JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

STUDY ON THE QUALITY OF SERVICE OF THE MOBILE-BASED JKN APPLICATION: A SENTIMENT ANALYSIS APPROACH

Muhammad Farid Wajdi*[1], Dedi I. Inan[2], Ratna Juita[3], Marlinda Sanglise[4]

1. Informatics Engineering, Faculty of Engineering, Papua University, Indonesia
2. Informatics Engineering, Faculty of Engineering, Papua University, Indonesia
3. Informatics Engineering, Faculty of Engineering, Papua University, Indonesia
4. Informatics Engineering, Faculty of Engineering, Papua University, Indonesia

Article Info
Keywords: Lexicon Classification; Mobile JKN; Sentiment Analysis; Service Quality

Article history:
Received 2 June 2024
Revised 20 July 2024
Accepted 4 August 2024
Available online 1 September 2024

DOI :
https://doi.org/10.29100/jipi.v9i3.5757

* Corresponding author.
Corresponding Author
E-mail address:
muh.faridwajdi201965028@gmail.com

ABSTRACT
This study aims to analyze user sentiment towards the quality of the mobile-based JKN (*Jaminan Kesehatan Nasional*) service. The sentiment analysis method used is lexicon classification. A total of 147,199 application reviews were collected through scraping techniques using Python. The reviews were classified into four service quality categories: performance, user interface, security, and customer service. The analysis results show that the majority of users are satisfied with the performance of the JKN mobile application, with 57.32% positive sentiment, 5.47% neutral, and 37.25% negative. Although the majority of service quality sentiment is positive, there are still aspects that need improvement, such as security and user interface. The calculation results of recall, precision, and F1-score indicate that the model has good sensitivity and balance in identifying relevant sentiments. This study provides feedback to the developers, namely *Badan Penyelenggara Jaminan Sosial* (BPJS) *Kesehatan*, for the improvement and enhancement of the Mobile JKN application, considering that the negative sentiment is still relatively significant.

.

I. INTRODUCTION

IN the current era of modern society, many mobile applications are used across various sectors, including the healthcare sector. One commonly used mobile application in Indonesia's healthcare sector is mobile JKN. This application was launched on November 15, 2017, by *Badan Penyelenggara Jaminan Sosial* (BPJS) *Kesehatan* for participants in *Jaminan Kesehatan Nasional* (JKN) program. The digital-based utilization of mobile JKN in Indonesia aims to achieve service excellence and increase participation in JKN BPJS & KIS (*Kartu Indonesia Sehat*) programs [1]. Mobile JKN provides facilities that make it easier for users to register and edit their personal information, as well as easily access family member data. Additionally, this application simplifies user access to facilities provided by BPJS [2]. As one of the healthcare service applications, mobile JKN is certainly required to maintain its service quality to enhance user satisfaction and ensure the continued use of the application. It is important to understand how users feel about this application because user satisfaction with an application is closely related to its quality [3]. Knowing user satisfaction with the service quality of mobile JKN allows for identifying areas that need improvement or enhancement within the application. By paying attention to user feedback, developers can make relevant improvements and developments to enhance the overall user experience. User satisfaction with service quality directly impacts how often and how long users will use the service.

Service quality is considered good if consumer expectations and needs can be met. Service quality in an application has a direct impact on user satisfaction [4]. Therefore, research on user sentiment towards the service quality of mobile JKN is necessary. This research provides information about user sentiment, particularly regarding the service quality of this application. This information can be used to improve the service quality of mobile JKN and address existing issues, thereby increasing the use of this application. One way to understand the sentiment of mobile JKN users is by analyzing the quality and satisfaction based on their experience using the application. This user experience is obtained from their reviews on the Google Play Store platform. The reviews and comments uploaded on this platform contain various things felt by each user and reflect the level of user satisfaction, the problems faced, and their views on the quality of this application. Therefore, this forms the basis for researchers to analyze user opinions on that platform to understand user sentiment towards the application. Sentiment analysis is a widely used method for capturing user opinions and emotions regarding a service [5]. In this case, it pertains to the opinions of JKN users provided through a mobile platform. Sentiment analysis of service quality can be used to determine how satisfied or dissatisfied users are with the features available on the JKN mobile app or if there are any issues in the application that require immediate attention. In this study, the sentiment analysis method used is lexicon-based classification with three data rating labels: positive, negative, and neutral. Using this method, sentiment sentence data can be more easily classified into positive, negative, and neutral labels. Lexicon classification works by selecting important words in the text and then grouping those words into labels such as positive, negative, and neutral based on a lexicon dictionary [6]. In lexicon classification, the lexicon dictionary is a list of words that have already been assigned specific sentiment labels. Words in the text that match entries in the lexicon dictionary will be given the same label as the word in the dictionary. For example, words like "good" or "excellent" are labeled as positive. Lexicon classification is very effective for analyzing short texts, such as app reviews on the Google Play Store. In sentiment analysis, the lexicon classification method can provide good results, especially for short documents [7]. In short texts, the number of words available for analysis is limited, making a simple yet effective approach like lexicon classification more suitable compared to other sentiment analysis methods. Additionally, this method is faster and does not require intensive computation, making it suitable for analyses that need quick and efficient results. Therefore, the lexicon classification method is considered more suitable for this study because the data used are short documents and do not require a large training dataset.

Several previous studies on sentiment analysis have been conducted. For example, sentiment analysis on users of the MyPertamina app only classified sentiments into two categories: positive and negative. The classification method applied used the Naïve Bayes Classifier (NBC) algorithm [8]. Sentiment analysis using review data from the Grab app, collected from the Google Play Store platform, employed the Support Vector Machine (SVM) classification method with two measurement scale labels, positive and negative, achieving an accuracy of 85.54% [9]. Research on user sentiment towards the Brimo app using the Naïve Bayes Classifier method resulted in an accuracy of 84.52%, precision of 82.51%, and recall of 87.62% [10]. Sentiment analysis of online transportation app users using the Support Vector Machine (SVM) algorithm achieved an accuracy of 87%, precision of 93%, and recall of 83% [11]. Sentiment analysis of the Mobile JKN app also follows a similar approach by using review data obtained from the Google Play Store platform and focusing on two measurement scale labels: positive and negative. This analysis applies the Support Vector Machine (SVM) algorithm to identify the sentiment of user reviews of the Mobile JKN app. The results of this study showed accuracy, precision, and recall values of 89.53%, 88.17%, and 45.96%, respectively [12].

However, from the five previous studies, it is known that they only used two sentiment labels. This means that those studies did not consider neutral sentiment, which many users may have. By adding a neutral sentiment label, we can obtain a more comprehensive understanding of users' opinions regarding the quality of the application. This study complements earlier research by incorporating three sentiment labels: positive, negative, and neutral. Thus, this study aims to fill the gap in previous research that only used two sentiment labels. The addition of a neutral sentiment label allows for a deeper analysis of users' feelings towards the Mobile JKN application. By expanding the sentiment categories, this research provides a more complete perspective on user perceptions, which in turn can help in the development and enhancement of the quality of Mobile JKN application services. Previous sentiment analyses also focused solely on general user sentiments towards the application usage, unlike this study, which focuses more on evaluating the quality of services provided, offering deeper insights into how users respond to features and performance of the Mobile JKN application.

This study aims to generate research data in the form of positive, negative, and neutral sentiments to assess user satisfaction with the quality of service of the Mobile JKN application. The results obtained from this research are important as they provide information regarding user sentiments towards the quality of this application's services, which can be used as feedback for the developer, BPJS Kesehatan. This research is expected to provide deep insights to BPJS Kesehatan regarding aspects of the application that need improvement, enabling targeted enhancements to increase user satisfaction. Moreover, this information can also aid in the development of new features that better meet user needs and expectations. This study is also anticipated to contribute to the literature by offering a more comprehensive analysis of user sentiments through the use of three sentiment labels (positive, negative, and neutral). This approach can serve as a reference for future research in sentiment analysis, particularly in the context of healthcare service applications.

II. METHOD

This research employs a quantitative approach. Quantitative approach is a research method that gathers and

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

analyzes data and presents it in numerical form [13]. The quantitative approach emphasizes systematic measurement and data analysis [14]. In the context of this research, the quantitative approach is applied by collecting review data from the Google Play Store and classifying them into positive, negative, and neutral sentiments. Numerical data is used to measure sentiments by calculating the frequency and distribution of each sentiment category. The results of this classification will be presented in the form of tables, diagrams, or graphs to provide a clear overview of user perceptions towards the Mobile JKN application.

In this research, sentiment analysis or opinion mining is essential to understand how users feel about the quality of Mobile JKN services using the lexicon classification method. The lexicon classification method is chosen because it can quickly identify and classify sentiments based on a predefined dictionary of words. This method is suitable for analyzing short text data, such as app reviews, due to its efficiency in processing large amounts of data with adequate accuracy compared to other methods like machine learning, which require extensive training data. With lexicon classification, words in review texts are identified and labeled as positive, negative, or neutral based on sentiment scores predefined in the lexicon dictionary.

The purpose of sentiment analysis is to conduct research on the emotions, attitudes, opinions, and evaluations expressed by the public or experts through various media about a product, brand, service, politics, or an institution [15]. In this research, sentiment analysis is utilized to classify reviews of the Mobile JKN application into positive, negative, and neutral sentiments.

The process of sentiment analysis is divided into three steps: classification, evaluation, and data visualization.

1. Classification

   Classification is used to retrieve and group comments or reviews of the Mobile JKN application from the Google Play Store platform. The benefit of this stage is to enable researchers to identify and classify the types of comments for further analysis [4].

2. Evaluation

   Evaluation involves reviewing the dataset through data preprocessing, classification method, and evaluation process [16].

3. Data Visualization

   In this stage, the processed data from the evaluation phase will be presented in the form of graphs, histograms, or matrices. Additionally, visualization can also take the form of a word cloud or interactive map, depending on the analysis results, as they can vary greatly [16].

A. Research Flow

The research flow in Figure 1 illustrates how this study was conducted. The research flow is designed to minimize the risk of chaos during the research process. It serves as a guide for researchers in planning and implementing their research agenda. The aim is to enable researchers to conduct structured and timely research, ensuring that the research progresses as expected [17].

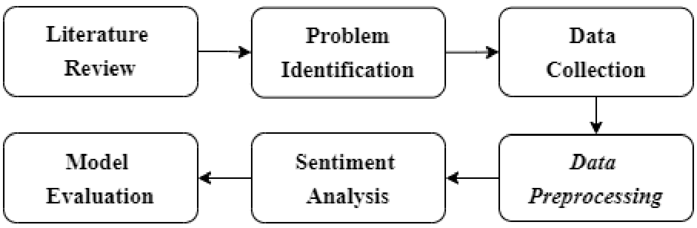


Fig. 1. Research Flow

The initial steps of this research are literature collection and problem identification. The literature review is conducted to strengthen and support existing theories as well as to gather sufficient information to complement the research [18]. This study employs a literature review, including journals related to the research being conducted. Meanwhile, problem identification is undertaken because this step determines the focus and direction of the research, making it crucial. By understanding the problem to be solved, researchers can formulate clear and specific objectives.

Following that, data collection is performed from the reviews of the Mobile JKN mobile application on the Google Play Store. The data collection process is carried out using data scraping techniques with Python programming. Scraping information is used to filter desired information, thereby facilitating the search process [19]. In this research, we utilized the google-play-scraper library to scrape review data from the Google Play Store. This library facilitates the rapid and efficient collection of review data. However, due to the substantial amount of data involved, specifically 147,199 reviews, the scraping process required a considerable amount of time to complete the data collection. Therefore, the solution implemented included using efficient scripts to reduce the scraping time.

Once all data is gathered, preprocessing is conducted to select the data to be used. Preprocessing is a data input management process that initially is unstructured, with the aim of transforming it into structured data before sentiment analysis [20]. In this preprocessing stage, there are five steps to be performed, namely case folding, cleaning, removing stopwords, tokenization, and stemming. After passing through this stage, the review data, which initially numbered 147,199, is filtered down to 119,783 review data.

Next is the sentiment classification process using the lexicon classification method. This stage begins with assigning sentiment labels to each review. This is done using an existing lexicon dictionary containing words with sentiment scores such as positive, neutral, and negative. Then, the lexicon dictionary is compared with the words in each review. After obtaining the sentiment labels, the next step is to classify words related to service quality. This is done by comparing each review that has been labeled with sentiment with an existing service quality dictionary, which contains service quality keywords. After obtaining the service quality keywords, the keywords are classified into predefined parameters. The next step is after the data has been classified, the data will be analyzed and visualized.

Next, we proceeded with calculating recall, precision, and F1-score to evaluate the results. These metrics recall, precision, and F1-score are relevant for evaluating the analysis results as they provide a deep understanding of how well the model has succeeded in accurately classifying review sentiments. Recall calculation is performed to assess how effective the classification method is in predicting positive comments among all actual positive comments in the dataset. In the context of sentiment analysis, a high recall indicates that the model is capable of correctly identifying a significant portion of reviews with the correct sentiment. Equation (1) is used to compute the recall value [21].

$$\text{Recall (R)} = \frac{\text{a}}{\text{a + c}} \times 100 \tag{1}$$

The calculation of precision is conducted to evaluate how accurately the model predicts positive, negative, and neutral data. High precision indicates that most reviews classified as positive, negative, or neutral by the model are indeed relevant to the intended sentiment. Equation (2) is used to calculate the precision value [21].

$$\text{Precision (P)} = \frac{\text{a}}{\text{a + b}} \times 100 \tag{2}$$

To measure the accuracy and balance between precision and recall, we calculate the F1-score. A high F1-score indicates that the model has a good balance between recall and precision. Equation (3) is used to compute the F1-score [21].

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

In general, a model is considered to have good performance if its recall, precision, and F1-score values approach 100%. High recall, precision, and F1-score indicate that the model is capable of accurately identifying and classifying review sentiments. Higher values indicate better performance. Recall and precision matrices can be seen in Table 1 [4].

TABLE I
RECALL AND PRECISION MATRIX

Document	Relevant	Irrelevant	Total
Found	a (hits)	b (noise)	a+b
Not Found	c (misses)	d (rejected)	c+d
Total	a+c	b+d	a+b+c+d

Explanation:
a (hits) = relevant document
b (noise) = irrelevant document
c (misses) = relevant document not found
d (rejected) = irrelevant document not found

III. RESULTS AND DISCUSSION

This section discusses the results of the research regarding user sentiment towards the quality of service of the Mobile JKN application. The results are derived from the collected data and also based on the analysis of user reviews available on the Google Play Store. This discussion begins by explaining the data collection and processing processes undertaken before sentiment classification and analysis. The results of the sentiment analysis that have been conducted will be explained, along with the evaluation of the performance of the model used. Subsequently, conclusions will be drawn based on the results of the analysis.

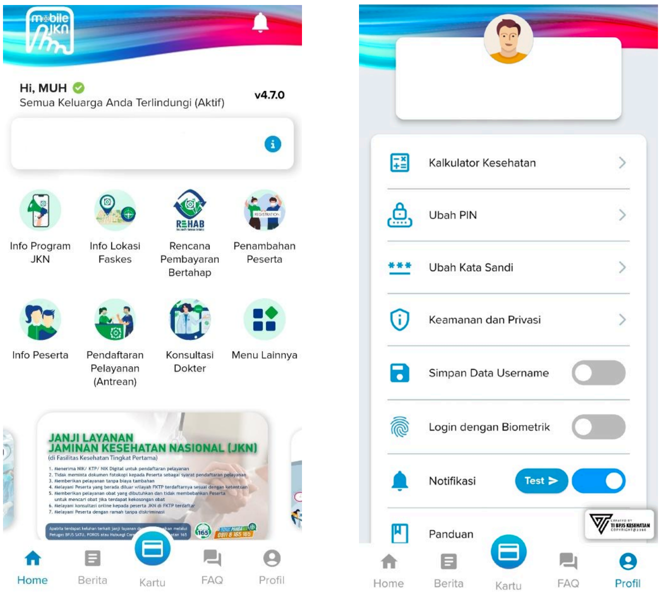
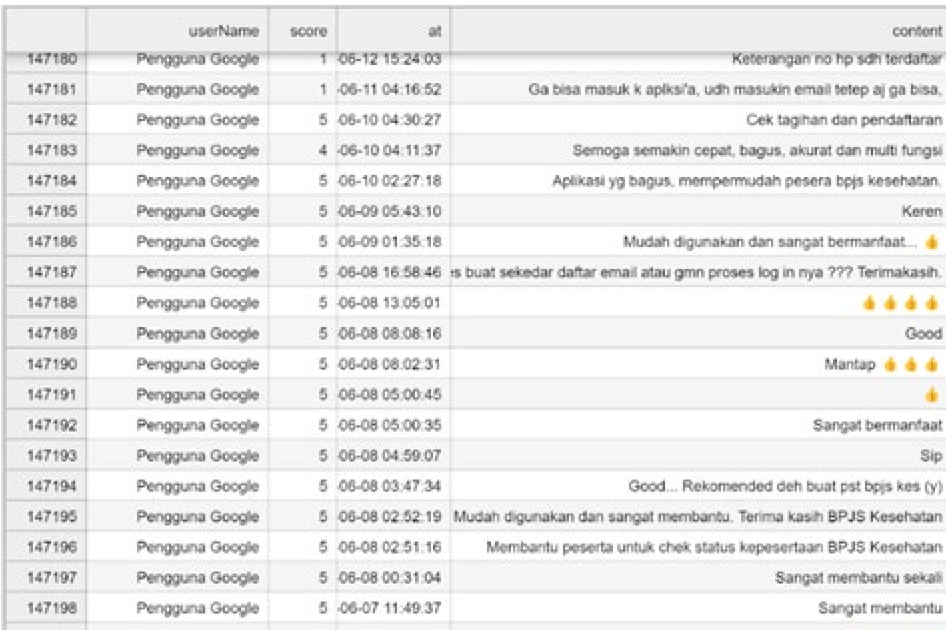


Fig. 2. Mobile JKN Application Interface

A. Data Retrieval

The reviews of the Mobile JKN mobile application in this study were obtained using scraping techniques with Python programming. The total number of successfully retrieved data was 147,199 reviews. The retrieved data includes the username of the reviewer, rating, the time when the user uploaded the review, and the content of the review. Out of the total 147,199 retrieved data, 119,783 were usable for analysis. This is because many reviews either lacked meaningful content or were composed solely of emoticons, rendering them unusable for research purposes. An example of the scraped data can be seen in Figure 3.



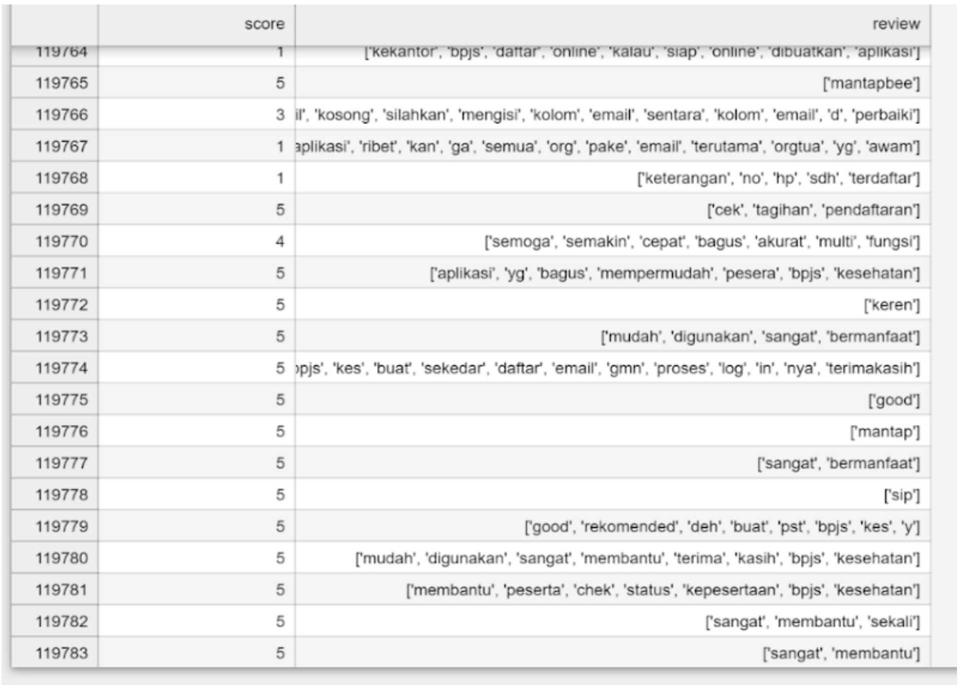
	userName	score	at	content
147180	Pengguna Google	1	-06-12 15:24:03	Keterangan no hp sdh terdaftar
147181	Pengguna Google	1	-06-11 04:16:52	Ga bisa masuk k apliksi'a, udh masukin email tetep aj ga bisa,
147182	Pengguna Google	5	-06-10 04:30:27	Cek tagihan dan pendaftaran
147183	Pengguna Google	4	-06-10 04:11:37	Semoga semakin cepat, bagus, akurat dan multi fungsi
147184	Pengguna Google	5	-06-10 02:27:18	Aplikasi yg bagus, mempermudah pesera bpjs kesehatan.
147185	Pengguna Google	5	-06-09 05:43:10	Keren
147186	Pengguna Google	5	-06-09 01:35:18	Mudah digunakan dan sangat bermanfaat... 👍
147187	Pengguna Google	5	-06-08 16:58:46	:s buat sekedar daftar email atau gmn proses log in nya ??? Terimakasih.
147188	Pengguna Google	5	-06-08 13:05:01	👍👍👍👍
147189	Pengguna Google	5	-06-08 08:08:16	Good
147190	Pengguna Google	5	-06-08 08:02:31	Mantap 👍👍👍
147191	Pengguna Google	5	-06-08 05:00:45	👍
147192	Pengguna Google	5	-06-08 05:00:35	Sangat bermanfaat
147193	Pengguna Google	5	-06-08 04:59:07	Sip
147194	Pengguna Google	5	-06-08 03:47:34	Good... Rekomended deh buat pst bpjs kes (y)
147195	Pengguna Google	5	-06-08 02:52:19	Mudah digunakan dan sangat membantu. Terima kasih BPJS Kesehatan
147196	Pengguna Google	5	-06-08 02:51:16	Membantu peserta untuk chek status kepesertaan BPJS Kesehatan
147197	Pengguna Google	5	-06-08 00:31:04	Sangat membantu sekali
147198	Pengguna Google	5	-06-07 11:49:37	Sangat membantu

Fig. 3. Scraped Data Result

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

B. Data Processing

After the data has been successfully retrieved, it proceeds to the data preprocessing stage. The data preprocessing process involves steps such as removing special characters, punctuation marks, and irrelevant conjunctions. Additionally, in some cases, text normalization is applied to standardize words with different spelling variations. For example, the word "*menggembirakan*" (delightful) can be normalized to its base form, "*gembira*" (happy). Subsequently, the data will be cleansed of stopwords, which generally do not contribute significantly to sentiment analysis, such as "*dan*" (and), "*atau*" (or), "*dari*" (from), and so on. In this stage, the existing data will be filtered to produce data ready for analysis. The review data, previously in sentence form, will be transformed into token or word form. These steps aim to improve the quality of the data to be used in subsequent sentiment analysis processes. The data after preprocessing can be seen in Figure 4.



	score	review
119764	1	['kekantor', 'bpjs', 'daftar', 'online', 'kalau', 'siap', 'online', 'dibuatkan', 'aplikasi']
119765	5	['mantapbee']
119766	3	il', 'kosong', 'silahkan', 'mengisi', 'kolom', 'email', 'sentara', 'kolom', 'email', 'd', 'perbaiki']
119767	1	aplikasi', 'ribet', 'kan', 'ga', 'semua', 'org', 'pake', 'email', 'terutama', 'orgtua', 'yg', 'awam']
119768	1	['keterangan', 'no', 'hp', 'sdh', 'terdaftar']
119769	5	['cek', 'tagihan', 'pendaftaran']
119770	4	['semoga', 'semakin', 'cepat', 'bagus', 'akurat', 'multi', 'fungsi']
119771	5	['aplikasi', 'yg', 'bagus', 'mempermudah', 'pesera', 'bpjs', 'kesehatan']
119772	5	['keren']
119773	5	['mudah', 'digunakan', 'sangat', 'bermanfaat']
119774	5	pjs', 'kes', 'buat', 'sekedar', 'daftar', 'email', 'gmn', 'proses', 'log', 'in', 'nya', 'terimakasih']
119775	5	['good']
119776	5	['mantap']
119777	5	['sangat', 'bermanfaat']
119778	5	['sip']
119779	5	['good', 'rekomended', 'deh', 'buat', 'pst', 'bpjs', 'kes', 'y']
119780	5	['mudah', 'digunakan', 'sangat', 'membantu', 'terima', 'kasih', 'bpjs', 'kesehatan']
119781	5	['membantu', 'peserta', 'chek', 'status', 'kepesertaan', 'bpjs', 'kesehatan']
119782	5	['sangat', 'membantu', 'sekali']
119783	5	['sangat', 'membantu']

Fig. 4. Preprocessed Data

C. Service Quality Dictionary

The service quality dictionary is a collection of keywords created based on user experience related to the quality of service of the application. The function of this service quality dictionary is to identify and classify sentiments found in user reviews. This dictionary contains words referring to predefined parameters, such as "performance", "user interface", "security", and "customer service". With these parameters, it is possible to efficiently understand how users respond to or evaluate the Mobile JKN application across various service aspects. For details, refer to Table 2.

As seen in Table 2, the existing keywords have been classified based on the predefined parameters. For example, in the review data, there is a sentence "*pelayanan baik, petugas ramah dan amanah*". The word "*ramah*" is categorized under the customer service parameter because it indicates how customer service in the Mobile JKN application is perceived to be good in treating users. Another example is the frequently commented word "error", as seen in the comment "*milih faskes* error *terus*". The word "error" refers to the performance of the Mobile JKN application, which is considered inadequate, thus it is included under the performance parameter.

TABLE II
SERVICE QUALITY LIBRARY

Service Quality	Library
Performance	update (9954), *mantap* (7845), good (5383), *oke* (2821), *eror* (1136), *perbaiki* (1132), *lemot* (1035), top (873), *puas* (751), *gagal* (641), *kurang* (695), error (555), *lumayan* (596), nice (478), *cepat* (467), *payah* (392), upgrade (380), *lancar* (397), *parah* (368), loading (313), *berat* (157), bad (149), bug (130), *rusak* (111), *lelet* (106), optimal (45), *kacau* (43), *canggih* (42), *bobrok* (35), amazing (27), maintenance (24), *efektif* (26), trouble (22), *berkualitas* (19)

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

User Interface	*bagus* (10509), *sesuai* (1327), *keren* (811), *perbaharui* (554), *aneh* (550), *jelek* (434), *tampilan* (210), *simpel* (201), *gampang* (161), *kece* (156), simple (132), *tidak jelas* (118), *rumit* (81), *ruwet* (68), *sempurna* (63), *nyaman* (59), easy (51), *cakep* (53), *sederhana* (24), ui/ux (2)
Security	*aman* (220), login (1799), captcha (1173), *verifikasi* (637), *kode* (246), password (185), *akun* (153), *registrasi* (72), register (63), logout (38), *sandi* (36)
Customer Service	*bantu* (12790), *susah* (5370), *mudah* (5593), *baik* (3499), *ribet* (2801), *berguna* (873), *memuaskan* (838), *tidak bisa* (802), *bayar* (704), *sulit* (623), *buruk* (523), *tingkatkan* (736), best (259), *praktis* (260), *ramah* (142), like (127), *lambat* (101), *rugi* (111), *berbayar* (100), *efisien* (82), repot (58), *tolol* (56), *nyebelin* (28), *terlayani* (10)

According to Table 2, for reviews based on "performance," the word "update" is the most frequent, with a total of 9,954 reviews, while the word "*berkualitas*" is the least frequent, with only 19 reviews. For reviews related to "user interface," the word "*bagus*" appears most frequently, totaling 10,509 reviews, whereas "*ui/ux*" is the least frequent, with only 2 reviews. In terms of "security," the word "*aman*" appears most frequently, with a total of 220 reviews, and "*sandi*" is the least frequent, with 36 reviews. Regarding "customer service," the word "*bantu*" has the highest frequency, with a total of 12,790 reviews, while "*terlayani*" has the lowest frequency, with only 10 reviews.

D. Analysis of Data

Based on the data collection results, the number of reviews is grouped according to ratings, as shown in Table 3.

TABLE III
NUMBER OF REVIEWS BASED ON SCORES

Rating	Number of Reviews
★	37.882
★ ★	6.732
★ ★ ★	6.551
★ ★ ★ ★	8.562
★ ★ ★ ★ ★	60.056
Total	119.783

In Table 4, the data is grouped into positive, neutral, and negative reviews. These reviews are categorized by dividing the scores into 3 categories: 1-2 stars are categorized as "negative", 3 stars are categorized as "neutral", and 4-5 stars are categorized as "positive".

TABLE IV
NUMBER OF REVIEWS BASED ON POSITIVE, NEUTRAL, AND NEGATIVE SENTIMENTS

Rating	Category	Number of Reviews
★ - ★ ★	Negative	44.614
★ ★ ★	Neutral	6.551
★ ★ ★ ★ - ★ ★ ★ ★ ★	Positive	68.618

Based on Table 4, there are 44,614 reviews with 1-2 stars categorized as negative sentiment, accounting for 37.25% of the total reviews. For 3-star ratings categorized as neutral, there are 6,551 reviews, or 5.47% of the total reviews. The 4-5-star ratings categorized as positive have 68,618 reviews, or 57.32% of the total reviews.

Table 5 is obtained from the grouping of keywords based on the quality of service parameters established in Table 2. This table provides an overview of how many users provide reviews related to specific aspects of the quality of service of the Mobile JKN application, such as performance, user interface, security, and customer service.

ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

TABLE V
NUMBER OF REVIEWS BASED ON SERVICE QUALITY

Service Quality	Number of Reviews
Performance	37.148
User Interface	15.564
Security	4.622
Customer Service	36.486
Others	25.963

According to the data in Table 5, "performance" has the highest number of reviews with a total of 37,148 reviews or 31% of the total reviews. Meanwhile, the "security" parameter has the lowest number, with 4,622 reviews or 3.86% of the total reviews. The "user interface" parameter has 15,564 reviews or 13% of the total reviews, and "customer service" has 36,486 reviews or 30.46% of the total reviews. As for "others," there are 25,963 reviews or 21.67% of the total data available. The "others" parameter includes reviews that do not have keywords based on the created service quality dictionary.

Next, the distribution of the number of reviews based on service quality will be visualized. The aim is to further explain the distribution of the existing data. Visualizations such as line graphs and pie charts are used in this study. The visualization of data can be seen in Figure 5 for the line graph and Figure 6 for the pie chart.

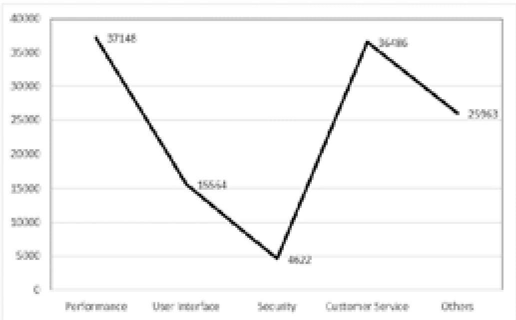


Fig. 5. Data Distribution Graphic

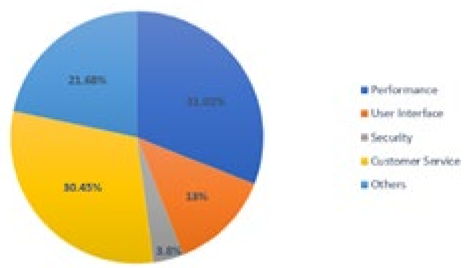


Fig. 6. Data Distribution Diagram

E. Sentiment Analysis of User Satisfaction

The keywords that have been grouped based on the quality of service parameters will be further categorized based on positive, neutral, and negative sentiments obtained from the lexicon classification results.

TABLE VI
NUMBER OF POSITIVE, NEUTRAL, AND NEGATIVE REVIEWS

Service Quality	Positive	Neutral	Negative
Performance	24.911	3.818	8.419
User Interface	12.696	741	2.127
Security	1.954	874	1.794
Customer Service	25.791	1.755	8.940
Others	16.497	2.019	7.447

In Table 6, reviews that have passed the lexicon classification stage will be grouped according to the sentiment results obtained. Based on this grouping, the results for "performance" show that there are 24,911 positive reviews, accounting for 67.06%, 8,419 negative reviews, accounting for 22.66%, and 3,818 neutral reviews, accounting for 10.28%. This indicates that the majority of users are satisfied with the performance of the Mobile JKN application. Although there are negative reviews about the application's performance, the number is relatively low compared to the positive reviews. A high percentage of positive sentiments indicates that the application's performance has met

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

the expectations of many users, such as speed, stability, and responsiveness. Keywords like "*mantap*" and "good" demonstrate user appreciation for the application's perfomance. However, despite the relatively low percentage of negative reviews, their number remains significant, with keywords such as "error," "*lemot*," and "*gagal*" indicating issues that need addressing, such as bugs, crashes, or long loading times. This data can be used to prioritize improvements and enhance application performance. For instance, focusing on optimizing application speed and reducing crash frequency would be beneficial.

For "user interface," there are 12,696 positive reviews, accounting for 81.57%, 2,127 negative reviews, accounting for 13.67%, and 741 neutral reviews, accounting for 4.76%. The majority of users consider the user interface of the Mobile JKN application to be good and easy to use. Although most users are satisfied with the application, there are still some negative reviews. This indicates that there is room for improvement in the interface design. The majority of users are satisfied with the user interface, indicating that the application is user-friendly and intuitive. Keywords such as "*bagus*", "*sesuai*", and "*keren*" support this view. However, there are 13.67% negative reviews that may indicate issues such as unappealing design, navigation difficulties, or inconsistent layout, with keywords like "*aneh*", "*jelek*", and "*rumit*". Improvements can be focused on specific aspects mentioned in the negative reviews. For example, enhancing visual aesthetics or optimizing the layout for a better user experience.

For "security," there are 1,954 positive reviews, accounting for 42.27%, 1,794 negative reviews, accounting for 38.81%, and 874 neutral reviews, accounting for 18.91%. The majority of users are satisfied with the security provided by the application. However, there is a significant percentage of negative reviews. This suggests that improvements in security aspects may be necessary to enhance user trust. More than half of the users are satisfied with the security aspects, indicating that some existing security measures are adequate. Keywords such as "*aman*" and "*verifikasi*" indicate this. However, nearly 40% of negative reviews suggest significant concerns regarding user data security, with keywords like "login" and "captcha" frequently mentioned. A comprehensive evaluation of the security system is necessary, including data encryption, protection against cyber attacks, and stronger privacy policies. Users should also be educated about the implemented security measures to enhance their trust.

For "customer service," there are 25,791 positive reviews, accounting for 70.68%, 8,940 negative reviews, accounting for 24.50%, and 1,755 neutral reviews, accounting for 4.81%. The positive percentage for customer service is quite high, indicating that the customer service of the application is highly appreciated by users. However, there is also a need to improve customer service to further enhance user satisfaction. The high percentage of positive reviews indicates that the customer service of this application is quite responsive and helpful in resolving user issues, with keywords such as "*bantu*" and "*baik*". However, the nearly one-quarter percentage of negative reviews suggests that there are aspects of customer service that need improvement, such as response time, quality of assistance, or service attitude, with keywords like "*susah*", "*ribet*", and "*buruk*". Regular evaluation of staff and user feedback should be conducted systematically. This aims to better understand user needs and improve services based on received feedback.

For "others," there are 16,497 positive reviews, accounting for 63.54%, 7,447 negative reviews, accounting for 28.68%, and 2,019 neutral reviews, accounting for 7.77%. This indicates that the majority of users are satisfied with the application beyond the quality of service. The high percentage of positive reviews outside the main service categories indicates that additional aspects of this application also satisfy users. However, the significant percentage of negative reviews suggests there are various other issues that may not have been identified within the main categories. Using feedback from these categories to identify and address minor or specific issues that may not be covered in the main categories, such as compatibility issues, additional features, or specific usability issues, can help improve overall service quality.

	Kualitas Layanan	Ulasan
1	Performance	hay ..kak admin mohon maaf ya kenapa aplikasi ini gk bisa di buka selalu eror kalau di gunakan,terlalu rumit menurut aku
2	User Interface	UI/UX harap ditingkatkan. GUI mengikuti mobile app tahun ini yang app lain gunakan.
3	Security	Ini kenapa setelah pembaruan, minta pin,, terus coba logout dan masuk lagi tdk bisa, salah pin terus
4	Customer Service	Muncul tagihan, padahal udah di bayar. CS LAMA
5	Others	apaan si gk jlz bner. bikin emosi doang

Fig. 7. Example of Negative Reviews Based on Service Quality

F. Analysis of Recall, Precision and F1-Score Calculation

Recall calculation is conducted to determine how sensitive the model is in detecting sentiments correctly. Meanwhile, precision calculation is performed to assess the model's precision level in finding truly relevant documents. Relevant and irrelevant documents, as well as documents that can be found and not found, can be seen in Table 7.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

TABLE VII
NUMBER OF RELEVANT AND IRRELEVANT DOCUMENTS

Document	Relevant	Irrelevant	Total
Found	65.352	28.468	93.820
Not Found	16.497	9.466	25.963
Total	81.849	37.934	119.783

Based on Table 7, recall, precision, and f1-score calculations are performed.

Recall calculation:

$$\text{Recall (R)} = \frac{a}{a + c} \times 100 \tag{1}$$

$$\text{Recall (R)} = \frac{65.352}{65.352 + 16.497} \times 100$$

$$\text{Recall (R)} = 79{,}84$$

Precision calculation:

$$\text{Precision (P)} = \frac{a}{a + b} \times 100 \tag{2}$$

$$\text{Precision (P)} = \frac{65.352}{65.352 + 28.468} \times 100$$

$$\text{Precision (P)} = 69{,}65$$

F1-score calculation:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

$$F1 - Score = 2 \times \frac{69{,}65 \times 79{,}84}{69{,}65 + 79{,}84}$$

$$F1 - Score = 74{,}23$$

From the calculation results, we obtained a recall value of 79.84%, precision of 69.65%, and an f1-score of 74.23%. With a recall value of approximately 79.84%, these results indicate that the model has relatively high sensitivity in detecting positive, negative, and neutral sentiments. The model can find most of the documents that should be relevant for sentiment analysis.

Although the recall value is quite high, the precision value is 69.65%, meaning that around 30% of the documents considered relevant may not be relevant. This value indicates that the model has a relatively high level of precision in finding truly relevant sentiments, although there is still a possibility that the model is less accurate in identifying truly relevant sentiments.

G. Discussion

This research differs from previous studies, where researchers analyzed sentiment in reviews of the Mobile JKN application focusing only on general usage and using only 2 sentiment labels: positive and negative [12]. In contrast, this study places greater emphasis on the quality of service of the Mobile JKN application and utilizes a more comprehensive set of three sentiment labels: positive, negative, and neutral. By concentrating on the quality of service of the application, it allows for a more comprehensive analysis of user responses regarding application quality aspects such as performance, user interface, security, and customer service. By incorporating more nuanced sentiment nuances, this research aims to provide a deeper understanding of user evaluations of the Mobile JKN application as a whole.

For this research, the initial step involved collecting review data. The reviews were gathered using scraping methods, resulting in a total of 147,199 reviews. Once the data was successfully collected, preprocessing was conducted. During the preprocessing stage, the original 147,199 reviews were filtered down to 119,783 reviews. Following this, a service quality dictionary was created containing relevant keywords to classify user sentiment towards the Mobile JKN application. This dictionary was developed based on parameters such as "performance," "user interface," "security," and "customer service," allowing researchers to understand how users responded to various aspects of the application's services.

After creating the dictionary, the next step is to classify sentiments based on the dictionary that has been established. Words in the reviews are categorized as positive, negative, or neutral sentiments according to the service quality dictionary. For example, the word "friendly" is classified under the customer service parameter, while "error" is associated with application performance. Following classification, the number of reviews for each sentiment category is analyzed. This includes creating Table 4, which categorizes reviews based on star ratings into "positive,"

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 3, September 2024, Pp. 1506-1517

"neutral," and "negative" sentiment categories. Furthermore, the distribution of reviews based on service quality is presented in Table 5, and visualizations such as line graphs and pie charts are used to illustrate data distribution more clearly. Further sentiment analysis is conducted by calculating the number of positive, neutral, and negative reviews for each service quality parameter, as shown in Table 6. These analyses provide insights into how users respond to each aspect of the Mobile JKN application.

In the final part of the study, recall, precision, and F1-score are calculated to evaluate the model's performance in accurately detecting sentiment. A recall value of 79.84% indicates that the model has high sensitivity in identifying relevant reviews. Meanwhile, a precision of 69.65% demonstrates a good level of accuracy in classifying the reviews identified as truly relevant. The F1-score result of 74.23% shows that the model achieves a balanced performance between recall and precision.

However, this research faces limitations in both the scraping method and the lexicon classification used. Firstly, concerning scraping, the primary limitation lies in the significant amount of duplicate data retrieved from the Google Play Store platform. Secondly, regarding lexicon classification, the main limitations are the accuracy and comprehensiveness of the keyword list used for sentiment classification. Lexicon classification relies on a predefined list of words associated with positive, negative, or neutral sentiments. However, lexicons may not include new words or variations of existing words, potentially resulting in the loss of information or richness of sentiments expressed by users.

These limitations can impact the analysis and interpretation in several ways. First, duplicated data can affect the number of keywords available for use, potentially reducing the accuracy of classification. Therefore, preprocessing is necessary before conducting analysis [22]. Second, imperfect lexicon classification can lead to errors in sentiment identification, such as categorizing neutral sentences as positive or negative. This can reduce the accuracy of interpretations regarding how users actually respond or evaluate the quality of service of the Mobile JKN application.

To improve the classification model's performance and achieve better recall, precision, and F1-score results in the future, several steps can be taken. Comprehensive text cleaning and handling of slang words can enhance sentiment classification accuracy. Maintaining the service quality dictionary is also crucial; it should be regularly updated by adding new keywords from recent reviews and removing irrelevant ones to improve accuracy. Additionally, exploring alternative methods is expected to enhance recall, precision, and F1-score outcomes.

This study also indicates the need for better methods to classify user review perceptions related to satisfaction and service quality. This is because the classification of neutral reviews is still significant, at nearly 5.5%, and reviews that are not categorized under service quality parameters amount to 21.67%. Implicitly, this highlights the necessity for representative keywords to accurately depict service quality. Additionally, improved methods are required to discern neutral opinions. This is crucial because cumulatively, these two aspects constitute a considerable portion but cannot be utilized as actionable insights for development to enhance the quality of the Mobile JKN application's services.

Further research suggests exploring alternative sentiment classification methods, such as employing machine learning or deep learning techniques. By utilizing these alternative methods, it is expected to achieve better performance in terms of recall, precision, and F1-score values.

IV. Conclusion

This study was conducted to understand user sentiment regarding the quality of service of the Mobile JKN application through user reviews on the Google Play Store platform. The method employed was lexicon classification with the aim of evaluating user satisfaction with the Mobile JKN application. From the analysis of the data, it was evident that the majority of users are satisfied with the performance of the Mobile JKN application. Most reviews (57.32%) expressed positive sentiments, while only a small fraction (37.24%) expressed negative sentiments. Furthermore, the analysis based on service quality parameters indicated that the majority of users rated the quality of Mobile JKN service as very good. However, there are areas that need improvement, such as user interface and security. Based on the results of recall, precision, and f1-score calculations, it was found that this model exhibits high sensitivity in detecting positive, neutral, and negative sentiments, with a recall value of 79.84%. The precision value of 69.65% indicates that the model has a relatively high level of precision in identifying truly relevant sentiments, although there is also a relatively high possibility that the model may inaccurately identify truly relevant sentiments, around 30%. However, the f1-score value of 74.23% suggests that the model achieves a relatively good balance between recall and precision. The findings of this research provide a better understanding of user sentiment regarding the quality of services provided by the Mobile JKN application. It is hoped that the results of this analysis can serve as a reference for developers to improve service quality, thereby enhancing user satisfaction with the Mobile JKN application.

REFERENCES

[1] H. Yuliastuti and M. Jawahir, "Analisis Efektivitas Pemanfaatan Aplikasi Layanan Kesehatan Mobile JKN Jaminan Kesehatan Nasional Di Indonesia," *J. Pengabdi. Kpd. Masy. Langit Biru Politek. Penerbangan Indones. Curug*, vol. 4, no. 01, pp. 28–40, 2023.

[2] A. Wulandari, I. S.-J. P. P. Vol, and undefined 2019, "Inovasi bpjs kesehatan dalam pemeberian layanan kepada masyarakat: Aplikasi mobile jkn," *core.ac.uk*, vol. 5, no. 2, 2019, Accessed: Sep. 10, 2023. [Online]. Available: https://core.ac.uk/download/pdf/280498391.pdf

[3] D. I. Inan *et al.*, "Service quality and self-determination theory towards continuance usage intention of mobile banking," *J. Sci. Technol. Policy Manag.*, vol. 14, no. 2, pp. 303–328, 2023, doi: 10.1108/JSTPM-01-2021-0005.

[4] A. N. Hidayanto, K. Y. P. M, and R. Randy, "JURNAL RESTI Quality on Playstore Data : A Case of Indodax," vol. 5, no. 158, pp. 1–6, 2023.

[5] S. K. Hasna, "Analisis Sentimen Data Ulasan Menggunakan Algoritma Support Vector Machine (Studi Kasus: Aplikasi Iflix)," 2021.

[6] N. S. Fathullah, Y. A. Sari, and P. P. Adikara, "Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi Naïve Bayes dengan Fitur Lexicon-Based," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 2, pp. 590–593, 2020, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6987

[7] E. Budianita, E. P. Cynthia, A. Pranata, and D. Abimanyu, "Pendekatan berbasis Machine Learning dan Leksikal Pada Analisis Sentimen," *Semin. Nas. Teknol. Informasi, Komun. dan Ind.*, pp. 99–104, 2022, [Online]. Available: https://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/19137

[8] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 42–48, 2023, doi: 10.54914/jtt.v9i1.609.

[9] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J. Inform.*, vol. 8, no. 2, pp. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.

[10] M. K. Khoirul Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 478–483, 2023, doi: 10.36040/jati.v7i1.6373.

[11] E. Suryati, Styawati, and A. A. Aldino, "Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM)," *J. Teknol. Dan Sist. Inf.*, vol. 4, no. 1, pp. 96–106, 2023, [Online]. Available: https://doi.org/10.33365/jtsi.v4i1.2445

[12] N. Maulida, N. Suarna, and W. Prihartono, "Analisis Ulasan Sentimen Aplikasi Mobile Jkn Dengan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 1651–1658, 2024, doi: 10.36040/jati.v8i2.9105.

[13] N. Wanma, D. I. Inan, and L. Y. Baisa, "Evaluasi User Experience Dan User Interface Aplikasi Laporkitong Dengan End User Computing Satisfaction something that is of great concern and determines the experience of using it for the people of," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, 2024.

[14] H. Ahyar *et al.*, *Buku Metode Penelitian Kualitatif & Kuantitatif*, no. March. 2020.

[15] R. R. Cahyani and R. Cahyani, "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019," *Matics*, vol. 12, no. 1, p. 79, 2020, doi: 10.18860/mat.v12i1.8356.

[16] S. Thomas, Yuliana, and Noviyanti. P, "Study Analisis Metode Analisis Sentimen pada YouTube," *J. Inf. Technol.*, vol. 1, no. 1, pp. 1–7, 2021, doi: 10.46229/jifotech.v1i1.201.

[17] M. Irfan Syahroni, "Prosedur Penelitian Kuantitatif," *eJurnal Al Musthafa*, vol. 2, no. 3, pp. 43–56, 2022, doi: 10.62552/ejam.v2i3.50.

[18] B. Firmansyah, D. I. Inan, R. Juita, and M. Sanglise, "Designing UI/UX Thrift shop Website Using User Centered Design (UCD) Method (Case Study: Backfold Market)," *G-Tech J. Teknol. Terap.*, vol. 8, no. 2, pp. 816–824, 2024, doi: 10.33379/gtech.v8i2.4040.

[19] D. D. A. Yani, H. S. Pratiwi, and H. Muhardi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *J. Sist. dan Teknol. Inf.*, vol. 7, no. 4, p. 257, 2019, doi: 10.26418/justin.v7i4.30930.

[20] R. Fero, "Analisa Metode Pre-Processing Teks Paling Optimal untuk Analisis Sentimen pada Media Sosial Twitter (Case : Produk Iphone," Universitas Telkom, S1 Informatika, 2019.

[21] T. Kanstren, "A Look at Precision, Recall, and F1-Score," Towards Data Science, 12 10 2020. [Online]. Available: https://towardsdatascience.com/a-look-at-precision-recalland-f1-score-36b5fd0dd3ec.

[22] A. Z. Rizquina and C. I. Ratnasari, "Implementasi Web Scraping untuk Pengambilan Data Pada Website E-Commerce," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 4, pp. 377–383, 2023, doi: 10.47233/jteksis.v5i4.913.