Factor analysis influencing Mobile JKN user experience using sentiment analysis

Muhammad Yazid Al Qahar, Yova Ruldeviyani, Ulfah Nur Mukharomah, Miftahul Agtamas Fidyawan, Ramadhoni Putra
Department of Information Technology, Faculty of Computer Science, University of Indonesia, Jakarta, Indonesia

Article Info

Article history:

Received May 8, 2023
Revised Sep 29, 2023
Accepted Oct 8, 2023

Keywords:

Mobile JKN
Naïve Bayes
Online service quality factor
Sentiment analysis
SMOTE
Support vector machine

ABSTRACT

Social security administration for health or Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan), as a public legal entity, has a critical role in the health of the Indonesian population. BPJS Kesehatan introduced the Mobile national health insurance or *jaminan kesehatan nasional* (JKN) application to enhance its services, enabling Indonesians to access it directly. Nevertheless, the rating of the Mobile JKN application on the Google Play Store has shown a gradual decline over time. Therefore, this study was conducted to analyze the factors influencing the user experience of the Mobile JKN application, utilizing the review data obtained from the Google Play Store. Sentiment analysis using the Naïve Bayes (NB) classification model and support vector machine (SVM) combined with synthetic minority oversampling technique (SMOTE) and slang word replacement. The results obtained an accuracy value of 93.33%, precision of 93.76%, recall of 93.33%, and F1-score of 93.43%. A further analysis was conducted using online service quality factors to obtain the main factors influencing the experience of Mobile JKN application users. The evaluation findings revealed that factors of security, ease of use, and timeliness are three fundamental aspects that should be given immediate attention by BPJS Kesehatan while improving the Mobile JKN application in the future.

Corresponding Author:

Muhammad Yazid Al Qahar
Department of Information Technology, Faculty of Computer Science, University of Indonesia
UI Salemba Campus, St. Salemba Raya No.4, DKI Jakarta 10430, Indonesia
Email: muhammad.yazid22@ui.ac.id

1. INTRODUCTION

Health is a fundamental societal need protected by the constitution [1]. Thus, the government has implemented several initiatives to enhance public health standards. One of the strategic programs launched by the government is the national health insurance or *jaminan kesehatan nasional* (JKN) through social security administration for health or Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) [2]. The government has made this program one of the requirements for its citizens to access public services [3].

To support this program, BPJS Kesehatan has introduced an innovation by creating the Mobile JKN application, a self-service mobile application based on information technology that participants can easily access anytime and anywhere. Mobile JKN provides services such as registering new participants, updating participant data, the digital membership card, the information channel, and submitting complaints [4]. Along with the high public demand for services on the Mobile JKN application, this is not in line with user satisfaction, as reflected in the Google Play Store application. Ratings for the Mobile JKN application tend to experience a downward trend from 2020 to 2022, as shown in Figure 1 [5]. In addition, the achievement target for Mobile JKN user registration set by BPJS Kesehatan management in 2021, namely 20 million users, was not achieved until December 2021, when only 14.8 million users were recorded [6].

Several research studies have been conducted to ascertain the factors that impact the user experience of the Mobile JKN application, which involve surveying users of the application through a questionnaire [7], [8]. This approach requires a long time and significant resources. Sentiment analysis can improve this situation by examining the review data readily accessible on the Google Play Store application, thereby decreasing the time required and resource expenditure [9].

Sentiment analysis commonly employs machine learning technologies such as support vector machine (SVM) and Naïve Bayes (NB) due to their ability to deliver accurate outcomes with fast processing times while necessitating minimal training data [10]. Several previous studies have conducted sentiment analysis using both classification models to analyze government applications using review data from the Google Play Store [11], in fintech applications using data from Twitter [12], [13], and analyzing the impact of the COVID-19 pandemic using Twitter data [14]–[16]. But in crawling data, class imbalance often occurs. The class imbalance can make the classification value decrease. One of the most frequently used methods for dealing with class imbalance is synthetic minority oversampling technique (SMOTE) [17]–[21].

Apart from conducting sentiment analysis, the stakeholders of a digital service provider company must evaluate the factors that affect the user experience of their services [12]. One of the most used methods is to perform factor analysis using online service quality factors. This approach can provide priority points that need to be considered by company management to improve the quality of its services and increase the company's overall profitability [22].

Based on the above, this research aims to analyze factors influencing the user experience when using the Mobile JKN application with sentiment analysis and online service quality factors. The stage begins by comparing the classification model between NB and SVM, combined with the SMOTE algorithm. Then it continues to perform factor analysis using online service quality factors to find out what topics are most discussed and need attention. The data for this study is sourced from reviews of the Mobile JKN application on the Google Play Store. The findings of this research are anticipated to furnish feedback and suggestions for future enhancements to the Mobile JKN application. This paper is organized as follows: section 2 describe the literature review, contains the theory that underlies this research, section 3 is method includes the method used in this research and how it was conducted, section 4 contains the results and discussion, which ends with recommendations that can be given to BPJS Kesehatan, lastly section 5 has the conclusions of this research.

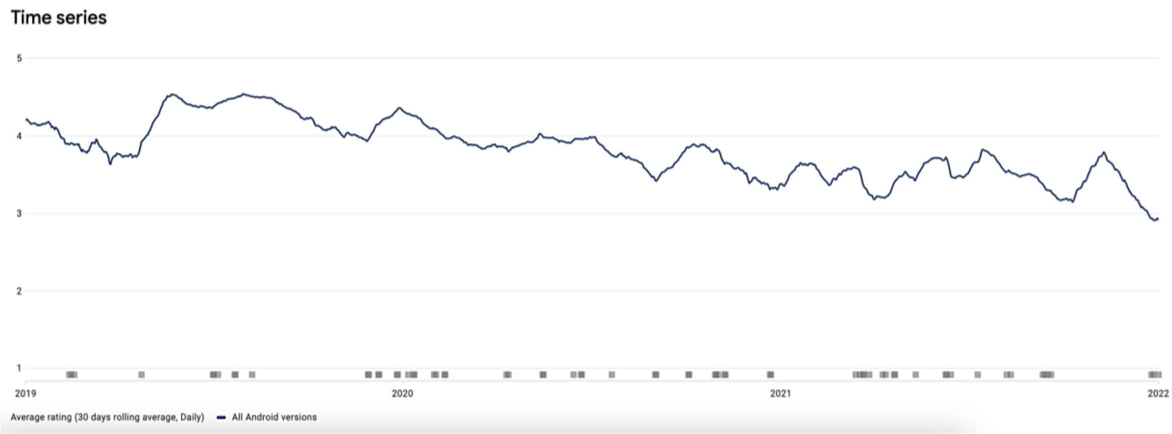

Figure 1. The trend of Mobile JKN ratings [5]

2. LITERATURE REVIEW

2.1. Naïve Bayes

The NB Classifier is a classification approach that utilizes the Bayes theorem [23]–[26], and it has gained popularity due to its straightforward and uncomplicated nature [11], [27]–[29]. The application where NB algorithms are frequently utilized is sentiment analysis [28]. The NB classifier uses probability theory to identify the highest probability of classification by examining the occurrence of each category in the training dataset [17], [30].

2.2. Support vector machine
One of the techniques employed for classification and regression prediction is SVM [11], [31]. SVM is a machine learning algorithm that utilizes the principle of structural risk minimization (SRM) to identify the optimal hyperplane that distinguishes between the two classes in the input space. In short, SVM looks for the best hyperplane that functions as a separator for two classes of data [30]–[32].

2.3. Synthetic minority oversampling technique
When data is crawling, there is usually an imbalance in the data called the imbalance class. This condition can happen because the number in one of the data classes is greater than the amount of data in another data class. The existence of an imbalanced class can reduce the quality of the classification results. One way to solve this problem is to apply the SMOTE method [17]–[19]. SMOTE can optimize the performance of the classification model algorithm when performing Sentiment analysis [19], [20]. This approach creates new minority class samples by combining neighboring models in a convex manner to balance the dataset [21], [30].

2.4. Online service quality factors
The quality of services provided to customers encompasses various factors such as the nature of the interaction, sources of information, and internet-based services, all of which are integrated with terms and conditions to enhance customer service relationships and build trust in the company [12]. Research by Yang and Fang [22] regarding the service quality dimensions was performed to broaden understanding of service quality and customer satisfaction in online securities brokerage services. The study results show the dimensions used for service quality which can be seen in Table 1.

Table 1. Service quality dimensions [22]

Category	Description
Responsiveness	Responsiveness and readiness to serve obstacles
Service reliability	Data suitability and ability to deliver promised service
Ease of use	Ease of use and related interfaces
Competence	Problem-solving ability
Access	Availability of services and media access
System reliability	System failure and errors
Timeliness	System speed and system updates
Security	System security and data security

3. METHOD
The research methodology employed in this study centers on a comprehensive analysis of the existing literature. The primary approach involves applying sentiment and factor analysis, explicitly referencing factors influencing online service quality, as visually depicted in Figure 2. The methodological journey unfolds through distinct stages: data collection, pre-processing, feature extraction, data division, data labeling, applying SMOTE, developing and evaluating the classification model, sentiment analysis, and culminating in factor analysis. The study culminates insights and recommendations derived from the outcomes of sentiment and factor analyses, thus contributing to a nuanced understanding of the subject matter.

3.1. Data collections
This study begins with data collection, achieved through the automated process of crawling data. Crawling data is getting data into documents automatically through an application [13]. Specifically, a Python library named Google Play scrapper is employed to gather data, focusing on app reviews related to the Mobile JKN application from Google Play sources. The outcome of this data crawling process manifests as a dataset conveniently stored in Excel or comma-separated values (CSV) format. This foundational step ensures the availability of organized and accessible data for subsequent analysis.

3.2. Pre-processing data
Pre-processing is an important stage in sentiment analysis [11]. The data will undergo a processing and cleaning process in pre-processing, making it easier to analyze. At this stage, a pre-processing process will be carried out. The raw data from the Mobile JKN application review will go through several processing stages, making the data lighter to process at the sentiment analysis stage. This study uses several steps in pre-processing: case folding, slang word replacement, tokenization, stop word removal, and stemming, as illustrated in Figure 3.

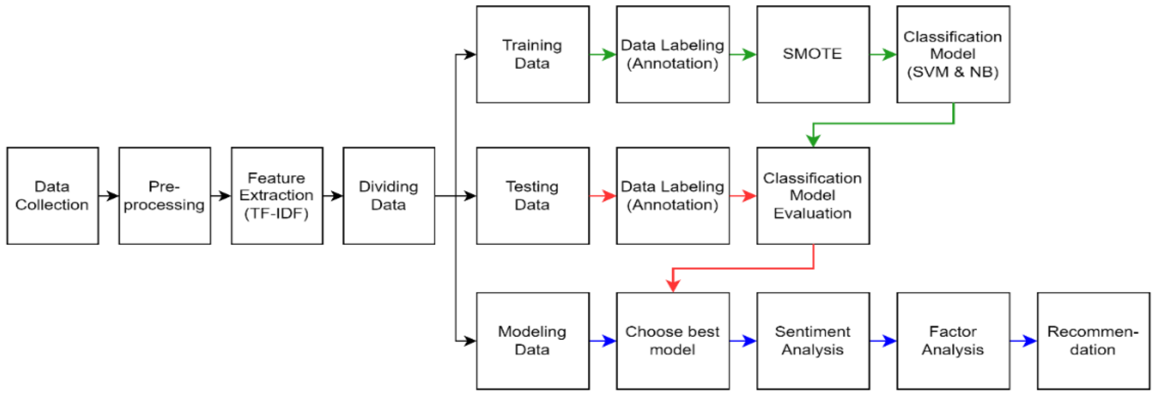

Figure 2. The model used to identify factors influencing the user experience of Mobile JKN

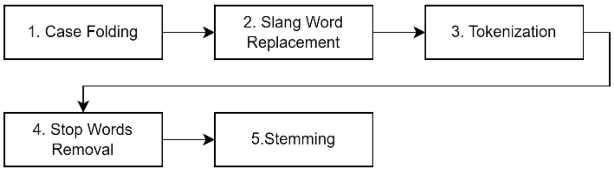

Figure 3. Data pre-processing steps

- Case folding: in sentiment analysis, if some words are the same, but some use capital letters in one of these words, then the word will be counted as more than one [11]. This redundancy can affect the analysis to be performed. Thus, it is necessary to have case folding so that there is no more redundancy in calculating the words. In this step, all data letters will be transformed into lowercase [11], [13], [20], [33]. Thus, no data use capital letters or a mixture of uppercase and lowercase letters. In addition, case folding can also increase the speed of comparing words during the indexing process [13]. The case folding process uses Python's basic syntax, the lower() function.
- Slang word replacement: slang has become a trend for people to communicate using social media. Slang saves typing time and can express pleasure or anger [34]. However, this word may not be recognized as the intended meaning in sentiment analysis, so it should be replaced with a more common word. In slang word replacement, abbreviated or non-standard words will be transformed into the appropriate word. Changing slang into common words is carried out using 15,396 Indonesian slang databases.
- Tokenization: the review sentence will be broken down into words separated by commas, then placed in an array [11]. The data that has been broken down can be called a token [13]. Tokenization is helpful in the process of calculating words that appear during the analysis process [20]. The process uses Python's basic syntax, the split() function.
- Stop words removal: words with no meaning will be deleted [20]. Usually, this process is also known as filtering. Words not matching the stopword list according to the language used will be deleted [11]. Stop-word removal is done because the words do not represent the meaning that will be used for analysis. Thus, the presence of stop word removal can also improve performance in conducting analysis [35]. At this stage, remove meaningless words in Indonesian and English using natural language toolkit (NLTK), a Python programming library.
- Stemming: words that have prefixes and suffixes are removed so that they return to their base words [15]. However, this procedure solely converts a word to its basic form and does not imply that it retains the same meaning as the original root word [20]. This process can increase the accuracy of the classification model because this process eliminates redundant words [34]. At this stage, the process uses the Python programming library, namely Sastrawi.

3.3. Feature extraction

The feature extraction process involves performing various calculations on the data, using techniques such as word weighting. One commonly used approach to word weighting is the term frequency-inverse document frequency (TF-IDF) feature. This feature can be implemented using the scikit-learn library in Python, providing a powerful tool for text analysis and natural language processing.

3.4. Dividing data

This stage will divide the data into training, tests, and modeling. Training data is used for dataset training as the learning stage of the algorithm used. Testing data is used to evaluate the performance of the algorithm. Modeling data is used for sentiment analysis, providing labels automatically through the system to the dataset.

3.5. Data labeling

This stage is the manual data labeling process by the internal team. The data will be labeled positive, negative, or neutral in the training and testing data. The sentiment labeling process will be done for the modeling data using the selected classification model algorithm.

3.6. Synthetic minority oversampling technique

This study proposes combining the SMOTE method with NB and SVM. This combination aims to address the issue of imbalanced datasets by over-sampling the minority class. This process uses an imbalanced-learn library in Python programming. By implementing this approach, the study seeks to improve the accuracy and robustness of the classification model.

3.7. Classification model

This study will compare two classification algorithms, NB and SVM. In Python, the classification model can be done using a scikit-learn library. This classification model process uses training data to train the classification model algorithm and testing data to measure the performance of the two algorithms.

3.8. Classification model evaluation

In this stage, an evaluation of the classification models of NB and SVM is carried out with the confusion matrix. This process is done with the Python programming library, namely sklearn-confusion-matrix. The classification model will be tested with several test scenarios to produce the best model. Some of the test scenarios to be evaluated are described in Table 2.

Table 2. Test scenario of model evaluation

Test	Slang replacement	SMOTE	Neutral sentiment
1	Yes	Yes	Yes
2	Yes	Yes	No
3	Yes	No	Yes
4	Yes	No	No
5	No	Yes	Yes
6	No	Yes	No
7	No	No	Yes
8	No	No	No

3.9. Sentiment analysis

During the sentiment analysis stage, the remaining dataset (data modeling) is processed and classified into different sentiments based on the model with the best performance. This model is selected based on its ability to accurately classify the sentiments in the dataset. The classification process involves analyzing the text data and identifying its emotions or opinions.

3.10. Factor analysis

At this stage, factor analysis is carried out based on the dimensions of online service quality factors. These dimensions are responsiveness, service reliability, ease of use, competence, access, system reliability, timeliness, and security [22]. The analysis is used to identify factors that influence user experience in using the Mobile JKN application. The factors that will fill each dimension are taken from the words that appear the most in the classification process. If there are words that have no meaning or do not match these factors, they will be eliminated and not included in any factor or dimension.

3.11. Recommendations

The final stage of this research is to provide recommendations based on three crucial factors. Factors that have the most negative sentiment will be given recommendations so that they can be improvements for the Mobile JKN application. So, application development can be better and improve user experience.

4. RESULTS AND DISCUSSION

4.1. Data collections

Data was taken on November 3, 2022, from the Google Play review application using the Google Play Scrapper library. The data sequence filter is "most relevant" for the last 3,000 data. The selected reviews use all-star ratings, namely 1 to 5, with details as shown in Figure 4. Collecting data using this configuration only takes 4 seconds. The method used in this study is much faster than using a questionnaire which takes seven weeks and only got 127 respondents, as was done by Handayani *et al.* [7].

4.2. Data pre-processing result

The second stage is to pre-process the data. Raw data from google play reviews will be processed with case folding, slang replacement, tokenization, stop words removal, and the steaming process. This pre-processing stage is essential to ensure the collected data is in a suitable format for further analysis. These measures aid in enhancing the accuracy and effectiveness of the subsequent analysis. The results of the pre-processing process can be seen in Table 3.

4.3. Data division and annotation for classification model

The clean data is divided into three distinct parts for analysis. The first part consists of 500 rows used as training data to train the learning system. The second part consists of 100 rows of data utilized as testing data to determine the performance of each modeling algorithm. Finally, the third part contains 2,400 rows of data that the system will automatically label using the selected modeling algorithm. The training and testing data are manually labeled by the research team, as depicted in Figure 5, to ensure accuracy and consistency in the analysis.

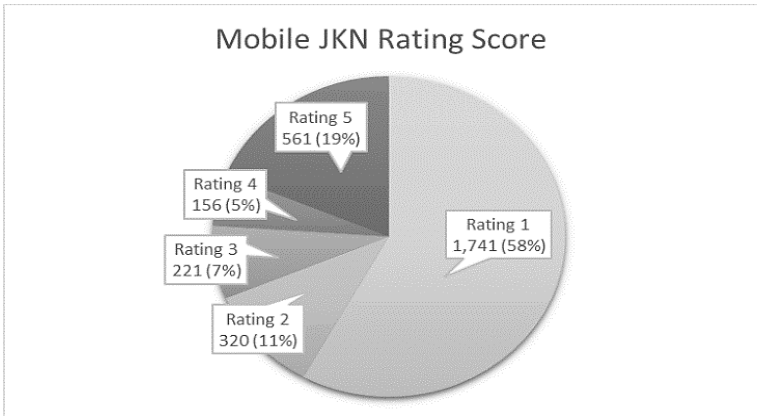

Figure 4. Distribution of rating scores from the Mobile JKN application

Table 3. Pre-processing results

Stages	Results
User reviews	The App isn't clear, there's still more regular credit, why's the notification always the num isn't active. The prev App was easier to access than the curr one which is too complicated
Case folding	the app isn't clear, there's still more regular credit, why's the notification always the num isn't active. the prev app was easier to access than the curr one which is too complicated
Slang replacement	the **application is not** clear, **there is** still more regular credit, **why is** the notification always the **number is not** active. the **previous application** was easier to access than the **current** one which is too complicated
Tokenizing	the | application | is | not | clear | there | is | still | more | regular | credit | why | is | the | notification | always | the | number | is | not | active | the | previous | application | was | easier | to | access | than | the | current | | one | which | is | too | complicated
Stop words removal	application | not | clear | still | regular | credit | notification | always | number | not | active | previous | application | easier | access | current | one | complicated
Stemming	application | not | clear | still | regular | credit | **notify** | always | number | not | active | previous | application | **easy** | access | current | one | **complicate**

4.4. Results of classification modeling evaluation

In this study, a performance comparison of the classification model algorithm was carried out against NB and SVM by combining three additional configuration variables: slang replacement, use of SMOTE algorithm, and use of neutral sentiment classification. Furthermore, this arrangement is anticipated to improve the performance score of the compared classification model algorithm. The classification model's performance was evaluated using the confusion matrix formula to determine the accuracy, precision, recall, and F1-score values. Table 4 describes the results of the tests that have been carried out.

Table 4 shows that the best classification model is when using NB and combined with using slang word replacement, SMOTE algorithm, and without using neutral data classification. The confusion matrix results using these combinations have accuracy values of 93.33%, precision of 93.76%, recall of 93.33%, and F1-score of 93.43%. A more detailed analysis related to the average performance evaluation of the classification model combined with several configuration variables can be seen in Table 5.

Table 5 shows that the average performance of the classification model in NB and SVM, when combined with the SMOTE algorithm, increases by 7.71% from the previous 78% to 86.04%. When replacing slang words with common words (slang replacement), the average performance also experienced an increase of 0.21% from 82.08% to 82.29%. When using neutral sentiment, the algorithm's performance decreased by 9.79% from 87.08% to 77.29%.

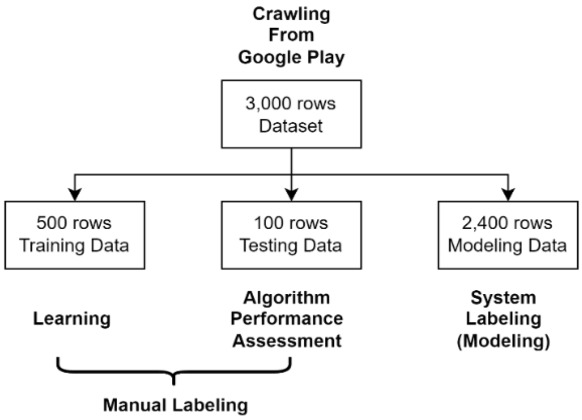

Figure 5. Data dividing and labeling process

Table 4. Confusion matrix results

Slang replacement	SMOTE	Neutral	Naive Bayes Accuracy (%)	Naive Bayes Precision (%)	Naive Bayes Recall (%)	Naive Bayes F1-score (%)	SVM Accuracy (%)	SVM Precision (%)	SVM Recall (%)	SVM F1-score (%)
True	True	True	77.50	87.15	77.50	81.60	85.83	84.91	85.83	85.01
True	True	False	93.33	93.76	93.33	93.43	93.33	93.55	93.33	93.13
True	False	True	76.67	77.77	76.67	71.72	88.33	84.66	88.33	85.83
True	False	False	81.67	85.44	81.67	78.52	93.33	93.55	93.33	93.13
False	True	True	80.00	85.42	80.00	82.48	85.83	84.91	85.83	85.01
False	True	False	93.33	93.76	93.33	93.43	93.33	93.55	93.33	93.13
False	False	True	75.00	76.88	75.00	69.28	89.17	85.31	89.17	86.71
False	False	False	80.00	84.40	80.00	76.02	92.50	92.80	92.50	92.23

Table 5. The average performance of the classification model

Classification model		Percentage	Result (%)
Slang replacement	True	82.29	UP 0.21
	False	82.08	
SMOTE	True	86.04	UP 7.71
	False	78.33	
Neutral	True	77.29	Down 9.79
	False	87.08	

4.5. Results of selected classification model

After selecting the classification model algorithm and the appropriate additional configuration variables, and after the modeling system is trained using training data, sentiment labeling can be performed automatically using the classification model system for 2,400 rows of review data. The results of this process will produce positive and negative sentiments. Negative sentiment dominates with 80.80% or 2,424 lines of data, while positive sentiment only gets 19.20% or 576 lines of Mobile JKN user review data, as seen in Figure 6.

4.6. Analysis of sentiment results

Based on 3,000 review data sets crawled from the Google Play Store for the Mobile JKN application, the words frequently appearing in all existing reviews are calculated. We took the 600 words that appear most often in the review collection and grouped them based on online service quality factors used in this study. The results of the word grouping can be seen in Table 6.

Labeling data that has been carried out sentiment analysis both manually and automatically using a classification model system, then regrouped using a service quality dictionary. From this process, we can analyze the factors influencing the experience of current Mobile JKN users, as shown in Table 7. This analysis helps to provide insights into the key factors that affect user satisfaction, allowing for more informed decision-making and improvements in service quality.

From these data, the first critical factor is related to security 2,173 reviews (72.43%) discuss this matter which consists of 188 positive reviews (8.65%) and 1,985 negative reviews (91.35%). The second factor is the ease of use, with a total of 1,935 reviews (64.50%) consisting of 426 positive reviews (22.02%) and 1,509 negative reviews (77.98%). The third factor is related to the timeliness of 1,381 reviews (46.03%) consisting of 279 positive reviews (20.20%) and 1,102 negative reviews (79.80%). In general, the factors influencing the user experience can be seen in Figure 7.

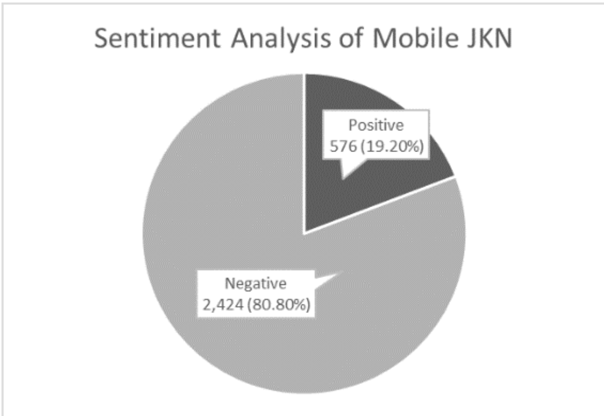

Figure 6. Sentiment analysis with the selected classification model

Table 6. Service quality dictionary

Service quality	Dictionary
Responsiveness	queue, suggestion, confirmation, report, response, responsive
Service reliability	data, appropriate, wrong, complete, clear, input, information, facilities, correct, quality, format
Ease of use	easy, hard, menu, good, complicated, difficult, feature, nice, use, bad, confused, convenient, troublesome, function, automatic, practical, appear, dizzy, understand, search, comfortable, simple, many times, tired, manual, push, annoyed, ok, tricky, satisfied, efficient, complex, done, button, effective, icon, screen
Competence	help, solution, finish, benefit, steady, success, problem, perfect
Access	access, office, online, phone, relationship, telephone, bank, offline, digital, handphone, internet, chat, call, signal, wifi, WhatsApp, up, off, public, broken
System reliability	fail, error, constraint, bug, interrupt, program, device, memory, reset, server, link, hardware, stable, developer, inhibit, exit, system, cancel, load, tough
Timeliness	update, change, direct, process, version, fast, upgrade, loading, smooth, edit, connection, late, slow
Security	register, key, enter, login, email, verification, number, account, identification, code, password, captcha, OTP, padlock, sms, password, logout, lost, registration, safe, property, lock, control, relogin, user, cache, gmail

4.7. Recommendations

After obtaining the three categories of main critical factors, further research can be carried out to get the words most frequently discussed in these three factors, as shown in Table 8. By identifying the most commonly discussed topics, researchers can gain insights into the underlying causes of user dissatisfaction and take appropriate measures to address these issues. This analysis can be used to develop targeted strategies to improve service quality and enhance customer experience.

Table 7. Sentiment analysis based on factor analysis results

Category	Positive	Negative	Sentiment
Responsiveness	20	105	125
Service reliability	187	634	821
Ease of use	426	1,509	1,935
Competence	224	237	461
Access	164	668	832
System reliability	53	481	534
Timeliness	279	1,102	1,381
Security	188	1,985	2,173

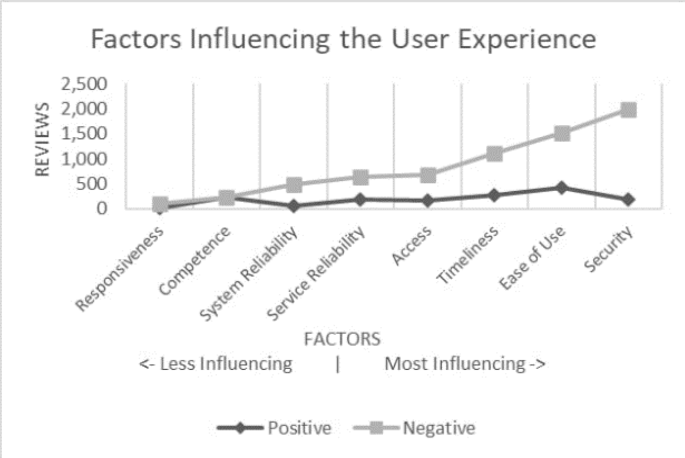

Figure 7. Factors influencing the user experience of the Mobile JKN application

Table 8. Top words on most critical factors

Security	Ease of use	Timeliness
register (806)	easy (490)	updates (736)
P:7.3% - N:92.7%	P:49.6% - N: 50.4%	P:8.2% - N:91.8%
key (557)	difficult (310)	change (247)
P:10.2% - N:89.8%	P:4.5% - N:95.5%	P:50.6% - N:49.4%
enter (475)	good (308)	process (85)
P:4.6% - N:95.4%	P:10.7% - N:89.3%	P:22.4% - N:77.6%
login (408)	menu (285)	direct (83)
P:5.1% - N:94.9%	P:10.9% - N:89.1%	P:39.8% - N:60.2%
verification (356)	hard (282)	fast (71)
P:5.3% - N:94.7%	P:8.9% - N:91.1%	P:73.2% - N:26.8%
email (302)	tricky (279)	version (63)
P:6.3% - N:93.7%	P:10.4% - N:89.6%	P:12.7% - N:87.3%
number (187)	features (232)	loading (43)
P:1.1% - N:98.9%	P:18.1% - N:81.9%	P:7.0% - N:93.0%
account (166)	nice (182)	upgrades (40)
P:0.6% - N:99.4%	P:32.4% - N:67.6%	P:22.5% - N:77.5%
identification (155)	use (142)	smooth (37)
P:1.9% - N:98.1%	P:28.9% - N:71.1%	P:37.8% - N:62.2%
code (147)	bad (67)	slow (16)
P:1.4% - N:98.6%	P:4.5% - N:95.5%	P:0.0% - N:100.0%

P: Positive, N: Negative

From these data, in the security factor, there are words such as register, key, enter, login, and verification, and most of them have negative sentiments. Hence, BPJS Kesehatan needs to provide convenience to the registration and user login process by adding additional options, such as registration or login with the Google Account API [36]. BPJS Kesehatan can also study this aspect with other companies with similar applications.

In the ease-of-use factor, there are words such as easy, difficult, good, menu, and hard. Most of these words also have negative sentiments, so BPJS Kesehatan needs further research on the user experience, features layout, and usability evaluation [37]–[39]. The process can use a prototyping design approach before doing development.

In the timeliness factor, there are words such as update, change, process, and direct, which also mostly have negative sentiments. Hence, BPJS Kesehatan needs to set the application update period so that it is not updated too often [40], [41]. In addition, his factor can be improved by using auto-scaling technology, such as microservices, to increase the speed of performance of the Mobile JKN application function [42].

5. CONCLUSION

This research aims to determine the factors that most influence the user experience of the Mobile JKN application by using a more appropriate sentiment analysis model. By comparing the performance of NB and SVM classification models, it was found that NB outperformed SVM in terms of precision and F1-score. When combined with the SMOTE algorithm, it resulted in a 7.71% increase in average accuracy. Additionally, the replacement of slang words increased accuracy by 0.21%. However, the use of neutral data decreased average accuracy by 9.79%. Based on that finding, the model used in this research combines NB with SMOTE and slang word replacement. These combinations resulted in an accuracy of 93.33%, a precision of 93.76%, a recall of 93.33%, and an F1-score of 93.43%. A further analysis based on 3,000 user reviews found that factors related to security, ease of use, and timeliness received overwhelmingly negative sentiment, with 91.35%, 77.98%, and 79.80% negative sentiment, respectively. BPJS Kesehatan needs to address these issues, as they are the most discussed and have negative connotations in the review data.

ACKNOWLEDGEMENTS

This research was carried out with the support of a fully funded scholarship program by BPJS Kesehatan, which the first author obtained for study at the University of Indonesia.

REFERENCES

[1] Government of the Republic of Indonesia, *The Constitution of the Republic of Indonesia of 1945*. Jakarta, Indonesia.

[2] President of the Republic of Indonesia, *The Republic of Indonesia Law Number 24 of 2011 Concerning Social Security Administrators*. Jakarta, Indonesia, 2011.

[3] President of the Republic of Indonesia, *Indonesian Presidential Instruction Number 1 of 2022 on Optimizing the National Health Insurance Program Implementation*. Jakarta, Indonesia, 2022.

[4] *BPJS Kesehatan Board of Directors Regulation Number 30 of 2017 Concerning the Implementation of the JKN Mobile Application*. Jakarta, Indonesia, 2017.

[5] BPJS Kesehatan, "Report on JKN Mobile Application Rating Trends (in Indonesia: Tren Rating Aplikasi Mobile JKN)," Unpublished internal company document, 2023.

[6] BPJS Kesehatan, "Report on the Target and Achievement of Registered Users of the Mobile JKN Application from 2020 to 2022 (in Indonesia: Target dan Pencapaian Pengguna Terdaftar Aplikasi Mobile JKN Tahun 2020 hingga 2022)," Unpublished internal company document, 2023.

[7] P. W. Handayani, D. A. Meigasari, A. A. Pinem, A. N. Hidayanto, and D. Ayuningtyas, "Critical success factors for mobile health implementation in Indonesia," *Heliyon*, vol. 4, no. 11, pp. 1-26, Nov. 2018, doi: 10.1016/j.heliyon.2018.e00981.

[8] A. Pratama, D. Ridwandono, T. L. M. Suryanto, E. M. Safitri, and A. Khusna, "Service Quality Analysis of M-Health Application Satisfaction And Continual Usage," *2021 IEEE 7th Information Technology International Seminar (ITIS)*, 2021, pp. 1-5, doi: 10.1109/ITIS53497.2021.9791639.

[9] P. Lak and O. Turetken, "Star ratings versus sentiment analysis - A comparison of explicit and implicit measures of opinions," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2014, pp. 796–805, doi: 10.1109/HICSS.2014.106.

[10] D. P. Javale and S. S. Desai, "Machine learning ensemble approach for healthcare data analytics," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 28, no. 2, pp. 926–933, 2022, doi: 10.11591/ijeecs.v28.i2.pp926-933.

[11] I. Firmansyah, M. H. Asnawi, S. A. Hasanah, R. Novian, and A. A. Pravitasari, "A comparison of support vector machine and Naïve Bayes classifier in binary sentiment reviews for PeduliLindungi application," in *2021 International Conference on Artificial Intelligence and Big Data Analytics, ICAIBDA 2021*, 2021, pp. 140–145, doi: 10.1109/ICAIBDA53487.2021.9689771.

[12] C. A. Haryani, A. N. Hidayanto, N. F. A. Budi, and Herkules, "Sentiment analysis of online auction service quality on Twitter Data: A case of E-Bay," *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, pp. 1–5, 2019, doi: 10.1109/CITSM.2018.8674365.

[13] S. Amrie, S. Kurniawan, J. H. Windiatmaja, and Y. Ruldeviyani, "Analysis of Google Play Store's sentiment review on Indonesia's P2P fintech platform," in *2022 IEEE Delhi Section Conference (DELCON)*, 2022, pp. 1-5, doi: 10.1109/DELCON54057.2022.9753108.

[14] M. B. Ressan and R. F. Hassan, "Naïve-Bayes family for sentiment analysis during COVID-19 pandemic and classification tweets," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 28, no. 1, pp. 375–383, 2022, doi: 10.11591/ijeecs.v28.i1.pp375-383.

[15] N. G. Ramadhan and F. D. Adhinata, "Sentiment analysis on vaccine COVID-19 using word count and Gaussian Naïve Bayes," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 3, pp. 1765–1772, 2022, doi: 10.11591/ijeecs.v26.i3.pp1765-1772.

[16] O. Chamorro-Atalaya *et al.*, "Supervised learning using support vector machine applied to sentiment analysis of teacher performance satisfaction," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 28, no. 1, pp. 516–524, 2022, doi: 10.11591/ijeecs.v28.i1.pp516-524.

[17] N. Santoso, W. Wibowo, and H. Himawati, "Integration of synthetic minority oversampling technique for imbalanced class," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, no. 1, pp. 102–108, 2019, doi:

10.11591/ijeecs.v13.i1.pp102-108.
[18] A. K. Hamoud *et al.*, "A prediction model based machine learning algorithms with feature selection approaches over imbalanced dataset," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 28, no. 2, pp. 1105–1116, 2022, doi: 10.11591/ijeecs.v28.i2.pp1105-1116.
[19] A. A. Alani and G. Cosma, "ArSL-CNN: A convolutional neural network for arabic sign language gesture recognition," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, pp. 488–499, 2020, doi: 10.11591/ijeecs.v22.i2.pp488-499.
[20] M. R. Pribadi, D. Manongga, H. D. Purnomo, I. Setyawan, and Hendry, "Sentiment analysis of the PeduliLindungi on Google Play using the random forest algorithm with SMOTE," in *2022 International Seminar on Intelligent Technology and Its Applications: Advanced Innovations of Electrical Systems for Humanity, ISITIA 2022 - Proceeding*, Aug. 2022, pp. 115–119, doi: 10.1109/ISITIA56226.2022.9855372.
[21] A. Muneer, R. F. Ali, A. Alghamdi, S. M. Taib, A. Almaghthawi, and E. A. Abdullah Ghaleb, "Predicting customers churning in banking industry: A machine learning approach," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 539–549, 2022, doi: 10.11591/ijeecs.v26.i1.pp539-549.
[22] Z. Yang and X. Fang, "Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services," *International Journal of Service Industry Management*, vol. 15, no. 3, pp. 302–326, 2004, doi: 10.1108/09564230410540953.
[23] N. M. Samsudin, C. F. B. Mohd Foozy, N. Alias, P. Shamala, N. F. Othman, and W. I. S. Wan Din, "Youtube spam detection framework using naïve bayes and logistic regression," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 3, pp. 1508–1517, 2019, doi: 10.11591/ijeecs.v14.i3.pp1508-1517.
[24] R. A. I. Alhayali, M. A. Ahmed, Y. M. Mohialden, and A. H. Ali, "Efficient method for breast cancer classification based on ensemble hoffeding tree and Naïve Bayes," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 2, pp. 1074–1080, 2020, doi: 10.11591/ijeecs.v18.i2.pp1074-1080.
[25] M. H. Jopri, M. R. Ab Ghani, A. R. Abdullah, T. Sutikno, M. Manap, and J. Too, "Naïve bayes and linear discriminate analysis based diagnostic analytic of harmonic source identification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 3, pp. 1626–1633, 2020, doi: 10.11591/ijeecs.v20.i3.pp1626-1633.
[26] L. M. Padirayon, M. S. Atayan, J. S. Panelo, and C. R. Fagela, "Mining the crime data using naïve Bayes model," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, pp. 1084–1092, 2021, doi: 10.11591/ijeecs.v23.i2.pp1084-1092.
[27] K. A. R. Torres and J. R. Asor, "Machine learning approach on road accidents analysis in Calabarzon, Philippines: An input to road safety management," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 2, pp. 993–1000, 2021, doi: 10.11591/ijeecs.v24.i2.pp993-1000.
[28] M. S. Arman, K. Sarker, A. K. Shakir, S. F. Hossain, and A. Hasan, "Medicine prediction based on doctor's degree: a data mining approach," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 2, pp. 1125–1134, 2022, doi: 10.11591/ijeecs.v26.i2.pp1125-1134.
[29] O. Somantri and D. Apriliani, "Opinion mining on culinary food customer satisfaction using naïve bayes based-on hybrid feature selection," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 1, pp. 468–475, 2019, doi: 10.11591/ijeecs.v15.i1.pp468.
[30] Hermanto, A. Y. Kuntoro, T. Asra, E. B. Pratama, L. Effendi, and R. Ocanitra, "Gojek and Grab user sentiment analysis on Google Play using Naive Bayes algorithm and support vector machine based smote technique," in *Journal of Physics: Conference Series*, Nov. 2020, vol. 1641, no. 1, doi: 10.1088/1742-6596/1641/1/012102.
[31] A. Khaleel Faieq and M. M. Mijwil, "Prediction of heart diseases utilising support vector machine and artificial neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 374–380, 2022, doi: 10.11591/ijeecs.v26.i1.pp374-380.
[32] C. Annubaha, A. P. Widodo, and K. Adi, "Implementation of eigenface method and support vector machine for face recognition absence information system," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 3, pp. 1624–1633, 2022, doi: 10.11591/ijeecs.v26.i3.pp1624-1633.
[33] Mussalimun, E. H. Khasby, G. I. Dzikrillah, and Muljono, "Comparison of K-N earest Neighbor (K-NN) and Naïve Bayes Algorithm for Sentiment Analysis on Google Play Store Textual Reviews," in *2021 8th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2021*, 2021, pp. 180–184, doi: 10.1109/ICITACEE53184.2021.9617217.
[34] S. F. Sayeedunnisa, "Using slang and emoticon for sentiment analysis of social media data," *International Journal of Engineering Research & Technology*, vol. 8, no. 15, pp. 69–72, Sep. 2020, doi: 10.17577/IJERTCONV8IS15016.
[35] Q. A. Xu, V. Chang, and C. Jayne, "A systematic review of social media-based sentiment analysis: Emerging trends and challenges," *Decision Analytics Journal*, vol. 3, pp. 1-16, Jun. 2022, doi: 10.1016/j.dajour.2022.100073.
[36] K. Kim, J. Ryu, H. Lee, Y. Lee, and D. Won, "Distributed and federated authentication schemes based on updatable smart contracts," *Electronics*, vol. 12, no. 5, 2023, doi: 10.3390/electronics12051217.
[37] A. David *et al.*, "Feasibility, ease-of-use, and operational characteristics of world health organization–recommended moderate-complexity automated nucleic acid amplification tests for the detection of tuberculosis and resistance to rifampicin and isoniazid," *The Journal of Molecular Diagnostics*, vol. 25, no. 1, pp. 46–56, 2023, doi: 10.1016/j.jmoldx.2022.10.001.
[38] X. Li, X. Zhao, W. (Ato) Xu, and W. Pu, "Measuring ease of use of mobile applications in e-commerce retailing from the perspective of consumer online shopping behaviour patterns," *Journal of Retailing and Consumer Services*, vol. 55, pp. 1-12, 2020, doi: 10.1016/j.jretconser.2020.102093.
[39] V. Hannus and J. Sauer, "Understanding farmers' intention to use a sustainability standard: The role of economic rewards, knowledge, and ease of use," *Sustainability*, vol. 13, no. 19, 2021, doi: 10.3390/su131910788.
[40] A. A. Al-Subaihin, F. Sarro, S. Black, L. Capra, and M. Harman, "App Store effects on software engineering practices," *IEEE Transactions on Software Engineering*, vol. 47, no. 2, pp. 300–319, 2021, doi: 10.1109/TSE.2019.2891715.
[41] J. Dai, D. Li, B. Gao, L. Zhou, and W. Chen, "Research on the release management technology of the mobile application for electricity dispatching," *Journal of Physics: Conference Series*, vol. 1971, no. 1, pp. 1–7, 2021, doi: 10.1088/1742-6596/1971/1/012055.
[42] F. Tapia, M. ángel Mora, W. Fuertes, H. Aules, E. Flores, and T. Toulkeridis, "From monolithic systems to microservices: A comparative study of performance," *Applied Sciences*, vol. 10, no. 17, 2020, doi: 10.3390/app10175797.

BIOGRAPHIES OF AUTHORS

Muhammad Yazid Al Qahar received his S.Kom. degree in Informatic Engineering from Indonesia Computer University, Bandung, in 2014. He is pursuing his master's degree in the Information Technology program at the University of Indonesia and works as an IT system specialist at BPJS Kesehatan. His research interests include data mining, software development, databases, project management, and mobile applications. He can be contacted at email: muhammad.yazid22@ui.ac.id.

Yova Ruldeviyani is a lecturer and researcher at the Faculty of Computer Science, University of Indonesia. She obtained her S.Kom. degree in Computer Science from the University of Indonesia in 1996, and her M.Kom. degree in Computer Science from the University of Indonesia in 2001. Her research interests include databases, business intelligence, enterprise resource planning, e-government, and software development. She can be contacted at email: yova@cs.ui.ac.id.

Ulfah Nur Mukharomah received her S.T. degree in Informatic Engineering from Widyatama University, Bandung, in 2017. She is pursuing her master's degree in the Information Technology program at the University of Indonesia and works as an IT system specialist at BPJS Kesehatan. Her research interests include software development and IT governance. She can be contacted at email: ulfah.nur21@ui.ac.id.

Miftahul Agtamas Fidyawan received his S.S.T. degree in Informatic Engineering from the State Polytechnic of Malang, Malang, in 2017. He is pursuing his master's degree in the Information Technology program at the University of Indonesia. He is a first-level computer expert at the Regional Development Planning Agency of Sumenep Regency. His research interests include software development, e-government, and data science. He can be contacted at email: miftahul.agtamas@ui.ac.id.

Ramadhoni Putra received his S.S.T. degree in Informatic Engineering from Sultan Syarif Kasim State Islamic University, Riau, in 2013. He is currently pursuing his master's degree in the Information Technology program at the University of Indonesia. He is a first-level computer expert at the Communication and Informatics Office of Rohul Regency. His research interests include infrastructure, e-government, IT networks, and software development. He can be contacted via email at ramadhoni.putra@ui.ac.id.