potential relationship between user sentiment expressed in reviews and perceived usability, suggesting that sentiment polarity may serve as an indirect indicator of user satisfaction and system performance [17].

Despite the growing body of research on sentiment analysis and usability evaluation, studies that integrate sentiment analysis using the VADER method with usability assessment using the SEQ instrument-particularly in the context of Indonesian super-apps such as Gojek-remain limited [18]. Therefore, this study aims to analyze public sentiment toward the Gojek application using VADER-based sentiment analysis and to evaluate perceived usability through derived SEQ scores. By combining sentiment polarity analysis with usability evaluation and statistical validation, this study seeks to provide a more comprehensive understanding of user experience and to support data-driven improvements in digital service quality.

METHODS

Several stages were carried out in this study and arranged systematically to produce a valid and accountable analysis. These stages included collecting user reviews of the Gojek application from the Google Play Store and then performing text preprocessing. The data were then analyzed using the Valence Aware Dictionary for Sentiment Reasoning (VADER) method, and the sentiment analysis results served as the basis for a usability evaluation using the Single Ease Question (SEQ) instrument. All of these stages are illustrated in the research flow diagram in Figure 1, which helps readers easily understand the steps taken and the interrelationships among the processes in this study.

Data collection

The research data were collected from user reviews of the Gojek mobile application available on the Google Play Store, which provides publicly accessible feedback reflecting direct user experiences and has been widely used as a data source for user experience and sentiment analysis of mobile applications [17], [19]. Data collection was conducted in September 2025 using a Python-based web scraping approach with the Google Play Scraper library, which has been commonly applied in prior studies to automate the large-scale collection of application reviews from the Google Play Store [1], [19].

A total of 30,000 reviews were retrieved using a chronological sampling strategy, starting from the most recent reviews available at the time of collection. Consequently, the dataset represents recent user experiences prior to and up to September 2025, capturing usability perceptions corresponding to the application's condition during that period. Although the exact posting dates of individual reviews were not available in the extracted dataset, focusing on the most recent reviews helps ensure temporal relevance, as commonly adopted in app review-based studies[17].

To ensure data quality, reviews were included if they contained textual feedback in Indonesian or English and were associated with a star rating, following common preprocessing practices in mobile application review analysis [1]. Reviews without text, duplicate entries, or non-informative content were excluded. The finalized dataset was stored in CSV format to support subsequent preprocessing, sentiment analysis, and usability evaluation using the SEQ framework.

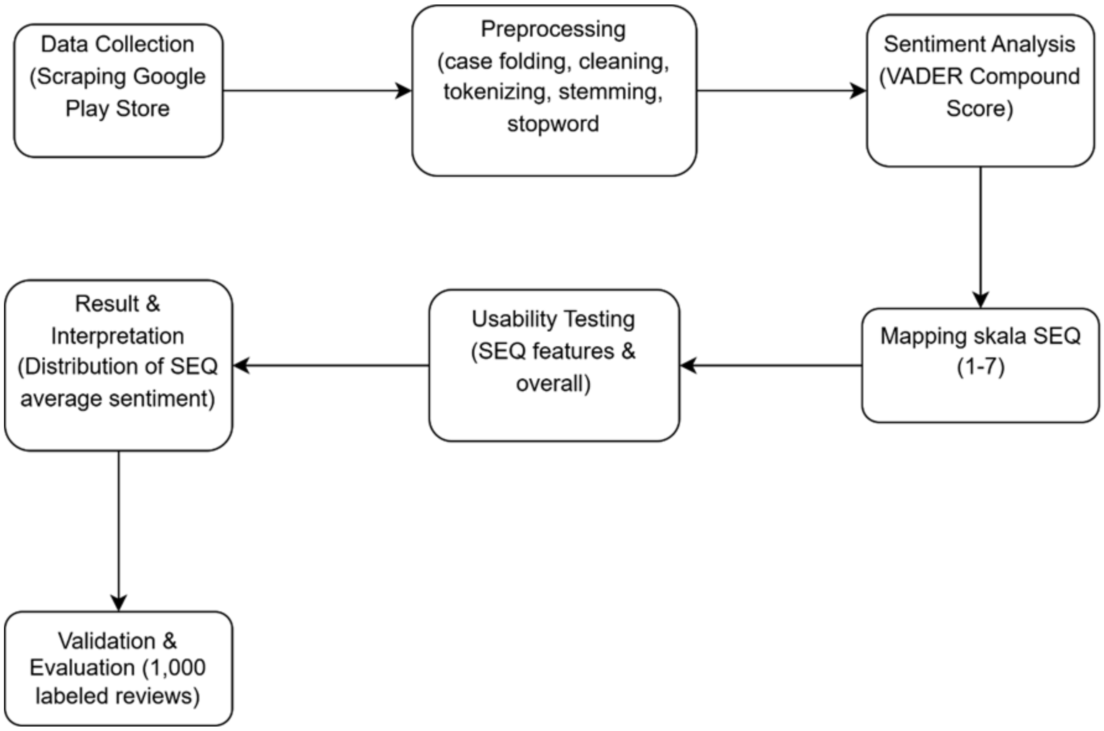

Figure 1. Research flow

Ethical considerations

This study analyzes publicly available user reviews obtained from the Google Play Store for academic research purposes. The data were collected in compliance with the platform's terms of service and used only in an aggregated manner. No personal or identifiable user information was collected or processed, and all data were handled to ensure user anonymity and privacy. As the study relies solely on publicly accessible content and does not involve direct interaction with human participants, formal ethical approval was not required.

Preprocessing

Prior to sentiment analysis, several text preprocessing steps were applied to improve data quality and to ensure that the textual input was suitable for sentiment classification. Text preprocessing is a standard procedure in natural language processing (NLP) tasks, particularly when dealing with user-generated content such as app reviews, which often contain noise and informal language [20], [21] .

The preprocessing steps applied in this study are as follows:

1. Case folding, which converts all characters in the text to lowercase. This step ensures consistency by treating words with identical meanings but different letter cases (e.g., "Driver" and "driver") as the same token.
2. Text cleaning, which removes irrelevant elements such as URLs, numbers, punctuation marks, and special characters. These elements frequently appear in online reviews but generally do not contribute to sentiment interpretation and may introduce noise into the analysis.
3. Tokenization, which splits sentences into individual words or tokens. This process allows each word to be analyzed independently and is a fundamental step in many text mining and sentiment analysis pipelines.
4. Stemming, which reduces words to their root or base form using the Sastrawi library. For example, the words *"berjalan"*, *"berjalanlah"*, and *"dijalankan"* are all normalized to the root word *"jalan"*. Stemming is particularly important in Indonesian language processing due to its rich morphological structure.

5. Stopword removal, which eliminates commonly used words such as *"yang"*, *"dan"*, and *"di"*. These words frequently appear in text but carry little semantic value for sentiment classification, serving mainly as syntactic connectors.

The output of this preprocessing stage is a simplified and normalized text representation that facilitates more accurate sentiment analysis using the VADER method. By reducing noise and linguistic variation, preprocessing helps improve the reliability of sentiment polarity detection in large-scale review data.

Sentiment analysis

The Valence Aware Dictionary for Sentiment Reasoning (VADER) method is one of the sentiment analysis methods chosen for this study because it is suitable for the characteristics of application review data, which is short, concise, and often uses informal language. VADER is a lexicon- and rule-based approach designed to analyze informal text, such as social media comments and application reviews [22]. The advantage of VADER lies in its combination of a rich sentiment dictionary with heuristics that take into account capitalization, punctuation (e.g., exclamation marks), intensifiers (strengthening words), and conjunctions that can change the polarity of sentiment, enabling it to capture contextual nuances in the text while remaining easy to use and having a good level of accuracy [23], [24]. In addition, VADER uses a compound score on a SEQ scale (1–7) [25]. The results of this sentiment analysis are then used in usability evaluation using the SEQ instrument.

Usability testing

The usability aspect of research is no less important than sentiment analysis, as it directly influences user experience and satisfaction, thereby determining the success of a digital application. One simple but effective instrument for measuring ease of use is the Single Ease Question (SEQ). In addition to being effective, SEQ was chosen because it is concise, practical, and is sensitive in distinguishing levels of ease across tasks [26]. The Single Ease Question (SEQ) is widely adopted in usability research because it consists of a single core question that effectively captures users' perceptions of a system's ease of use [27]. According to practical guidelines, SEQ scores range from 1 to 7, with higher scores indicating greater ease of use [28].

The results of sentiment analysis using VADER were then used to map user perceptions of various Gojek features, such as Goride, Gocar, Gofood, Gosend, Gopay, and the app in general. Feature categorization was performed using keyword-based matching on review text, where reviews were assigned to specific application features based on the resence of predefined service-related keyword. The SEQ score was then calculated as the primary indicator of the app's ease of use.

Integration of sentiment analysis and usability evaluation

This study proposes a rule-based integration framework to systematically relate sentiment polarity extracted from user reviews to perceived usability levels measured using the Single Ease Question (SEQ) scale. At this stage, the integration is presented as a conceptual and procedural mechanism, without implying direct equivalence between the sentiment score and usability measurement.

The compound score generated by the VADER sentiment analysis was selected as the primary indicator of overall sentiment polarity, as it effectively summarizes positive, negative, and neutral components into a single continuous measure suitable for user-generated text analysis. Sentiment analysis on mobile application reviews has been widely applied in recent research to derive insights from user feedback on apps, highlighting its usefulness in understanding user perceptions and expectations of mobile services [29]. To enable comparison with the ordinal SEQ scale, compound scores were mapped to discrete usability levels using predefined threshold intervals. The Single Ease Question (SEQ) scale is a widely used metric in usability research for capturing participants' perceived ease of use on a seven-point scale. This approach aligns with common practices in usability evaluation, where ordinal usability metrics are related to users' subjective perceptions [30].

The mapping scheme converts VADER compound scores into SEQ usability categories through a rule-based threshold function. Threshold values were determined based on the standard interpretation of VADER sentiment polarity and the empirical distribution of sentiment scores observed in the dataset. This mapping mechanism serves as an operational step to prepare for subsequent statistical validation rather than as direct evidence of usability measurement.

Table 1. Mapping of compound vader scores to the seq scale

Compound score range	Label sentiment	SEQ score	Usability interpretation
≤ -0.4	Very negative	1	Veri difficult
-0.4 < × ≤ -0.1	Negative	2	Difficult
-0.1 < × ≤ -0.05	Quite negative	3	Quite difficult
-0.05 < × ≤ 0.00	Neutral	4	Quite
0.00< × ≤ 0.3	Quite positive	5	Quite easy
0.3 < × ≤ 0.6	Positive	6	Easy
0.6	Very positive	7	Very easy

Table 1 presents the mapping rules associating ranges of compound sentiment scores with corresponding SEQ usability levels. This mapping framework enables further inferential analysis, such as correlation testing, to assess the relationship between sentiment polarity and perceived usability in later sections of the manuscript.

RESULTS AND DISCUSSIONS

Results

This section presents the outcomes of sentiment analysis and usability evaluation based on Gojek user review data. The results are reported using tables and graphical visualizations, including the distribution of sentiment polarity, SEQ usability scores, and feature-based statistical summaries. To ensure clarity, quantitative findings are first presented in dedicated result-oriented subsections, followed by a discussion that interprets the results in relation to usability perception, statistical validation, and findings from previous studies.

Sentiment analysis result

The VADER sentiment analysis model was applied to 30,000 Gojek user reviews to obtain sentiment polarity scores. VADER generates a compound score ranging from −1 (strongly negative) to +1 (strongly positive), representing the overall sentiment polarity expressed in each review.

Based on the compound scores, the reviews were classified into positive, neutral, and negative sentiment categories. The results show that positive sentiment constitutes the largest proportion of the dataset, followed by neutral sentiment, while negative sentiment represents the smallest proportion.

Analysis of the compound score distribution indicates that most values are concentrated around neutral to moderately positive ranges, with fewer instances of extreme positive or negative scores. This distribution reflects variability in sentiment intensity across user reviews.

These findings provide a descriptive overview of sentiment polarity patterns observed in large-scale textual user reviews. At this stage, the results strictly represent sentiment analysis outputs and do not imply usability evaluation or task-level performance.

SEQ score distribution

This subsection presents the descriptive results of the Single Ease Question (SEQ) analysis derived from Gojek user reviews. The analysis aims to summarize the overall distribution and characteristics of SEQ scores before further validation and inferential analysis. The results are reported using descriptive statistics, frequency distributions, and proportional visualizations.

Table 2. Overall SEQ statistics

Statistic	Score
Mean	4.11
Median	4.0
Standar deviation	0.83
Number of reviews	30,000

Table 2 summarizes the descriptive statistics of the SEQ scores obtained from 30,000 user reviews. The average SEQ score is 4.11, with a median of 4.0 and a standard deviation of 0.83, indicating moderate variability in the SEQ values across the dataset

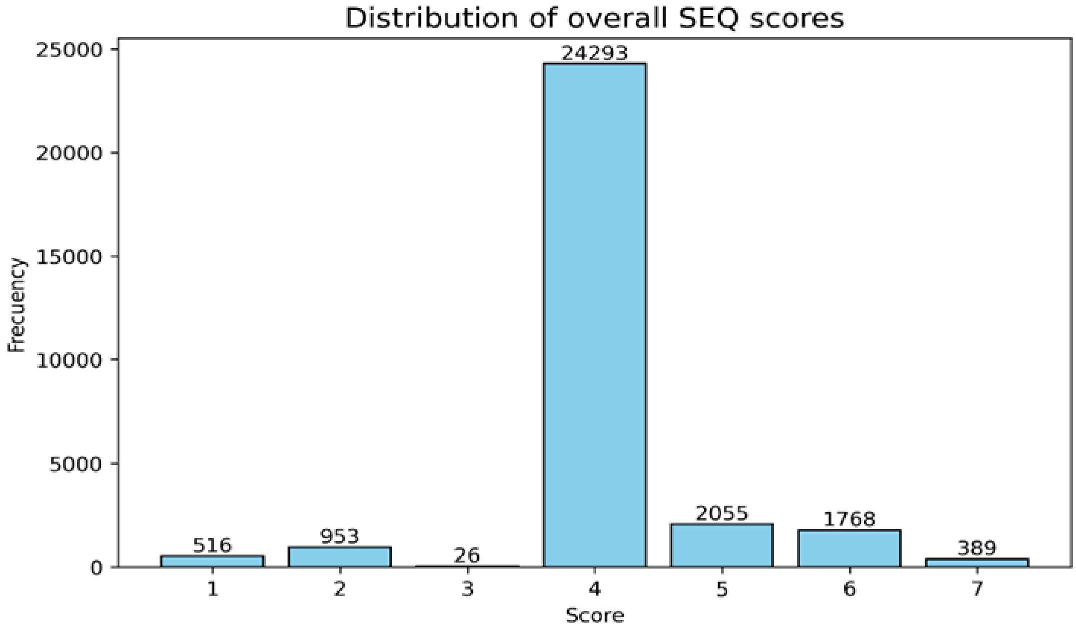

Figure 2. Distribution of overall SEQ scores

Figure 2 shows the distribution of overall SEQ scores from 30,000 Gojek user reviews. The results are strongly concentrated at score 4, indicating that most users perceived the application as having moderate ease of use. Higher scores (5–6) appear less frequently, while very low (1–3) and very high (7) scores are relatively rare. Overall, the distribution suggests that user experiences tend to cluster around the midpoint of the SEQ scale rather than at extreme evaluations.

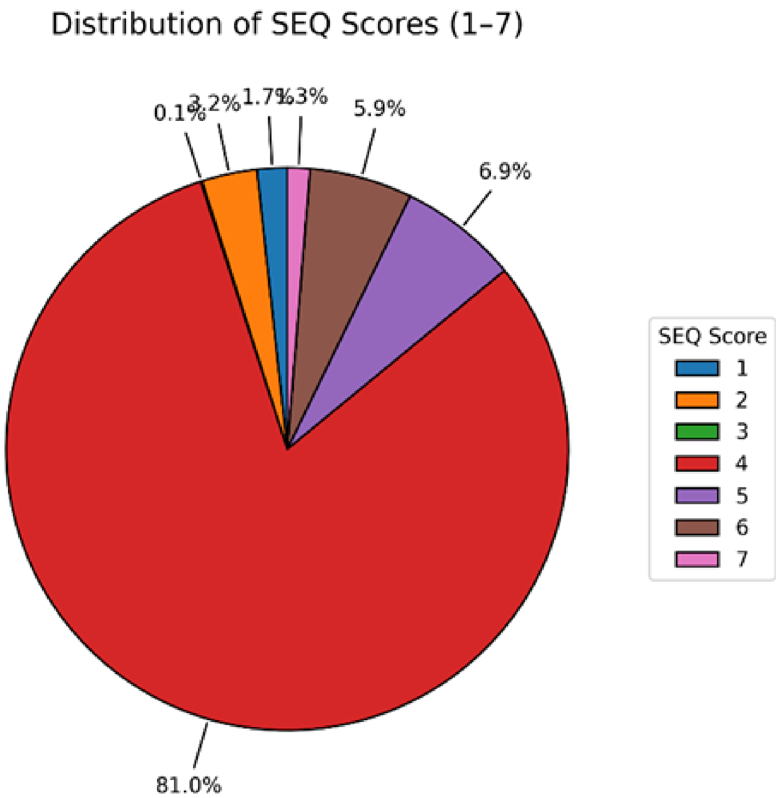

Figure 3. Overall SEQ score proportions

Figure 3 presents the proportional distribution of SEQ scores across all reviews. SEQ score 4 dominates the dataset, accounting for approximately 81,0% of the total observations. Higher scores (5 and 6) and lower scores (1–3) each constitute smaller proportions of the dataset

Validation of the relationship between vader sentiment scores and seq scores (spearman correlation)
To examine whether sentiment polarity is associated with the derived SEQ scores, a Spearman rank correlation analysis was conducted between the VADER compound sentiment scores and the SEQ scores. Spearman correlation was selected because it is a non-parametric method and does not require the assumption of normally distributed data.

The analysis resulted in a Spearman correlation coefficient of $\rho = 0.716$, with a statistically significant p-value ($p < 0.05$), indicating a strong positive association between sentiment polarity and SEQ scores. This suggests that reviews with higher sentiment scores tend to be associated with higher SEQ values.

It is important to emphasize that this result reflects an association rather than a causal relationship. The correlation analysis provides empirical support for the relationship between sentiment polarity and SEQ scores, but it does not imply that sentiment directly determines usability perceptions.

Statistical results for each application feature
This subsection presents the descriptive statistical results of SEQ scores across different Gojek application features. The analysis aims to examine how the average SEQ scores vary among features based on user reviews, without drawing evaluative or comparative conclusions.

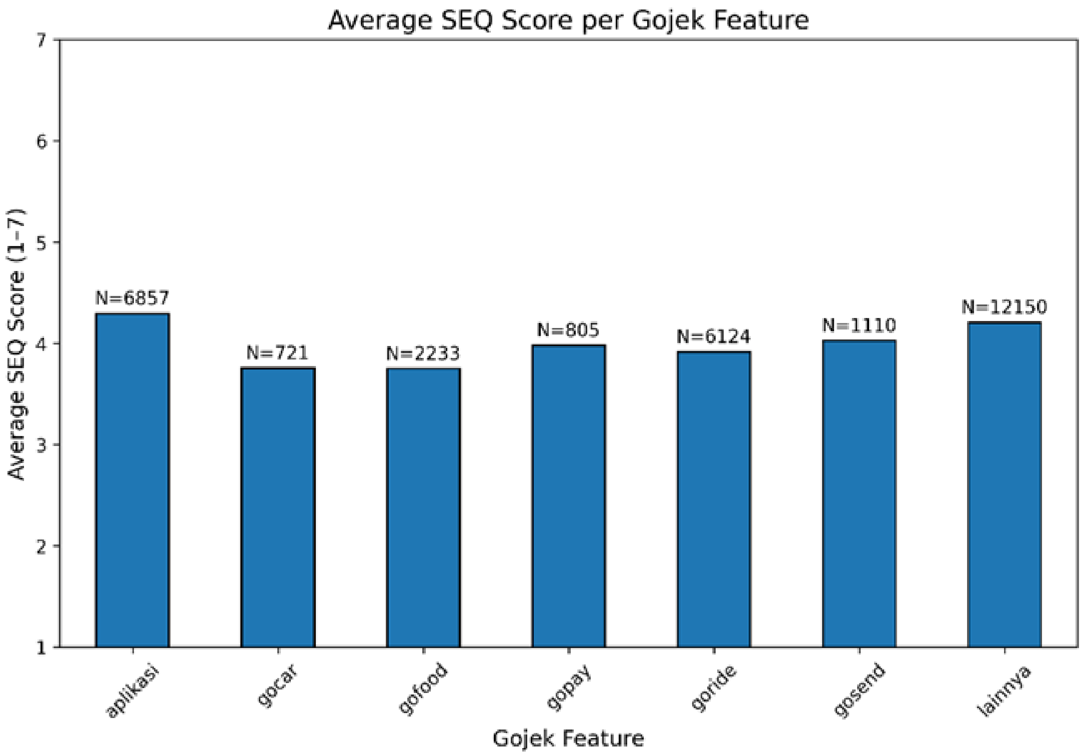

Figure 4. Average SEQ score for each Gojek feature

Figure 4 shows the average SEQ scores for each Gojek feature based on user reviews. The results indicate that the average SEQ scores across features range from approximately 3.7 to 4.3 on a 1–7 scale.

The "aplikasi" category shows an average SEQ score of approximately 4.3 based on 6,857 reviews, while "gocar" and "gofood" have lower average SEQ scores, each around 3.7, based on 721 and 2,233 reviews, respectively. The "gopay" feature has an average SEQ score of approximately 4.0 from 805 reviews, and "goride" shows a similar average score based on 6,124 reviews.

The “gosend” feature records an average SEQ score slightly above 4.0 based on 1,110 reviews, while the “lainnya” category shows an average score of approximately 4.2, representing 12,150 reviews. The number of reviews for each feature is indicated above the corresponding bars.

Discussion

Based on the sentiment classification results using the VADER method, a large proportion of Gojek user reviews exhibit positive sentiment, alongside neutral and negative opinions that reflect variability in user perspectives. These findings support the suitability of the VADER algorithm for large-scale sentiment analysis of informal and unstructured user-generated text, such as app store reviews. Previous studies have reported comparable performance of VADER when applied to short, expressive, and non-standard textual feedback, emphasizing its efficiency and reliability in similar contexts [31], [32], [33]. Nevertheless, as noted in prior research, VADER remains limited in detecting nuanced linguistic expressions such as sarcasm, irony, or implicit dissatisfaction, which should be taken into account when interpreting sentiment-based results [34], [35].

The usability analysis based on the derived Single Ease Question (SEQ) scores indicates that the overall usability of the Gojek application is centered around the midpoint of the 1–7 SEQ scale. The average SEQ score of 4.11 suggests that user interactions with the application are generally perceived as moderately easy, but do not yet reflect the highest level of perceived ease. The dominance of a score of 4 in the overall SEQ distribution implies that, while users are able to use the application without major difficulty, certain aspects of the user experience may still require improvement to better meet user expectations. Similar interpretations of mid-scale SEQ values have been reported in prior usability studies, where such scores often indicate acceptable but improvable system performance [36].

To further examine the relationship between sentiment polarity and perceived usability, a Spearman rank correlation analysis was conducted between VADER compound sentiment scores and SEQ usability scores. The results reveal a strong and statistically significant positive association ($\rho = 0.716$, $p < 0.05$), indicating that reviews expressing more positive sentiment tend to be associated with higher SEQ scores. This finding provides empirical support for the existence of a meaningful relationship between sentiment expressed in user reviews and derived usability evaluations [35]. However, it is important to emphasize that this relationship represents an association rather than a causal effect, and sentiment polarity alone should not be interpreted as a direct measure of usability.

Further analysis across application features reveals variations in average SEQ scores. Features such as GoCar and GoFood exhibit comparatively lower average SEQ values, suggesting that user interactions related to these services may involve greater challenges, potentially associated with navigation complexity, system responsiveness, or service-related factors [37]. In contrast, higher average SEQ scores observed for the general application interface and GoPay feature indicate a closer alignment with user expectations in terms of interface design and feature integration [38]. These feature-level differences suggest that future usability improvements may benefit from prioritizing services with lower SEQ scores, while maintaining and refining design elements that already perform relatively well.

Limitation

This study has several limitations that should be considered when interpreting the results. First, the usability evaluation relies on SEQ scores derived indirectly from sentiment analysis of user reviews, rather than from controlled usability evaluation scenarios. Although a strong and statistically significant association between VADER sentiment scores and SEQ values was observed, the derived usability scores may not fully capture task-specific usability issues experienced by users during actual interactions with the application [36], [39].

Second, user reviews collected from the Google Play Store may contain content that is not strictly related to interface usability, such as complaints about service providers, delivery delays, pricing, or external factors beyond the application’s interface. While large-scale data helps reduce individual bias, the presence of such non-usability-related content may still influence the derived SEQ scores.

Third, the sentiment analysis process relies on the VADER lexicon-based model, which is known to have limitations in handling sarcasm, mixed sentiment, and implicit expressions of dissatisfaction. As a result, certain reviews may be misclassified, particularly when users express usability concerns indirectly or ambiguously [34].

In addition, this study does not provide an openly accessible code repository for replication. However, the preprocessing and sentiment analysis pipeline can be reproduced using standard Python libraries commonly employed in sentiment analysis research, and the implementation scripts can be shared upon reasonable request to support transparency and reproducibility.

Finally, this study focuses on a single mobile application within a specific platform and time frame. Therefore, the findings may not be directly generalizable to other applications, platforms, or cultural contexts [40]. Future research may benefit from incorporating multilingual sentiment models, task-based usability evaluation, or human-labeled datasets to further validate and extend the proposed approach.

CONCLUSION

This study proposed an integrated approach that combines VADER sentiment analysis with the Single Ease Question (SEQ) framework to evaluate the usability of the Gojek mobile application based on large-scale user reviews. By analyzing 30,000 reviews collected from the Google Play Store, this research provides an empirical overview of user sentiment patterns and derived usability perceptions across different application features.

The sentiment analysis results indicate that positive sentiment dominates user feedback, suggesting generally favorable user experiences when interacting with the application. The usability evaluation based on derived SEQ scores shows that the overall usability of the Gojek application is perceived as moderately good, with an average SEQ score of 4.11 on a 1–7 scale. This finding implies that while the application is relatively easy to use, there is still room for improvement to achieve higher levels of perceived ease of use.

The validation using Spearman rank correlation demonstrates a strong and statistically significant positive association between VADER sentiment scores and SEQ usability scores ($\rho = 0.716$, $p < 0.05$). This result provides evidence that sentiment polarity expressed in user reviews is meaningfully associated with perceived usability, supporting the use of sentiment-based approaches as a complementary method for large-scale usability evaluation. However, this relationship reflects an association rather than causation, and sentiment analysis should not be interpreted as a direct substitute for traditional usability evaluation.

Feature-level analysis further reveals variations in usability perceptions across different services within the application. Lower average SEQ scores observed in features such as GoCar and GoFood indicate potential areas for usability enhancement, while higher scores in the general application interface and GoPay highlight strengths in interface design and feature integration. These findings suggest that targeted improvements in specific features may help improve overall user experience.

Overall, this study contributes to the growing body of research on automated usability evaluation by demonstrating how sentiment analysis and usability metrics can be integrated to extract insights from large-scale user feedback. Future work may extend this approach by incorporating human-labeled datasets, multilingual sentiment models, or task-based usability evaluations to further strengthen the robustness and generalizability of the findings.

REFERENCES

[1] Y. I. Kurniawan *et al.*, "Naive Bayes Classifier with SMOTE for Sentiment Analysis of Blibli App Reviews on The Google Play Store," *Jurnal Penelitian Inovatif*, vol. 5, no. 3, pp. 2675–2688, Sep. 2025, doi: 10.54082/jupin.1842.

[2] Tarwoto, R. Nugroho, N. Azka, and W. Sayudha Rendra Graha, "Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google PlayStore Menggunakan IndoBERT," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 3, no. 2, pp. 495–505, Apr. 2025, doi: 10.35870/jtik.v9i2.3340.

[3] M. H. Hoti and J. Ajdari, "Sentiment Analysis Using the Vader Model for Assessing Company Services Based on Posts on Social Media," *SEEU Review*, vol. 18, no. 2, pp. 19–33, Dec. 2023, doi: 10.2478/seeur-2023-0043.

[4] H. Basri, M. B. S. Junianto, and I. Kusyadi, "Enhancing Usability Testing Through Sentiment Analysis: A Comparative Study Using SVM, Naive Bayes, Decision Trees and Random Forest," *Jurnal Teknologi Sistem Informasi dan Aplikasi*, vol. 7, no. 4, pp. 1603–1610, Oct. 2024, doi: 10.32493/jtsi.v7i4.45117.

[5] A. Iskoko, I. Tahyudin, and P. Purwadi, "Analisis Sentimen Ulasan Pengguna Aplikasi Sinaga Mobile pada Google Play Store Menggunakan Algoritma Naive Bayes," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 6, pp. 1635–1645, Jun. 2025, doi: 10.52436/1.jpti.843.

[6] L. M. Gandy, L. V. Ivanitskaya, L. L. Bacon, and R. Bizri-Baryak, "Public Health Discussions on Social Media: Evaluating Automated Sentiment Analysis Methods," *JMIR Form. Res.*, vol. 9, 2025, doi: 10.2196/57395.

[7] M. Tsiourlini, K. Tzafilkou, D. Karapiperis, and C. Tjortjis, "Text Analytics on YouTube Comments for Food Products," *Information (Switzerland)*, vol. 15, no. 10, Oct. 2024, doi: 10.3390/info15100599.

[8] R. Kurniawan, H. O. L. Wijaya, and R. P. Aprisusanti, "Sentiment Analysis of Google Play Store User Reviews on Digital Population Identity App Using K-Nearest Neighbors," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 2, pp. 170–178, Jun. 2024, doi: 10.32736/sisfokom.v13i2.2071.

[9] H. Rahmajati, M. Chandra Saputra, and A. D. Herlambang, "Penerapan Aspect-Based Sentiment Analysis untuk Identifikasi Masalah Kualitas Layanan Aplikasi Mobile iPusnas berdasarkan Persepsi Pengguna," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 7, pp. 2548–964, 2025, [Online]. Available: http://j-ptiik.ub.ac.id

[10] A. Subiyakto, R. Aisy, B. G. Sudarsono, M. Sihotang, D. Setiyadi, and A. Sani, "Empirical evaluation of user experience using lean product and process development: A public institution case study in Indonesia," in *AIP Conference Proceedings*, American Institute of Physics Inc., Apr. 2021. doi: 10.1063/5.0041676.

[11] R. Srivastava, P. K. Bharti, and P. Verma, "Comparative Analysis of Lexicon and Machine Learning Approach for Sentiment Analysis," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, pp. 71–77, 2022, [Online]. Available: www.ijacsa.thesai.org

[12] R. A. Putra, R. Novita, T. K. Ahsyar, and Zarnelly, "Implementation of Classification Algorithm for Sentiment Analysis: Measuring App User Satisfaction," *Teknika*, vol. 13, no. 2, pp. 204–212, Jun. 2024, doi: 10.34148/teknika.v13i2.827.

[13] G. Tamami, W. A. Triyanto, and S. Muzid, "Sentiment Analysis Mobile JKN Reviews Using SMOTE Based LSTM," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 19, no. 1, p. 13, Jan. 2025, doi: 10.22146/ijccs.101910.

[14] X. Liu and K. Park, "Construction and Ranking of Usability Indicators for Medical Websites Based on Website User Experience," *Applied Sciences (Switzerland)*, vol. 14, no. 13, Jul. 2024, doi: 10.3390/app14135465.

[15] S. Pambudi, P. Setiaji, and W. A. Triyanto, "Sentiment Analysis of Fizzo Novel Application Using Support Vector Machine and Naïve Bayes Algorithm with SEMMA Framework," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 4, pp. 1861–1880, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.4875.

[16] T. Sari Ningsih, T. Iman Hermanto, I. Ma, ruf Nugroho, and S. Tinggi Teknologi Wastukancana Purwakarta, "Sentiment Analysis of Mobile Provider Application Reviews Using Naive Bayes Algorithm and Support Vector Machine," *Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 2, 2024, doi: 10.33395/v8i2.13469.

[17] C. Rahmi Anissa, K. Ditha Tania, and W. Kurnia Sari, "Sentiment Analysis on Google Play Store Reviews to Measure User Perception of the Gojek Application Using CNN," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 6, p. 3322, 2025, doi: 10.30871/jaic.v9i6.11084.

[18] P. Amri, D. M. Suri, and Syuhada, "The analysis of ride hailing user characteristics from app reviews," *Jurnal Siasat Bisnis*, pp. 241–262, Nov. 2024, doi: 10.20885/jsb.vol28.iss2.art7.

[19] H. Aldabbas, A. Bajahzar, M. Alruily, A. A. Qureshi, R. M. Amir Latif, and M. Farhan, "Google Play Content Scraping and Knowledge Engineering using Natural Language Processing Techniques with the Analysis of User Reviews," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 192–208, Jan. 2020, doi: 10.1515/jisys-2019-0197.

[20] D. Setiawan, D. Utari Iswavigra, and E. Anggiratih, "Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates," *Scientific Journal of Informatics*, vol. 12, no. 3, 2025, doi: 10.15294/sji.v12i3.26360.

[21] M. Rahayu Ningsih, "Sentiment Analysis on SocialMedia Using TF-IDF Vectorization and H2O Gradient Boosting for Student Anxiety Detection," *Scientific Journal of Informatics*, vol. 11, no. 4, 2024, doi: 10.15294/sji.v11i4.20582.

[22] K. Barik and S. Misra, "Analysis of customer reviews with an improved VADER lexicon classifier," *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/s40537-023-00861-x.

[23] Y. A. Reddy, S. Agarwal, V. Parashar, and A. Arora, "Visualizing Public Opinion on X: A Real-Time Sentiment Dashboard Using VADER and DistilBERT," *ArXiv*, Jun. 2025, [Online]. Available: http://arxiv.org/abs/2504.15448

[24] A. Oad *et al.*, "VADER Sentiment Analysis without and with English Punctuation Marks," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, no. 2, pp. 1483–1488, Apr. 2021, doi: 10.30534/ijatcse/2021/1371022021.

[25] V. Vought, R. Vought, A. S. Lee, I. Zhou, M. Garneni, and S. A. Greenstein, "Application of sentiment and word frequency analysis of physician review sites to evaluate refractive surgery care," *Advances in Ophthalmology Practice and Research*, vol. 4, no. 2, pp. 78–83, May 2024, doi: 10.1016/j.aopr.2024.03.002.

[26] M. W. Wronikowska *et al.*, "Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems," *J. Eval. Clin. Pract.*, vol. 27, no. 6, pp. 1403–1416, Dec. 2021, doi: 10.1111/jep.13582.

[27] A. L. Ridho, D. Dwi, J. Suwawi, and R. R. Riskiana, "Redesigning the User Interface of a University Laboratory Website Using the User-Centered Design Approach," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 378–387, Aug. 2023, doi: 10.30865/klik.v4i1.1172.

[28] B. K. Dewi, M. Defriani, and M. A. Sunandar, "Design Of Health Service Mobile Application Interface Using User Centered Design Method," *Sinkron*, vol. 8, no. 1, pp. 231–240, Jan. 2023, doi: 10.33395/sinkron.v8i1.11931.

[29] M. R. F. Rahmatullah, P. N. Andono, Affandy, and M. A. Soeleman, "Improving Random Forest Performance for Sentiment Analysis on Unbalanced Data Using SMOTE and BoW Integration: PLN Mobile Application Case Study," *Scientific Journal of Informatics*, vol. 12, no. 1, pp. 1–10, Apr. 2025, doi: 10.15294/sji.v12i1.19295.

[30] M. C. Untoro and M. Farhan, "Sentiment Analysis of Public Opinion on BAWASLU Using Random Forest and Particle Swarm Optimization," *Scientific Journal of Informatics*, vol. 12, no. 1, pp. 171–182, Jun. 2025, doi: 10.15294/sji.v12i1.22234.

[31] T. Olaleye, P. Ugege, O. Okewale, O. Akinade, S. Akintunde, and C. Akparanta, "Evaluation of Vader and MultiLingual sentiment analyzers for opinion analytics using graphical illustrations," in *COLINS-2022:6th International Conference on Computational Linguistics and Intelligent Systems*, Gliwice, Poland: CEUR Workshop Proceedings, May 2022.

[32] G. Mihai, "Sentiment Analysis of Students' Feedback Using the VADER Model for Teaching Process Assessment," *Annals of Dunarea de Jos University of Galati. Fascicle I. Economics and Applied Informatics*, vol. 30, no. 3, pp. 196–203, Dec. 2024, doi: 10.35219/eai15840409445.

[33] Y. A. Singgalen, "Sentiment and Toxicity Analysis of Digital Content Using Perspective, Vader, and TextBlob: Tourism and Birdwatching," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 5, no. 1, pp. 142–153, 2024, doi: 10.30865/klik.v5i1.2091.

[34] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 4, Apr. 2024, doi: 10.1016/j.jksuci.2024.102048.

[35] R. Alabduljabbar, "User-centric AI: evaluating the usability of generative AI applications through user reviews on app stores," *PeerJ Comput. Sci.*, vol. 10, 2024, doi: 10.7717/peerj-cs.2421.

[36] K. Shear *et al.*, “Usability Testing of an Interoperable Computerized Clinical Decision Support Tool for Fall Risk Management in Primary Care,” *Appl. Clin. Inform.*, vol. 14, no. 2, pp. 212–226, Mar. 2023, doi: 10.1055/a-2006-4936.
[37] F. A. Kusumo, D. R. S. Saputro, and P. Widyaningsih, “SENTIMENT ANALYSIS OF REVIEWS ON X APPS ON GOOGLE PLAY STORE USING SUPPORT VECTOR MACHINE AND N-GRAM FEATURE SELECTION,” *Barekeng*, vol. 19, no. 2, pp. 1037–1046, Jun. 2025, doi: 10.30598/barekengvol19iss2pp1037-1046.
[38] M. Kumar, L. Khan, and H. T. Chang, “Evolving techniques in sentiment analysis: a comprehensive review,” *PeerJ Comput. Sci.*, vol. 11, 2025, doi: 10.7717/PEERJ-CS.2592.
[39] A. Al-Adaileh, M. Al-Kfairy, M. Tubishat, and O. Alfandi, “A sentiment analysis approach for understanding users’ perception of metaverse marketplace,” *Intelligent Systems with Applications*, vol. 22, Jun. 2024, doi: 10.1016/j.iswa.2024.200362.
[40] P. Weichbroth, “Usability Issues With Mobile Applications: Insights From Practitioners and Future Research Directions,” *ArXiv*, Feb. 2025, doi: 10.1109/ACCESS.2025.3573503.