

Sentiment Analysis Naive Bayes Method on SatuSehat Application

Shahmirul Hafizullah Imanuddin^{1*}, Kusworo Adi², Rahmat Gernowo²

¹Information Systems Masters Study Program, Graduate School Faculty, Universitas Diponegoro, Semarang, Indonesia

²Physics Study Program, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

Received: May 29, 2023

Revised: June 15, 2023

Accepted: July 25, 2023

Published: July 31, 2023

Corresponding Author:

Hafizullah

mirulhafiz@students.undip.ac.id

DOI: [10.29303/jppipa.v9i7.4054](https://doi.org/10.29303/jppipa.v9i7.4054)

© 2023 The Authors. This open access article is distributed under a (CC-BY License)



Abstract: The SatuSehat application is an application that provides health services to users. This application is a development of the PeduliLindungi application which is used to handle vaccination history in the new normal era. Therefore, it is important to classify user reviews into positive and negative sentiments using the Naïve Bayes method. The use of this method because it can produce a model that is quite accurate and effective. The results of data collection in this study were 25,000 of which 18,359 were negative and 6,641 were positive. The results of the Naïve Bayes accuracy test are 97% with negative sentiment results, namely precision has a value of 98%, recall has a value of 98% and f1-score has a value of 98%, while positive sentiment results, namely precision has a value of 94%, recall has a value of 94% and f1-score has a value of 94%. This study aims to classify user reviews of the SatuSehat application into positive and negative sentiments and assess the performance of the Naïve Bayes method regarding public opinion on the use of the SatuSehat application based on reviews from the Google Playstore application.

Keywords: Google Playstore; Naïve Bayes; SatuSehat

Introduction

Application is software made by a computer company to do certain tasks, for example Microsoft Word, Microsoft Excel. The application comes from the word application which means the application of application usage. (Saarikko et al., 2020) an application is an application, storing something, data, problems, work into a facility or media that can be used to apply or implement existing things or problems so that they turn into a new form without losing values. the basis of data matters, problems, and the work itself. so the application is a transformation of a problem or work in the form of things that are difficult to understand to be simpler, easier and understandable for users. So that with the application, a problem will be helped more quickly and precisely (Salvagno et al., 2023). Applications have many types, among them are desktop applications that operate offline and web applications that operate online. A web application is an application that uses browser

technology to run applications and is accessed via a computer network, (Fauzan et al., 2022).

Meanwhile, Alsaffar et al. (2022) a web application is a program that is stored on a server and sent via the internet and accessed through a browser interface. From the above understanding, it can be concluded that a web application is an application that is accessed using a web browser via the internet or intranet network. A web application is also computer software coded in a programming language that supports web-based software such as HTML, JavaScript, CSS, Ruby, Python, PHP, Java and other programming languages. Around 1,200,000 mobile applications are currently available on the Google Playstore, thus offering users many choices. Therefore, it is important for the user to know the functionality of the application to be installed, what other consumers think about it, and then decide whether to install it or not. User ratings usually consist of two parts: the rating and the review.

In this case, the value of ratings and reviews can reach thousands or even millions which makes it

How to Cite:

Imanuddin, S. H., Adi, K., & Gernowo, R. (2023). Sentiment Analysis Naive Bayes Method on SatuSehat Application. *Jurnal Penelitian Pendidikan IPA*, 9(7), 5524–5531. <https://doi.org/10.29303/jppipa.v9i7.4054>

difficult for application developers to find ways to improve application performance based on thousands of text comments (Alzubaidi et al., 2021). One of them is the SatuSehat application, which is an upgraded version of the PeduliLindungi application in the new normal era. The purpose of this application is to monitor people who have been vaccinated, need to carry out activities in crowded places (Nurmansyah et al., 2022). The SatuSehat application helps manage people's vaccination history and saves a history of places visited by people using the QR code scanning feature. In addition, the SatuSehat application provides registration opportunities for those who have not received vaccinations or do not know where to register for their vaccinations.

In its development, the SatuSehat application raises pros and cons which can be seen in the comments on the Google Playstore application. Given these problems, sentiment analysis is a perfect field study to process a collection of reviews and classify them into positive and negative reviews. Sentiment analysis, also known as opinion mining, involves processing data about an individual's opinion, behavior and emotions towards an entity (Shaik et al., 2023). These entities are individuals, or topics that are happening around us. The results of data processing can be in the form of information about how people view entities extracted from sentences (Yang et al., 2022).

The method that will be used to classify user reviews of the SatuSehat application is Naïve Bayes. Naïve Bayes can provide better performance than other methods, if the number of samples available for model training is limited. This method tends to be more computationally efficient and requires less training data to produce a good model. In addition, the algorithm in Naïve Bayes is simple but has relatively high accuracy and speed (Diba & Nugraha, 2020). Several studies have been conducted previously using the Naïve Bayes method, such as research by (Salsabila et al., 2022) who conducted a sentiment analysis about the COVID-19 vaccine using the Naïve Bayes algorithm with a total of 3,780 data obtained. The results of the analysis show that the majority of user reviews are positive (60.3%), while neutral reviews (34.4%) outnumber negative reviews (5.4%). The accuracy value generated by the Naive Bayes Classifier algorithm is 0.93 (93%).

Other studies have also analyzed public opinion regarding vaccination with a total of 845 data by comparing the Naïve Bayes and SVM methods. The results obtained are that the classification of the Naïve Bayes method has an accuracy rate of 85.59%, which is higher than the SVM, which is 84.4% (Syahputra et al., 2022). Another research is research conducted by Illia et al. (2022), who conducted a sentiment analysis on the

PeduliLindungi application with a total data of 4,636. The results of the Naïve Bayes classification show an accuracy value of 90% with a negative precision of 10%, a positive precision value of 99%, a negative recall value of 68% and a positive recall value of 91%.

The results of this study indicate that the positive reviews are higher than the negative reviews. However, this research was conducted in 2022 and before the application was developed into SatuSehat. Therefore, the purpose of this research is to classify user reviews of the SatuSehat application into positive and negative sentiments and assess the performance of the Naïve Bayes method regarding public opinion on the use of the SatuSehat application based on reviews from the Google Playstore application. In addition, it aims to display a visualization to find the words that appear most often in user reviews.

Method

Therefore, it is important to classify user reviews into positive and negative sentiments using the Naïve Bayes method This research flowchart is described as Figure 1.

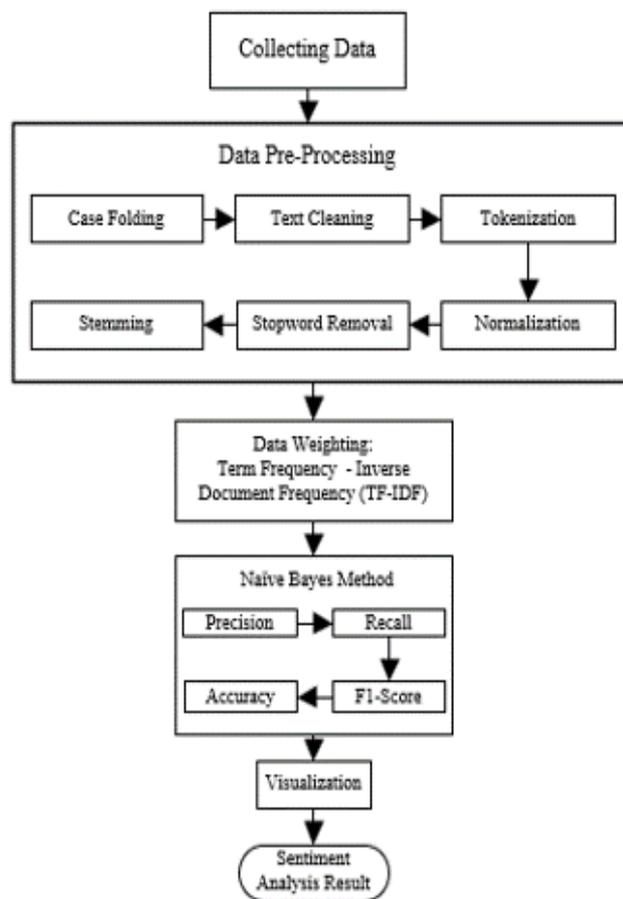


Figure 1. Research Flowchart

Figure 1, show a research flowchart in which the first stage is data collection taken through user reviews of the SatuSehat application on the Google Playstore application. After that, text processing is carried out to select data so that it is more organized with the stages of case folding, text cleaning, tokenization, normalization, stopword removal and stemming (Syahputra et al., 2022). Case folding is the process of manipulating text by changing all uppercase letters to lowercase. Text Cleaning is a step to clean unnecessary text data, such as deleting null values, removing duplicate values, removing unwanted characters and symbols in text, and choosing which variables to use.

Tokenization divides a sentence or paragraph into individual word blocks, or what are called tokens. Normalization refers to the stage of changing words in abbreviated, colloquial and non-standard forms into standard forms. Stopword removal removes words which are usually many and are considered meaningless (Sarica & Luo, 2021) and stemming is the process of replacing words in basic words with affixes. The next stage is weighting using TF - IDF (Term Frequency Inverse Document Frequency). This process is carried out with the aim of converting text into numeric and counting how many words appear in the document. The results for TF - IDF are obtained from the results of TF multiplied and the formula can be seen in equations 1, 2 and 3 (Alfarizi et al., 2022).

Result and Discussion

Data Collecting

Features reviews from users is one of the features that Google Play has that can be used to see reviews from other users (Sutino & Siahaan, 2019). Apart from that, Google Play has a feature to analyze user reviews application making it easier for application makers to monitor the state of the application based on words or topics that appear frequently in related reviews design, privacy, profile, resource usage, speed, stability, updates, uninstall and usability of the application (Fuad & Al-Yahya, 2021).

The data used in this study came from reviews of the SatuSehat application on the Google Playstore application and for a period of three months, from January 2023 to March 2023. The data collected consisted of 25,00 reviews with 18,35 belonging to the negative class and 6,641 belonging to the positive class. The results of data collection are imported into the Jupyter Notebook software using python. In this process, rating data is converted into sentiment which can be seen in Table 1.

Changing the rating to sentiment in Table 1 aims to categorize text or data based on the sentiment contained in it, namely in this study, sentiment with a scale of one to three is defined as negative sentiment, and positive sentiment is on a scale of four to five. Sentiment analysis also uses an algorithm to process and classify built data. There are many algorithms can be used in analytical research sentiment. There are the usual top ten algorithms used, among them was C4.5 (Wankhade et al., 2022).

Table 1. Changing Ratings to Sentiments

Text	Score	Tag
Can't be opened.. The application asks for an update but can't be updated.	3	Negative
The app update is still running, so you can't use it optimally (still learning how to use this app)	5	Positive
This ugly application has been updated and can't even be opened	2	Positive
Why can't the verification code be used????? I've used different codes many times but it still doesn't work???	1	Negative
Good	4	Positive

Pre-Processing

The stages in preprocessing are that the original data that has been collected will be further processed with the aim of cleaning the data, namely removing noise, clarifying features, and converting the original data according to research needs (Fan et al., 2021). In this study, pre-processing was carried out by case folding, text cleaning, tokenization, normalization, stopword removal and stemming. The first process is case folding which aims to make changes to letters that were

previously mixed between capital letters and lowercase letters to become only lowercase letters. The case folding stage in this study can be seen in Table 2.

The second process is the text cleaning process which aims to clean sentences from link or hyperlinks, punctuation marks, mentions of username (@username), urls and numbers that should not be in the dataset. The stages of text cleaning in this study can be seen in Table 3.

Table 2. Case Folding

Before Case Folding	@Satusehat The error... uninstalled it with the same result... https://t.co/xxUp5fDYaG @KAI121 Is it possible to show a vaccine certificate only through the Satu Sehat application or not, admin? No need to use letters gt @KAI121
After Case Folding	@ satusehat the error... I've uninstalled it with the same results... https://t.co/xxup5fdyag @ kai1 21 Can you show a vaccine certificate only through the Satu Sehat application or not, admin? No need to use letters gt @kai121

Table 3. Text Cleaning

Before Text Cleaning	@satusehat the error. I've uninstalled it with the same results... https://t.co/xxup5fdyag @kai121 can you show a vaccine certificate only through the Satu Sehat application or not, admin? No need to use letters gt @kai121
After Text Cleaning	The error has been uninstalled, the result is the same If you show a vaccine certificate only through the Satu Sehat application, can you or no, admin, you don't have to use the GT letter

Table 4. Tokenization

Before Tokenization	the error has been uninstalled, the result is the same If you show a vaccine certificate only through the Satu Sehat application, can you or no, admin, you don't have to use the GT letter
After Tokenization	['error', 'sdh', 'uninstall', 'same', 'sj', 'result'] ['if', 'show', 'certificate', 'vaccine', 'only', 'through', 'application', 'onehealthy', 'doang', 'can', 'no', 'min', 'gausah', 'use', 'letter', 'gt']

The next process is the tokenization process which can be seen in Table 4. At this stage all the words that have been collected will have their punctuation removed such as symbols, characters and all that are not letters and the purpose is to separate sentences into words.

Tokenization is the separation or decoding of data textual into tokens, which are smaller means components (Verma et al., 2018). Principle of tokenization Konashevych (2020), namely separating

each word that composes sentence in the document. It can be classified into sentence tokenization and word tokenization. The normalization process is a process that is carried out after everything has gone through the tokenize process. The normalization process is the process of changing words that were originally in the form of abbreviations or slang words back into standard words (Rianto et al., 2021). The normalization stage can be seen in Table 5.

Table 5. Normalization

Before Normalization	['error', 'sdh', 'uninstall', 'same', 'sj', 'result'] ['if', 'show', 'certificate', 'vaccine', 'only', 'through', 'application', 'satusehat', 'doang', 'can', 'no', 'min', 'gausah', 'use', 'letter', 'gt']
After Normalization	['error', 'already', 'uninstall', 'same', 'only', 'result'] ['if', 'show', 'certificate', 'vaccine', 'only', 'through', 'application', 'satusehat', 'doang', 'can', 'no', 'min', 'gausah', 'use', 'letter', 'that']

Database normalization is rarely performed on small databases, and is considered unnecessary for personal use (Albarak & Bahsoon, 2018). However, as the information contained in a database develops, the normalization process will be very helpful in saving the space used by each table in it (Shorten & Khoshgoftaar, 2019), as well as speeding up the process of requesting data. The purpose of database normalization is to reduce data repetition Arini et al. (2019), ensure data is in the right position, no amount of data enters the database,

there is deletion of inappropriate data and data that is changed inappropriately. At the stopwords removal stage, if there are conjunctions or irrelevant words, they will be removed, such as first, second and third person pronouns, a name, conjunctions and question words. These words are words that have no meaning when separated from other words and are not related to adjectives related to sentiment which can be seen in Table 6.

Table 6. Stopword Removal

Before Stopword Removal	[if, 'show', 'certificate', 'vaccine', 'only', 'through', 'application', 'satusehat', 'doang', 'can', 'no', 'min', 'gausah', 'use', 'letter', 'that']
After Stopword Removal	[if, 'show', 'certificate', 'vaccine', 'only', 'through', 'application', 'doang', 'can', 'use', 'letter', 'so']

The last process is stemming, which aims to replace word forms into basic words according to the structure of the Indonesia Dictionary. such as words that have

initial and final affixes such as me, mem, meng, meng, di, per, ber, an, kan, i, nya etc. The affix words will be changed to basic words as shown in Table 7.

Table 7. Stemming

Before Stemming	[error, 'already', 'uninstall', 'same', 'only', 'result']
After Stemming	[if, 'show', 'certificate', 'vaccine', 'only', 'via', 'application', 'doang', 'can', 'use', 'letter', 'so']

Weighting

Table 8. Weighting word using TF-IDF method

Document	TF-IDF	Label
['enhance', 'hopefully', 'smooth', 'going', 'download', 'vaccine', 'certificate']	[0.4670, 0.2693, 0.2765, 0.2484, 0.4347, 0.1452, 0.1557]	54%15 Positive
['very', 'good']	[0.7669, 0.6417]	Positive
['email', 'account', 'sign in', 'update', 'difficult']	[0.2691, 0.5337, 0.3711, 0.1618, 0.2199]	Negative
['really', 'enter', 'complicated']	[0.6098, 0.4416, 0.6581]	Negative

The weighting will be carried out using Term Frequency - Inverse Document Frequency (TF-IDF). TF-IDF which aims to give weight to the relationship of a word (Term) to a document before doing calculations on the algorithm and overcome problems in classifying data into positive or negative sentiments. The stages of the TF-IDF that have been carried out can be seen in Table 8.

Naïve Bayes Classification

```
[19]: # Naive Bayes
print(classification_report(data_test['Target'], nbpredictions))
```

	precision	recall	f1-score	support
NEGATIVE	0.98	0.98	0.98	18173
POSITIVE	0.94	0.94	0.94	6396
accuracy			0.97	24569
macro avg	0.96	0.96	0.96	24569
weighted avg	0.97	0.97	0.97	24569

Figure 2. Naïve Bayes Classification result

Figure 2 shows the results of the Naïve Bayes classification which shows an accuracy of 97% with negative sentiment results, namely precision has a value of 98%, recall has a value of 98% and f1-score has a value of 98%, while positive sentiment results, namely precision has a value of 94%, recall has a value of 94% and f1-score has a value of 94%. Furthermore, data that has gone through the weighting stage will be classified into two classes, namely the positive class and the negative class with the requirement that if the weight is > 0 then it enters the positive class while the weight < 0 will enter the negative class in the form of a pie chart. The percentage graph in this study can be seen in Figure 3.

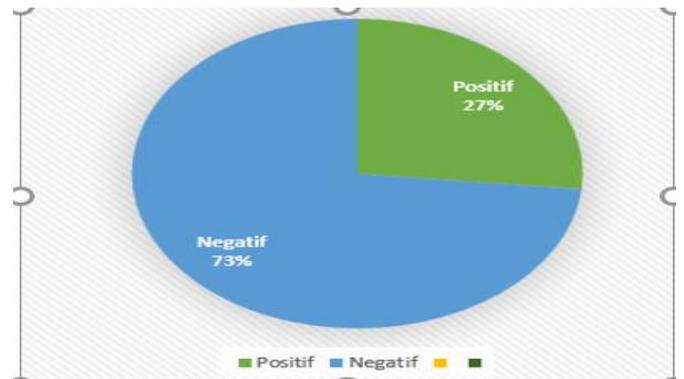


Figure 3. Positive and Negative ratio

Based on Figure 3, it can be seen that the results of sentiment generate more negative sentiment, which is equal to 73.4%, compared to positive sentiment, which is equal to 26.6%. From these results it can be concluded that the community still has an unfavorable view of the SatuSehat application. Therefore, it is hoped that in the future, the SatuSehat application can improve its quality so that users have a positive view of using it. Because a successful application is an application that can make

users feel satisfied, useful and make users want to use it again in the future (Lee & Trimi, 2018).

Visualization

Visualization aims to extract information in the form of topics that are often discussed by SatuSehat application users in the Google Playstore application review column.



Figure 4. WordCloud for Positive word

Figure 4. shows positive wordclouds that are often used by SatuSehat application users to provide reviews on the Google Playstore application. In wordcloud, the larger the word size indicates that the more often these words are used by SatuSehat application users as topics of conversation. Some of the topics that are often discussed by SatuSehat application users are good, "good" means fine, "good" means nice, and "mantap" means excellent.



Figure 5. WordCloud for Negative word

Furthermore, Figure 5 shows negative word clouds that are often used by SatuSehat application users to provide reviews on the Google Playstore application. Some of the topics that are often discussed by SatuSehat application users in negative reviews are the words update, difficult to mean "difficult", "strange" to mean strange, log in, bugs, and so on.

Conclusion

This study aims to classify user reviews of the SatuSehat application into positive and negative sentiments and assess the performance of the Naive Bayes method regarding public opinion on the use of the SatuSehat application based on reviews from the Google Playstore application. In addition, this study aims to display a visualization to find the words that appear most frequently in user reviews with the data obtained as many as 25,000. The results obtained in this study are that using the Naive Bayes method is able to provide high accuracy results, namely 97% with negative sentiment results, namely precision has a value of 98%, recall has a value of 98% and f1-score has a value of 98%, while sentiment results are positive, namely precision has a value of 94%, recall has a value of 94% and f1-score has a value of 94%. Furthermore, the results of the percentage chart show that sentiment results generate more negative sentiment, with 73.4% compared to positive sentiment, which is equal to 26.6%. Therefore, the results of this study can be a recommendation for the company to improve the quality of the SatuSehat application so that it is better in the future.

Acknowledgments

Thanks to all parties who have supported the implementation of this research. I hope this research can be useful.

Author Contributions

Conceptualization: Shahmirul Hafizullah Imanuddin, data curation: Kusworo Adi, funding acquisition: Rahmat Gernowo, methodology, visualization: Shahmirul Hafizullah Imanuddin, writing-original draft: Kusworo Adi, writing-review & editing: Rahmat Gernowo.

Funding

This research was independently funded by researchers.

Conflicts of Interest

No Conflicts of interest.

References

Albarak, M., & Bahsoon, R. (2018). Prioritizing technical debt in database normalization using portfolio theory and data quality metrics. *Proceedings of the 2018 International Conference on Technical Debt*, 31-40. <https://doi.org/10.1145/3194164.3194170>

Alfarizi, M. I., Syafaah, L., & Lestandy, M. (2022). Emotional Text Classification Using TF-IDF (Term Frequency-Inverse Document Frequency) And LSTM (Long Short-Term Memory). *JUITA: Jurnal Informatika*, 10(2), 225. <https://doi.org/10.30595/juita.v10i2.13262>

- Alsaffar, M., Aljaloud, S., Mohammed, B. A., Al-Mekhlafi, Z. G., Almurayziq, T. S., Alshammari, G., & Alshammari, A. (2022). Detection of Web Cross-Site Scripting (XSS) Attacks. *Electronics*, 11(14), 2212. <https://doi.org/10.3390/electronics11142212>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Arini, F. Y., Arifudin, R., & Aris, M. (2019). Applied structured database in a small project. *Journal of Physics: Conference Series*, 1321(3), 032130. <https://doi.org/10.1088/1742-6596/1321/3/032130>
- Diba, S. F., & Nugraha, J. (2020). Implementation of Naive Bayes Classification Method for Sentiment Analysis on Community Opinion to Indonesian Criminal Code Draft. In *Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019)*, 186-192. <https://doi.org/10.2991/assehr.k.201010.027>
- 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, 652801. <https://doi.org/10.3389/fenrg.2021.652801>
- Fauzan, R., Krisnahati, I., Nurwibowo, B. D., & Wibowo, D. A. (2022). A Systematic Literature Review on Progressive Web Application Practice and Challenges. *IPTEK The Journal for Technology and Science*, 33(1), 43. <https://doi.org/10.12962/j20882033.v33i1.13904>
- Fuad, A., & Al-Yahya, M. (2021). Analysis and Classification of Mobile Apps Using Topic Modeling: A Case Study on Google Play Arabic Apps. *Complexity*, 2021, 1-12. <https://doi.org/10.1155/2021/6677413>
- Illia, F., Eugenia, M. P., & Rutba, S. A. (2022). Sentiment Analysis on PeduliLindungi Application Using TextBlob and VADER Library. *Proceedings of The International Conference on Data Science and Official Statistics*, 2021(1), 278-288. <https://doi.org/10.34123/icdsos.v2021i1.236>
- Konashevych, O. (2020). General Concept of Real Estate Tokenization on Blockchain: The Right to Choose. *European Property Law Journal*, 9(1), 21-66. <https://doi.org/10.1515/eplj-2020-0003>
- Lee, S. M., & Trimi, S. (2018). Innovation for creating a smart future. *Journal of Innovation & Knowledge*, 3(1), 1-8. <https://doi.org/10.1016/j.jik.2016.11.001>
- Nurmansyah, M. I., Rosidati, C., Yustiyani, Y., & Nasir, N. M. (2022). Measuring the Success of PeduliLindungi Application Use for Supporting COVID-19 Prevention: A Case Study among College Students in Jakarta, Indonesia. *Kesmas: Jurnal Kesehatan Masyarakat Nasional*, 17(sp1). <https://doi.org/10.21109/kesmas.v17isp1.6057>
- Rianto, Mutiara, A. B., Wibowo, E. P., & Santosa, P. I. (2021). Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation. *Journal of Big Data*, 8(1), 26. <https://doi.org/10.1186/s40537-021-00413-1>
- Saarikko, T., Westergren, U. H., & Blomquist, T. (2020). Digital transformation: Five recommendations for the digitally conscious firm. *Business Horizons*, 63(6), 825-839. <https://doi.org/10.1016/j.bushor.2020.07.005>
- Salsabila, S. K., Jondri, J., & Astuti, W. (2022). Analysis of Community Sentiment on Twitter towards COVID-19 Vaccine Booster Using Ensemble Stacking Methods. *Building of Informatics, Technology and Science (BITS)*, 4(2), 467-473. <https://doi.org/10.47065/bits.v4i2.1902>
- Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care*, 27(1), 75. <https://doi.org/10.1186/s13054-023-04380-2>
- Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *PLOS ONE*, 16(8), e0254937. <https://doi.org/10.1371/journal.pone.0254937>
- Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). Sentiment analysis and opinion mining on educational data: A survey. *Natural Language Processing Journal*, 2, 100003. <https://doi.org/10.1016/j.nlp.2022.100003>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Sutino, Q. L., & Siahaan, D. O. (2019). Feature extraction from app reviews in google play store by considering infrequent feature and app description. *Journal of Physics: Conference Series*, 1230(1), 012007. <https://doi.org/10.1088/1742-6596/1230/1/012007>
- Syahputra, R., Yanris, G. J., & Irmayani, D. (2022). SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter. *Sinkron*, 7(2), 671-678. <https://doi.org/10.33395/sinkron.v7i2.11430>
- Verma, M., Bridges, R., & Hollifield, S. (2018). ACTT: Automotive CAN Tokenization and Translation. *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 278-

283.

<https://doi.org/10.1109/CSCI46756.2018.00061>

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780.

<https://doi.org/10.1007/s10462-022-10144-1>

Yang, Y., Wu, Z., Yang, Y., Lian, S., Guo, F., & Wang, Z. (2022). A Survey of Information Extraction Based on Deep Learning. *Applied Sciences*, 12(19), 9691. <https://doi.org/10.3390/app12199691>