Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. 5. No. 1; e-ISSN: 2808-4519

Sentiment Analysis of Jakarta Kini (JAKI) Application Reviews using the Naive Bayes Method

Romi Satrio Wicaksono[1]*, Ade Priyatna[2]

[1,2]*Universitas Bina Sarana Informatika, Depok, Indonesia*
romisatriowicaksono29@gmail.com[1*], ade.aeq@bsi.ac.id[2]

Abstract

Jakarta Kini (JAKI) is a super-app developed by Jakarta Smart City to simplify public service access in the DKI Jakarta Province. As an application widely used by the public, JAKI has received thousands of user reviews on the Google Play Store, reflecting public opinion on its features and performance. This study aims to classify user reviews into positive, negative, and neutral sentiment categories using the Naive Bayes algorithm. The research method includes collecting review data through web scraping, text preprocessing, sentiment analysis, data labeling, model building, and model evaluation using the Orange platform. The results show that the Naive Bayes algorithm successfully classified 3,226 review data with a perfect accuracy of 100%, as confirmed by the confusion matrix and other evaluation metrics (precision, recall, F1-score, MCC). The sentiment distribution reveals that most reviews are neutral, followed by negative and then positive sentiments. This indicates that public perception of the JAKI application tends to be moderate, highlighting the need for developers to improve the quality of digital public services. This research is expected to serve as a reference for utilizing machine learning-based sentiment analysis in evaluating public service applications

Keywords: *JAKI Application, Naive Bayes, Sentiment Analysis, Orange, Google Play Reviews*

1. Introduction

The era of increasingly rapid digital transformation requires the government to keep pace with these developments by implementing technology to improve the quality of public services. One innovation that can be implemented in this progress plan is the Jakarta Kini (JAKI) application, developed by Jakarta Smart City under the auspices of the Jakarta Provincial Communications, Informatics, and Statistics Office. As an integrated service center, JAKI was created with the primary goal of helping the public more easily access a wide range of government-provided information and services, ranging from public complaints and information on natural disasters such as floods, food prices, taxes, and air quality. With its integrated concept, within a single platform, JAKI is expected to capture public attention, foster public participation, foster bureaucratic transparency, and facilitate faster access to technology-based services [1].

The rapid development of digital technology is encouraging the government to provide technology-based public services. One implementation of this innovation is the Jakarta Kini (JAKI) application, developed by Jakarta Smart City under the auspices of the Jakarta Provincial Communications, Informatics, and Statistics Office. The JAKI application serves as a digital public service center that provides various features such as citizen reporting (JakLapor), food price information (JakPangan), weather and flood information, vaccination services, and air quality, all integrated into a single platform.

2. Literature Review

2.1. Sentiment Analysis

Sentiment analysis is a process in Natural Language Processing (NLP) that aims to identify, extract, and classify opinions or feelings from text to determine whether the sentiment is positive, negative, or neutral. [2]

using the Naïve Bayes method and SVM, sentiment analysis is a technique used to understand user perceptions based on opinions expressed in textual form, particularly on social media or app reviews. This technique is crucial because it helps organizations or companies make strategic decisions based on public or customer opinion.
Various classification algorithms have been used in sentiment analysis. Among them, Naïve Bayes remains widely used due to its speed, simplicity, and effectiveness in Indonesian text [3]. The development of transformer-based models such as BERT demonstrates progress in sentiment analysis. However, traditional methods like Naïve Bayes remain relevant because they are computationally less demanding and still provide competitive results [3].

2.2. Jakarta Today (JAKI)

JAKI is a super-app developed by Jakarta Smart City under the DKI Jakarta Provincial Government. This application aims to make it easier for Jakarta residents to access various public services and information related to the city of Jakarta in one digital platform. With various features such as JakLapor for reporting problems such as garbage, flooding, or damaged roads, JakRespons for displaying the status of responses to citizen reports, and JakPangan which provides information on staple food prices in Jakarta markets, JAKI functions to improve the efficiency of public services. In addition, this application also provides services such as vaccination registration, health information, and public satisfaction surveys through the JakSurvei feature. With the aim of increasing transparency and public participation, JAKI not only functions as a platform for submitting complaints but also as a means to accelerate the government's response to existing problems in Jakarta [4].

2.3. User Reviews and Ratings on Google Play Store

User reviews and ratings on the Google Play Store are feedback provided by app users about the apps they download and use. Reviews typically consist of text describing the user's experience or opinion about the app, while ratings are numerical scores assigned using a star system (1 to 5 stars). Ratings provide a quick overview of an app's quality, while reviews provide deeper insights into its strengths, weaknesses, and user expectationsReviews on the Google Play Store have been widely used to evaluate app quality because they are open and dynamic[4]

Reviews and ratings play a crucial role in app development, as they provide direct feedback from users that can be used to improve app quality. Reviews and ratings also influence an app's reputation, where high ratings can increase the trust of new users, while negative reviews can encourage developers to make improvements. Furthermore, the data is dynamic and continuously updated as reviews increase, and includes important information such as app ID, user name, review date, rating, review content, and app version. With organized data structures, both in CSV and JSON formats, developers and researchers can leverage them to evaluate application performance and improve features and services to suit user needs.

2.4. Naive Bayes Classifier

The Naive Bayes Classifier is a probabilistic-based classification algorithm that utilizes Bayes' Theorem, assuming that each feature is independent. Recent studies confirm that Naive Bayes remains relevant despite the emergence of many new algorithms [2]. The Naive Bayes algorithm is a simple yet effective text classification method, including for app reviews . Naive Bayes is also known for its efficiency in document classification. In the context of sentiment analysis, this algorithm is widely used due to its efficiency and simplicity in processing text data.

2.5. Orange

Orange is an open-source, visual programming-based software used for data analysis and machine learning. This application provides a drag-and-drop interface, allowing
users to build data analysis pipelines without writing programming code. Orange supports various machine learning methods such as classification, regression, clustering, and interactive data visualization. Orange's strength lies in its ease of use and its ability to integrate various algorithms such as Naive Bayes, Support Vector Machines, and Random Forests in a single, integrated platform [5]. The use of machine learning in sentiment analysis on social media demonstrates

3. Research Process and Steps

This study uses a quantitative approach with a text classification method based on sentiment analysis.Variousvisual programming platforms have been developed to support data mining experiments. However, this study focuses on the application of Naive Bayes, which can be implemented efficiently without requiring additional complexity (Sonk & Tunger, 2024). The quantitative approach was chosen because this study aims to process numerical and statistical data to obtain an objective picture of public perception of the JAKI (Jakarta Kini) application. Analysis was conducted on user reviews taken from the Google Play Store using the Naive Bayes classification algorithm, known as a probability-based statistical method and widely used in text classification tasks. The algorithm implementation was carried out using Orange Data Mining software, which provides an interactive visual environment for data processing and machine learning.

The steps taken are shown in Figure 1

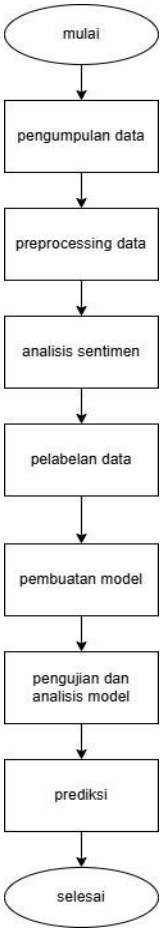

Fig. 1: Research Process and Steps

The research flow in this study consists of a series of structured stages, starting with a literature review to strengthen the theoretical framework, followed by review data collection, text data preprocessing, sentiment analysis, data labeling, classification model development, model performance evaluation, and conclusion drawing. Each stage is interconnected and structured to systematically support the achievement of the research objectives. This flowchart facilitates understanding of the research process from start to finish, while ensuring that each methodological step is carried out in a planned and documented manner.

3.1. Research Object

The object of this research is the Jakarta Kini (JAKI) application, a superapp developed by Jakarta Smart City under the auspices of the Jakarta Provincial Communication, Informatics, and Statistics Agency. This application is designed to provide various integrated public services in one digital platform, such as reporting citizen problems through the JakLapor feature, food price information through JakPangan, flood and weather information, and vaccination and air quality services. The JAKI application was chosen as the research object because it has a large number of users and diverse reviews on the Google Play Store, which makes it a relevant data source for sentiment analysis. Through these user reviews, public perceptions of the performance, ease of use, and effectiveness of the services in the application can be evaluated. This research specifically focuses on collecting and processing JAKI review data from the Google Play Store using web scraping methods, which are then automatically analyzed using the Naive Bayes algorithm to classify user sentiment into three categories: positive, neutral, and negative. Thus, the JAKI application as a research object makes an important contribution to understanding public opinion towards technology-based local government digital services.

3.2. Review Data Collection

Data scraping using Python and Google Colab is the process of automatically extracting data from online sources. Using Google Colab, users can run Python scripts directly from the cloud without the need for additional software installation. This process allows data, such as app reviews, to be collected in large quantities and stored in a format ready for analysis, such as a CSV file. Google

Colab simplifies this process by providing an organized interface, access to relevant Python libraries, and the ability to save or download scraping results directly from the work environment.

3.3. Data Processing and Analysis Methods

There are two main stages in sentiment analysis using Orange Data Mining. The first stage is data pre-processing, where review data from files is imported and processed through several processes, such as text transformation using the Corpus and Preprocess Text nodes, and sentiment analysis using the Sentiment Analysis node. The results are visualized using Word Clouds and Bar Plots to understand word patterns and sentiment distribution. The processed data is then stored or forwarded for further analysis. The second stage is classification and model evaluation, where the pre-processed data is used to build a classification model using the Naive Bayes algorithm. This model is evaluated using the Test and Score nodes, with the evaluation results visualized in a Confusion Matrix to assess classification accuracy. Additionally, the Predictions and

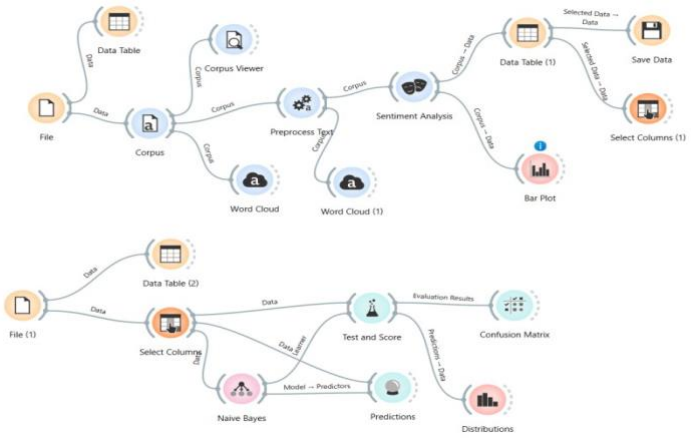

Fig. 2: Source: Personal Author Document, Orange, Distributions nodes are used to display the sentiment prediction results and probability distributions of the model.

3.3. 1 Preprocessing

The process begins with data loading via the File widget, which is then passed to Corpus to convert the raw data into an analyzable text format. This data can be viewed via the Corpus Viewer and visualized as a Word Cloud. The Preprocess Text widget performs several important steps, such as Transformation, which includes lowercase conversion and accent removal to standardize the text. Next, the Tokenization process is performed using the RegExp \w+ pattern, which separates words based on letters and numbers. The Filtering stage removes stopwords (although still set to Kazakh, it should be changed to Indonesian), removes symbols using a regular expression, and sets word frequency limits to ensure only relevant tokens are processed. This process also limits tokens to nouns and verbs, which are crucial for meaning analysis. Finally, in the Normalization stage, the Porter Stemmer method is used to truncate words into their basic forms to maintain word consistency. All these preprocessing stages aim to clean and simplify the text data so that it is ready for accurate and efficient use in sentiment analysis. Preprocessing steps such as tokenization, stopword removal, and stemming significantly impact classification model accuracy [1]. Combining preprocessing with TF-IDF has been shown to improve Naive Bayes accuracy [2]

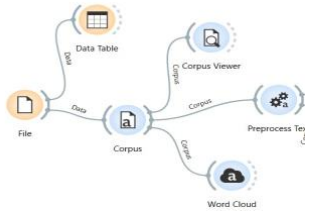

Fig. 3: Source: Personal Author Document, Orange,

3.4. Sentiment Analysis

In this configuration, the selected method is Multilingual Sentiment with Indonesian, which is suitable for analyzing Indonesian-language text data such as JAKI app reviews. This method generates a numeric sentiment value (polarity) for each text, where a

positive value indicates positive sentiment, a negative value indicates negative sentiment, and a value of zero indicates neutral sentiment.
These values are numeric and reflect the strength of user opinion, so they can be used as a basis for automatic sentiment labeling or visualization of opinion distribution in further analysis.

3.5. Data Labeling

The results of the sentiment analysis process are used to label each review data set. Labeling is performed both automatically and semi-manually, using a sentiment scoring scheme as a reference: values above zero are categorized as positive sentiment, values of zero as neutral, and values below zero as negative. These labels are then added to the dataset file and used in training and testing the Naive Bayes model. This labeling process also ensures that the data has the ground truth necessary for evaluating the performance of the classification algorithm.

3.6. Model Creation

This dataset has undergone a sentiment analysis phase, resulting in a sentiment column as a numeric feature, a label as a classification target (sentiment category: positive, neutral, or negative), and content as a meta file containing the original review text. In the workflow, data is imported from an external file using the File node and then displayed via a Data Table for exploration. Next, the Select Columns node is used to select relevant columns, namely the sentiment column as a feature and the label as a target. A classification model is built using the Naive Bayes algorithm to learn the relationship between sentiment values and labels, so that it can be used to predict sentiment categories in new data.

3.7. Model Testing and Analysis

The Test and Score node is used to test the performance of the Naive Bayes model created by evaluating metrics such as accuracy, precision, recall, and F1-score to assess the quality of sentiment classification. The output of this node is the model evaluation results, which can be forwarded to other nodes for further analysis. One of the nodes receiving this output is the Confusion Matrix, which visualizes the test results in a matrix showing the number of correct and incorrect predictions for each sentiment category: positive, neutral, and negative. This matrix is useful for understanding the distribution of classification errors and provides a more detailed picture of the model's ability to accurately classify sentiment.

3.8. Prediction

The Predictions node generates sentiment prediction labels (positive, neutral, or negative) based on a trained model, such as Naive Bayes. These results provide information about the sentiment classification for new data. Meanwhile, the Distributions node displays the model's prediction probabilities, indicating the model's confidence level in each prediction.
In the context of data analytics and machine learning, "Prediction" is the fundamental process of utilizing mathematical or algorithmic models, trained on historical data with known outcomes, to estimate or forecast unknown future values or outcomes. This process begins with data pre-processing to ensure quality, continues with building and training a model (e.g., using regression or artificial neural networks) on historical data, then validating the model to ensure its generalizability, and finally applying it to new data to generate predictions. These predictions, whether numerical or categorical, are highly significant for strategic decision-making in various sectors such as economics, healthcare, and risk management.

4. Results and Discussion

This study was conducted to classify user sentiment toward the Jakarta Kini (JAKI) application using the Naive Bayes algorithm. Review data was collected from the Google Play Store using a scraping method using the google-play-scraper library, then further processed in the Orange application.

4.1. Research Results

4.1.1. Data Preprocessing Results

The analysis process consists of several stages: from preprocessing, sentiment analysis, classification, to model performance evaluation. The data is analyzed by dividing the sentiment value into three categories: if the value <0 then it is categorized as negative sentiment, if = 0 as neutral, and if >0 as positive. Figure IV.1 displays the results of processing review data that has been given a sentiment value and labeled based on that category. The table shows how the system groups the review text into three types of sentiment based on the numerical scores generated by the system**.**

sentiment	content	label
8	bagus	positif
8	bagus	positif
-10	tidak puas karena laporan dan yang ngerjain laporan tidak sesuai sama apa yang di kerjakan	negatif
-3	servernya digedein dong. mikir.	negatif
-14.2857	tidak jelas nih aplikasiii kuota penuh terus	negatif
-6	Proses daftar di persulit	negatif
100	Terima kasih banyak	positif
0	Kenapa cek kendaraan tdk terperinci lagi ?????	netral
-9.09091	Kurang bagus ada laporan tindak lanjuti lama	negatif
-4.7619	aplikasi gk update	negatif
8	Mantul	positif
0	gimana cara pesannya?	netral

Fig. 4: Data Preprocessing Results

4.1.2. Test Results

After the data was sentimentally labeled, the classification process was performed using the Naive Bayes algorithm available on the Orange platform. Model testing was conducted using 20-fold cross-validation and stratified data sampling to ensure a balanced data distribution. The results of the model performance evaluation are displayed in the form of a confusion matrix, as shown in Figure IV.2. Based on the confusion matrix, it is known that the Naive Bayes model successfully predicted all data accurately without misclassification. A total of 595 negative data were successfully classified as negative, 1707 neutral data were correctly classified as neutral, and 924 positive data were also correctly predicted. With no misclassified data, the model accuracy reached 100%.

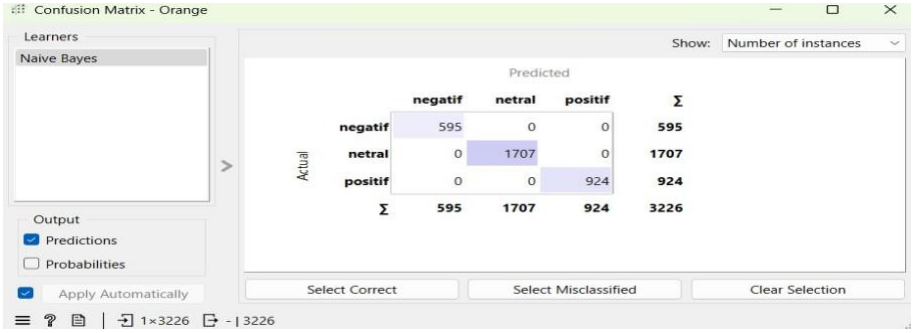

Fig. 5: Naive Bayes Model Evaluation Results in Orange (Test and Score)

The model evaluation scores for accuracy, precision, recall, and other metrics all show a perfect score (1,000), indicating excellent classification performance.

The 100% accuracy result obtained in this study is likely due to data leakage between features and labels. In the current flow, the sentiment column (the polarity number generated by the Sentiment Analysis module) serves as a feature, while class labels are derived directly from the same value, with the rule: sentiment > 0 is categorized as positive, sentiment = 0 as neutral, and sentiment < 0 as negative. Thus, the features and labels essentially come from identical information sources. This condition prevents the model from truly learning; instead, it simply "reads the threshold" to consistently produce accurate predictions. Therefore, the evaluation metric obtained is perfect (100%).

Evaluation of the performance of classification algorithms on mobile app sentiment analysis shows that feature selection significantly impacts model accuracy. This aligns with the results of this study using Naive Bayes, where the quality of input features plays a significant role in classification performance.

5. Conclusion

Based on the research results, it can be concluded that:

1. The Naive Bayes algorithm is capable of effectively classifying the sentiment of user reviews of the Jakarta Kini (JAKI) app. The classification process is carried out through systematic stages, starting with collecting review data using scraping techniques from the Google Play Store, followed by a preprocessing stage that includes text cleaning, normalization, tokenization, and stopword removal. Next, sentiment analysis is performed using a numerical polarity approach to categorize reviews into positive, negative, and neutral sentiments. These sentiment labels are then used to train a classification model using the Naive Bayes algorithm on the Orange platform.
2. The results of the model performance test show that the Naive Bayes algorithm has a perfect accuracy rate of 100%, with 3,226 reviews successfully classified without error, as demonstrated by the confusion matrix and other evaluation metrics (precision, recall, F1-score, and MCC), all of which scored 1.000. This demonstrates the Naive Bayes algorithm's reliability in classifying short texts such as app reviews quickly and efficiently.

3. The distribution of review sentiments indicates that the majority of users gave neutral reviews, followed by negative and then positive ones.
4. Based on this distribution, the hypothesis chosen as the final conclusion is: "The majority of user reviews of the JAKI application are neutral or negative."
5. This indicates that although the JAKI application offers various integrated digital public service features, user perceptions tend to be less than enthusiastic. Therefore, application developers and the government need to pay attention to improving the quality of the services offered

6. Suggestions

1. Based on the research results, which showed 100% accuracy due to possible data leakage between features and labels, several improvements are recommended for future research:
2. Use review text (content) as a feature source with representations such as bag-of-words, TF-IDF, or n-grams. Labels should be determined through manual annotation or using star ratings as a proxy, rather than derived from the same sentiment score.
3. Apply training and test data division using the train/test split method or k-fold cross-validation after all preprocessing stages are included in the pipeline to prevent information leakage.
4. Compare the performance of Naive Bayes with other algorithms such as Support Vector Machine (SVM), Logistic Regression,or Random Forest to obtain a more comprehensive picture of model performance.
5. Conduct periodic evaluations on new data to ensure the model is more adaptable to changing user review characteristics.
6. Recognizing that perfect accuracy is nearly impossible to achieve in real-world Natural Language Processing (NLP) tasks, if this occurs, it is necessary to re-evaluate the possibility of data leakage or issues with the dataset. the possibility of data leakage or issues with the dataset.

References

[1] Andriyanto, D., Said, F., Titiani, F., & Erni, E. (2021). Analisis Kesuksesan Aplikasi Jakarta Kini (JAKI) Menggunakan Model Delone and McLean. *Paradigma - Jurnal Komputer Dan Informatika*, *23*(1). https://doi.org/10.31294/p.v23i1.10018

[2] A. U. Khasanah and A. Febriyanti, "Sentiment Analysis of JNE User Perception using Naïve Bayes Classifier Algorithm," *OPSI*, vol. 15, no. 1, p. 124, Jun. 2022, doi: 10.31315/opsi.v15i1.7179.

[3] E. Rahman, N. Namora, and L. Anas, "SENTIMENT ANALYSIS USING NAIVE BAYES ALGORITHM CASE STUDY ON AMAZON E-COMMERCE PRODUCT REVIEWS," *JURTEKSI (jurnal Teknologi dan Sistem Informasi)*, vol. 11, no. 3, pp. 457–462, Jun. 2025, doi: 10.33330/jurteksi.v11i3.3637.

[4] M. Rizqi, A. Rinaldi Dikanda, and D. Rohman, "Journal of Artificial Intelligence and Engineering Applications Naive Bayes Algorithm to Enhance Sentiment Analysis of Coursera Application Reviews on Google Play Store," 2025. [Online]. Available: https://ioinformatic.org/.

[5] Ardika, I. N. A. K., & Wibawa, I. G. A. (2022). *Analisis Sentimen Ulasan Pengguna Aplikasi Pelayanan Masyarakat Dengan Menggunakan Algoritma Random Forest*.

[6] City, J. S. (2021). JAKI Jadi Platform Digital Terpadu DKI Jakarta. https://smartcity.jakarta.go.id/blog/jaki-jadi-platform-digital-terpadu

[7] Dewi, A., & Rahman, T. (2021). Implementasi TF-IDF dan Naive Bayes untuk Analisis Sentimen Ulasan Produk. Jurnal Teknologi Informasi Dan Ilmu Komputer, 8(1), 45–52.

[8] Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2020). Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile) (Vol. 1, Issue 10). http://j-ptiik.ub.ac.id

[9] Liu, B. (2020). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (2nd ed.). Cambridge University Press.

[10] Manurung, D. D. E., Matondang, N. H., & Prasvita, D. S. (2022). Analisis Sentimen pada Ulasan Aplikasi Jakarta Terkini (JAKI) di Google Play Store Menggunakan Metode Support Vector Machine. Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA), 589–598.

[11] Nugroho, B., & Sari, M. (2020). Analisis Sentimen terhadap Aplikasi E-Learning Menggunakan SVM. Jurnal Teknologi Informasi, 9(2), 55–62.

[12] Nugroho, D., Hartati, S., & Prasetyo, E. (2020). Klasifikasi Polaritas Sentimen pada Ulasan Produk Menggunakan Algoritma Naïve Bayes. Jurnal Ilmiah Komputer Dan Informatika, 10(1), 45–50.

[13] Prasetia, I. P. S. D., Wijaya, K. A. S., Yudartha, I. P. D., & Savitri, R. (2024). Analisis Persepsi Masyarakat Terhadap Aplikasi JAKI (Jakarta Kini) di Provinsi DKI Jakarta. Ganaya : Jurnal Ilmu Sosial Dan Humaniora, 7(1), 1–12. https://doi.org/10.37329/ganaya.v7i1.2941

[14] Setiawan, D., & Lestari, A. (2022). Evaluasi Performa Algoritma Klasifikasi pada Analisis Sentimen Aplikasi Mobile. Jurnal Informatika Dan Sistem, 11(3), 88–96.

[15] Sonk, R., & Tunger, D. (2024). Trend mining and topic modelling for future research in urban mobility using Orange data mining tool. *European Journal of Futures Research*, *12*(11–14. https://doi.org/10.1186/s40309-024-00229-1

[16] Suyanto. (2020). *Machine Learning dengan Python dan Scikit-Learn*.

[17] Wijaya, K., & Ramadhan, R. (2023). Perbandingan Metode Naive Bayes dan Random Forest untuk Analisis Sentimen Media Sosial. *Jurnal Ilmu Komputer Terapan*, *15*(1), 44–53.

[18] Sonk, R., & Tunger, D. (2024). Trend mining and topic modelling for future research in urban mobility using Orange data mining tool. *European Journal of Futures Research*, *12*(11–14. https://doi.org/10.1186/s40309-024-00229-1