International Seminar and Workshop on Public Health Action
https://swopha.dinus.ac.id

Sentiment Analysis of Satu Sehat Application Usage Reviews on Google Play Store

Ayulia Putri[1], Mieke Nurmalasari[2]*

[1, 2] Health Information Management, Faculty of Health Sciences, Esa Unggul University, West Jakarta, Indonesia

Article Info
Article History
Submitted:
Revised:
Accepted:
Keywords:
Satu Sehat; Sentiment Analysis; Google Play Store

Abstract

The Satu Sehat application is a digital platform that supports access to health information and services in Indonesia. This research analyzes the sentiment of user reviews on the Google Play Store to evaluate perceptions and obstacles faced. A total of 10,000 reviews were collected using Google-Play-Scraper and analyzed with a lexicon-based approach using TextBlob in Python. The results showed 56.9% of positive reviews were related to ease of use and access to health data, while 43.1% of negative reviews were due to technical constraints, such as login difficulties and app bugs. These findings highlight the need for technical improvements to enhance system stability and user satisfaction.

Correspondence Address:
Jl. Arjuna Utara No.9, Duri Kepa, Kb. Jeruk, Jakarta Barat, Indonesia
E-mail:
mieke@esaunggul.ac.id

eISSN 3063-2439

Introduction

Advances in information technology have brought significant changes to various aspects of life, including the healthcare sector. One emerging innovation is mobile-based health applications that enable users to access health information and medical services more easily. The Satu Sehat application represents one such innovation designed to facilitate access to health information and services for the Indonesian population. Through this application, users can obtain health information, conduct consultations with medical professionals, and access various other healthcare services.

Satu Sehat is a public health application provided by the Ministry of Health that offers diverse health features and services with data connected through a single platform called Indonesia Health Service (HIS), commonly known as Satu Sehat (Lubis & Sudradjat, 2024). The Satu Sehat application, previously known as Peduli Lindungi, was first utilized during the Covid-19 pandemic in 2019 as a mandatory application for citizens to verify Covid-19 vaccination certificates. In March 2023, the application was renamed to Satu Sehat. This name change was accompanied by the addition of new features, such as a medical resume feature that enables detailed and well-structured health history monitoring (Nugraha et al., 2024).

The success of an application is determined not only by the features it offers but also by user perceptions and experiences. User reviews on application distribution platforms, such as Google Play Store, can provide valuable insights regarding satisfaction levels, encountered problems, and suggestions for further improvements. Reviews provided by Satu Sehat application users on Google Play Store often consist of positive suggestions and negative complaints. However, sorting these reviews can be challenging due to the large volume of reviews available on Google Play Store. Performing this task manually would require considerable time. Therefore, a method is needed that can automatically and rapidly separate all reviews into positive and negative categories. Thus, the aim of this study was to identify light-sensitive injectables in one of Singapore's tertiary restructured hospitals, determine the reasons for and impact of light pro- tection, and suggest appropriate light protective measures.

Previous research related to sentiment analysis of Satu Sehat app reviews conducted by Nugraha, M. T. et al (2023) aimed to obtain knowledge-based information patterns that could help app developers understand user needs and improvements that could be made through the Knowledge Discovery in Databases (KDD) method. The results showed negative sentiment, as many users complained about the Satu Sehat app having numerous bugs and OTP codes that frequently failed to be sent to users when logging into the Satu Sehat app (Nugraha et al., 2024).

Therefore, this research project aims to analyze the sentiment of Satu Sehat app user reviews using user review data on Google Play Store with the Python programming language to evaluate user satisfaction levels and identify problems encountered. It is hoped that the results of this study can provide valuable insights for app developers to improve the quality and performance of the Satu Sehat app.

Methods

The data used in this study consists of primary data regarding user reviews collected from Google Play Store. Data scraping was performed using the google-play-scraper package, with a total of 10,000 data points collected. Google-Play-Scraper is a scraping tool for extracting data from Google Play Store without external dependencies using Python programming language. In this study, the data collected includes reviews and ratings from the Satu Sehat application.

The data analysis employed is sentiment analysis with a lexicon-based approach. Lexicon-based is a feature or collection of words that have been assigned values based on a dictionary or lexicon to detect sentiment strength in text, with the assumption that a single sentence of text can contain two sentiments simultaneously (Nurmalasari & Temesvari, 2020).

Google-Play-Scraper performs data scraping by extracting the application ID found in the URL, such as the Satu Sehat application ID 'com.telkom.tracencare'. After data scraping, the process continues to the data preprocessing stage, which consists of cleaning, case-folding, tokenizing, filtering, and stemming. The purpose of this preprocessing stage is to clean the data from noise or remove irrelevant data or information that would interfere with model accuracy.

Subsequently, the noise-free data is labeled and scored, where the positive category is assigned a score > 0, and a score ≤ 0 for the negative category. The results of the sentiment analysis are visualized in the form of pie charts and word clouds.

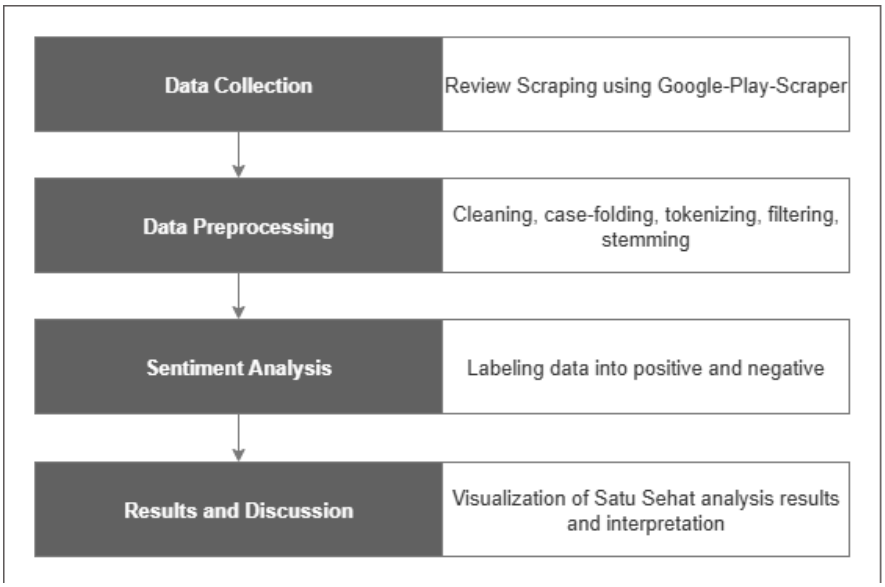

Figure 1. Sentiment analysis model schema.

The Google-Play-Scraper API is a method for extracting data from Google Play Store without external dependencies using Python programming language. The data that can be extracted includes application information such as application title, developer URL, application category, overall rating and reviews, description, thumbnail, content rating, and application screenshots. In addition to application information, the Google-Play-Scraper API can also extract user review data such as name, photo, rating, date, comment likes, and comments (Larasati et al., 2022).

Python is a high-level programming language developed by Guido Van Rossum and first released in 1991. In recent years, Python has become extremely popular among researchers and practitioners. Python is a versatile programming language that can be used for various applications, including Machine Learning and Deep Learning (Alfarizi et al., 2023). In Python code development, various Integrated Development Environments (IDEs) can be used, such as Visual Studio Code, Sublime Text, and PyCharm. Additionally, there are also web-based IDEs that support collaboration and cloud-based processing, such as Jupyter Notebook and Google Colab, which increasingly expand accessibility and convenience in the development and deployment of Python-based applications (Alfarizi et al., 2023).

Word Cloud is a visual representation of text that highlights the frequency or importance of words in a collection of text data. Words that appear more frequently in the text will be displayed larger or more prominently in the word cloud (Julianto, 2022).

In this study, data was collected from user reviews of the Satu Sehat application on Google Play Store through scraping using google-play-scraper in Google Colab within the time frame from January 1, 2020, to June 30, 2024. The total amount of data collected was 10,000 data points and saved in CSV format. The following are the steps for scraping user reviews of the Satu Sehat application on Google Play Store: (1) Install the Google-Play-Scraper API by running 'pip install google-play-scraper'; (2) Import packages 'from google_play_scraper import app', 'import pandas as pd', 'import numpy as np'; (3) Extract the application ID found in the URL such as 'com.telkom.tracencare'; (4) Perform review scraping; (5) Insert the scraping results into a data frame.

Data preprocessing is the systematic technique of reshaping raw, disparate data into a clean, coherent dataset by implementing data purification, normalization, and integration methods (van Atteveldt et al., 2021). Data preprocessing is a stage performed to ensure that data containing reviews used for sentiment analysis facilitates modeling (Giovani et al., 2020). Several stages involved in data preprocessing are as follows:

Cleaning

Incomplete, incorrect, incomplete, inaccurate, or irrelevant data is identified during the data cleansing process. This type of dirty data can be replaced, modified, or deleted using special techniques. The results of data that has undergone the cleaning process include the removal of irrelevant attributes in the text, such as numbers, symbols, URLs, hashtags, emojis, punctuation marks, and excessive spaces. This process aims to obtain cleaner and easier-to-analyze data (Hadiprakoso et al., 2023).

Case-Folding

The data results that have undergone the case folding stage are data where all capital letters (upper case) present in the text, including at the beginning, middle, and end of sentences, have been converted to lowercase letters (lower case) (Sriani et al., 2024).

Tokenizing

The data results that have undergone the tokenizing stage are data that have been processed by separating sentence text into individual words based on their constituent components. The tokenizing results yield review data as a list of individually separated words or word-by-word from the originally complete sentence text (Wati & Ernawati, 2021).

Filtering

The data results that have undergone the filtering stage are data that have undergone the removal of words that are not beneficial and hold no value in the research process. The filtering results are obtained using the remove stopwords method in the NLTK Library (Joshi & Patel, 2021).

Stemming

The data results from stemming involve reducing the frequency of occurrence of derivative words by removing affixes found at the beginning, end, or both in a word (Fan et al., 2021).

Sentiment analysis is a process of analyzing and processing textual data related to opinions about products, applications, cases, services, individuals, or organizations to generate global information within a data collection. This includes affective states (the writer's feelings when writing) and emotional expressions (the expressions the writer wants to communicate to readers) (Rahayu et al., 2022). Sentiment analysis represents an NLP methodology employed to identify and categorize the affective orientation of text as either positive, negative, or neutral (Aftab et al., 2023).

In the sentiment analysis of Satu Sehat application usage reviews on Google Play Store, this study uses 'TextBlob', which is a Natural Language Processing library in Python that facilitates text manipulation and language data processing. The following are the steps for conducting sentiment analysis: (1) Install 'TextBlob' by running '!pip install textblob'; (2) Subsequently, label the data/reviews into 2 categories: 'score > 0' meaning 'positive' and 'score ≤ 0' meaning 'negative'; (3) After the data is completed and successfully labeled, visualize the data using pie charts and word clouds.

Results

From the 10,000 scraped data points, data preprocessing was subsequently performed, followed by sentiment analysis categorized into 2 categories: positive and negative sentiment. Figure 2 shows that from 10,000 reviews, 56.9% of reviews were positive and 43.1% were negative. The highest percentage was in positive sentiment, indicating that more than half of users have positive views or responses toward the Satu Sehat application. However, the percentage of negative sentiment is also quite significant at over 40%, indicating the presence of problems or deficiencies in certain areas or aspects of the Satu Sehat application. Therefore, it is necessary to identify and fix problems or bugs that are frequently reported so

that the percentage of negative sentiment toward the Satu Sehat application can be reduced and positive sentiment can increase further.

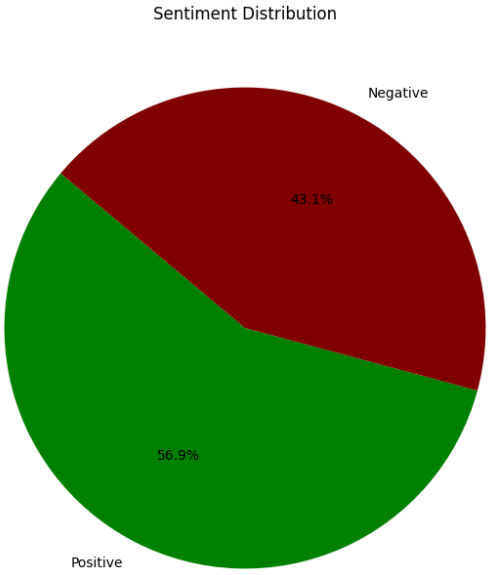

Figure 2. Percentage of positive and negative sentiment reviews

The following is the result of the sentiment analysis visualization of Satu Sehat reviews in the form of a word cloud.

Figure 3. Word cloud positive and negative sentiment

Figure 3 shows the word cloud for positive sentiment and negative sentiment. The words or topics that are most dominant or frequently appear in positive reviews regarding the Satu Sehat application are "bagus" (good), "membantu" (helpful), "memudahkan" (simplifies), "bermanfaat" (beneficial), and "menarik" (interesting), which illustrate that the Satu Sehat application is effective and flexible to use in healthcare services. The words or topics that are most dominant or frequently appear in negative reviews regarding the Satu Sehat application are "buruk" (bad), "susah" (difficult), "dipersulit" (made difficult), "gagal" (failed), and "jelek" (poor), which indicate that the Satu Sehat application frequently experiences bugs or errors that cause users or the public to have difficulty accessing information available in the Satu Sehat application.

Table 1. Positive sentiment and negative sentiment

Type of Sentiment	Examples of Several Reviews
Positive Sentiment	"tampilan baru, sistem lebih lancar dan terlihat lebih fresh"
	"tampilannya lebih friendly dan informatif"
	"ini review kedua untuk app ini versi terbaru ini jauh lebih baik dari versi sebelumnya. lengkap dan mudah digunakan"
	"kami sangat terbantu menggunakan apk ini karena mudah mengakses kartu vaksin terima kasih"
	"sebuah aplikasi yang sangat membantu untuk eksistensi kewarganegaraan dengan kesehatan bebas covid19"
	"sangat membantu aplikasi ini karena tidak ribet membawa kertas kemana mana dan sangat efisien"
Negative Sentiment	"ngga bisa login email tidak terdaftar pas mau bikin email akun sudah ada, padahal akses loginnya cuma email sama nomber hp nomber hp yg lama sekarang uda ganti padahal butuh banget surat vaksin"
	"aplikasi satu sehat ini booster saya tidak ada sertifikat nya padahal saya sudah melakukan vaksin sampai booster tapi di apk ini sertifikat yg ada hanya sampai vaksin tiga atau booster pertama saja sudah melakukan pengaduan pun masih belum ada muncul data dan sertifikat nya"
	"aplikasinya gak bisa masuk. masukin email gak ada no verifikasi masuk ke email masukin pake no hp malah gak masuk-masuk. ini gimana sih aplikasinya dari pada sering error kaya gini mah minding balik kaya yang dulu aya model aplikasinya ini setelah gnti nama jadi satu sehat malah banyak masalah"
	"tolong secepatnya diperbaiki ini kenapa pas login akun tidak ditemukan, pas daftar juga malah akun sudah ada udah minta bantuan email malah bantuan sedang bemasalah aneh hampir semua aplikasi sekelas pemerintah gini amat gak ada pemeliharaan berkala"
	"coba login masuk email bisa tapi nik ga sesuai mau di ganti jadi susah udh coba beberapa kali masi susah ganti nik jadi susah buat verifikasi email"
	"sumpah bad banget tolong dong perbaiki sistemnya masa mau login aja susah nya ampun usah masuk email dan nomor hp malah di bilangnya akun tidak terdaftar giliran daftar di bilangnya akun sdh terdaftar"

Table 1 shows examples of reviews that fall into the positive and negative sentiment categories. In terms of positive sentiment, it can be seen that the public appreciates the Satu Sehat app for its ease of access to detailed and well-structured health history data. The new interface and services offered are also considered attractive and informative. On the other hand, negative sentiment indicates that the public is experiencing issues with the Satu Sehat app, such as login difficulties, server errors, app updates, and frequently unregistered user accounts, as well as problems with the issuance of COVID-19 vaccination cards.

Discussion

The results of the study show a sentiment distribution of 56.9% positive and 43.1% negative from 10,000 reviews of the Satu Sehat app, reflecting the condition of digital health apps with great potential but still facing implementation challenges. The dominance of positive sentiment indicates that the concept of the Satu Sehat app as a centralized digital health platform has been well received by the Indonesian public. This is due to the ease of access to integrated health services, official support from the Ministry of Health which enhances credibility, and the public's need for accelerated digital health transformation in the wake of the COVID-19 pandemic. However, the proportion of negative sentiment, which reached 43.1%, indicates serious problems in the technical implementation of the application. This negative sentiment is most likely caused by the instability of the application, with frequent bugs and crashes, slow performance, a user interface that is not very intuitive, especially for senior users, and system integration issues that result in data inconsistencies with health facilities. Limited integration coverage, complicated registration processes, and digital infrastructure constraints in some areas also contribute to a less than optimal user experience. Despite these challenges, the results show positive potential, with the Indonesian public ready to adopt digital health technology but expecting more mature implementation. Significant negative sentiment can be used as valuable feedback for continuous improvement. Strategies that need to be prioritized include optimizing technical stability, simplifying user experience, expanding healthcare

facility integration, and improving outreach programs. With proper handling of the issues causing negative sentiment, the Satu Sehat application has the potential to achieve higher user satisfaction and become the foundation for the success of Indonesia's digital health transformation.

Conclusion

Based on the sentiment analysis results of 10,000 Satu Sehat application reviews obtained from Google Play Store from January 1, 2020, to June 30, 2024, it can be concluded that the sentiment distribution shows 56.9% of reviews are positive and 43.1% are negative. The positive sentiment analysis results for the Satu Sehat application demonstrate public appreciation for the ease of access to well-structured health history data, the attractive new interface of the Satu Sehat application, and informative services. Meanwhile, the negative sentiment analysis results for the Satu Sehat application show user complaints regarding difficulties in using the application, including login problems, server errors, problematic application updates, and unregistered accounts. Additionally, issues with Covid-19 vaccination card issuance are also frequently reported, along with unresponsive servers. Users experience problems when logging in or registering because OTP codes are not sent to users. Based on the sentiment analysis of Satu Sehat application usage, positive sentiment was obtained, as evidenced by the many users who feel helped by the presence of the Satu Sehat application in their daily lives.

Several suggestions for Satu Sehat application developers are that although the Satu Sehat application has positive aspects appreciated by some users, there are a significant number of complaints and dissatisfaction that need to be addressed. Focus on technical improvements and enhancing application stability should be the main priority to increase user satisfaction and reduce negative sentiment. Therefore, it is hoped that the Satu Sehat application can be more effective and beneficial in providing healthcare services to the community.

Author Contributions

Conceptualization (AP, MN), methodology (MN), data analysis (AP), writing and editing the manuscript (AP, MN), interpreting the result (AP, MN). The final version of the manuscript was approved by all authors.

References

Aftab, F., Bazai, S. U., Marjan, S., Baloch, L., Aslam, S., Amphawan, A., & Neo, T. K. (2023). A Comprehensive Survey on Sentiment Analysis Techniques. International Journal of Technology, 14(6), 1288–1298. https://doi.org/10.14716/ijtech.v14i6.6632

Alfarizi, M. R. sirfatullah, Al-farish, M. Z., Muhamad, T., Ardiansah, G., & Elgar, M. (2023). Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning. Karya Ilmiah Mahasiswa Bertauhid (KARIMAH TAUHID), 2(1), 1–6.

Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data. Frontiers in Energy Research, 9, 1–17. https://doi.org/10.3389/fenrg.2021.652801

Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. Jurnal Teknoinfo, 14(2), 115.

Hadiprakoso, R. B., Setiawan, H., Yasa, R. N., & Girinoto. (2023). Text Preprocessing for Optimal Accuracy in Indonesian Sentiment Analysis Using a Deep Learning Model with Word Embedding. AIP Conference Proceedings, 2680(1). https://doi.org/10.1063/5.0126116

Joshi, A. P., & Patel, B. V. (2021). Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process. Oriental Journal of Computer Science and Technology, 13, 78–81. https://doi.org/10.13005/ojcst13.0203.03

Julianto, I. T. (2022). Analisis Sentimen Terhadap Sistem Informasi Akademik Institut Teknologi Garut. Jurnal Algoritma, 19(1), 449–456. https://doi.org/10.33364/algoritma/v.19-1.1112

Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 6(9), 4305–4313.

Lubis, B. O., & Sudradjat, A. (2024). Evaluasi Aplikasi Satusehat Dengan Metode Use Questionnare dan IPALubis, Baginda Oloan, & Sudradjat, Adjat. (2024). Evaluasi Aplikasi Satusehat Dengan Metode Use Questionnare dan IPA. 8(3), 2988–2995. 8(3), 2988–2995.

Nugraha, M. T., Nina, S., & Enri, U. (2024). Analisis Sentimen Ulasan Aplikasi Satu Sehat Pada Google Play Store Menggunakan Naïve Bayes Classifier. JATI (Jurnal Mahasiswa Teknik Informatika), 7(5), 3593–3601.

Nurmalasari, M., & Temesvari, N. A. (2020). Analisis Sentimen terhadap Opini Masyarakat dalam Penggunaan Mobile-JKN untuk Pelayanan BPJS Kesehatan Tahun 2019. Indonesian of Health ..., 8(1), 35–44. https://inohim.esaunggul.ac.id/index.php/INO/article/view/208

Rahayu, S., MZ, Y., Bororing, J. E., & Hadiyat, R. (2022). Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP. Edumatic: Jurnal Pendidikan Informatika, 6(1), 98–106.

Sriani, Suhardi, & Saufa Yardha, L. (2024). Analisis Sentimen Pengguna Aplikasi Mobile Jkn Menggunakan Algoritma Naïve Bayes. Journal of Science and Social Research, 4307(2), 555–563.

van Atteveldt, W., van der Velden, M. A. C. G., & Boukes, M. (2021). The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms. Communication Methods and Measures, 15(2), 121–140. https://doi.org/10.1080/19312458.2020.1869198

Wati, R., & Ernawati, S. (2021). Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python. Jurnal Teknik Informatika UNIKA Santo Thomas, 06, 240–247.