JSCE (Journal of System and Computer Engineering)
Vol.6, No.4, October 2025, pp. 296-305
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i4.2062

Sentiment Analysis of Gojek User Reviews using TF-IDF and Machine Learning in Orange Platform

Silvia Putri Yuliani [a]; Ari Ati Putri Muharani[b]; Riyana Qori Fatmawati[c]; Faisal Fahmi[*]

Department of Information and Library Science, Airlangga University, Jl Airlangga No 4-6, Surabaya, 60115, Indonesia

[a] silvia.putri.yuliani-2022@fisip.unair.ac.id; [b] ari.ati.putri-2022@fisip.unair.ac.id; [c]riyana.qori.fatmawati-2022@fisip.unair.ac.id

** Corresponding author: faisalfahmi@fisip.unair.ac.id*

Abstract

The present study aims to identify sentiments expressed in user reviews of the Gojek application on the Google Play store. The methodology employed entails the utilization of Natural Language Processing (NLP) techniques and machine learning algorithms through the Orange platform. A total of 3,615 reviews were analyzed, consisting of 80% training data and 20% testing data. The research process was comprised of four primary stages: data collection and labeling, text preprocessing, feature transformation using TF-IDF, and testing five classification algorithms: The following machine learning algorithms were considered: neural networks, naïve Bayes, random forests, decision trees, and k-nearest neighbors. The findings indicated that the Neural Network model demonstrated optimal performance, with an accuracy of 93.20%, an F1-score of 93.00%, and an MCC of 75.80%. These findings suggest that NLP approaches can be effectively utilized to comprehend user perceptions of the app. This research is expected to assist Gojek app developers in improving service quality based on user feedback.

Keywords: *Gojek, Sentiment Analysis, NLP, Orange*

1. Introduction

Gojek, a pioneering entity in the realm of on-demand services within Indonesia, has undergone a rapid phase of expansion since its inception. The company's comprehensive suite of services, encompassing transportation, food delivery, and digital payments, has enabled Gojek to amass a user base that exceeds millions (Rahman et al., 2024). Gojek is an online transportation service platform that is headquartered in Indonesia. According to data from Katadata (2019), the Gojek application has been downloaded over 142 million times, making it one of the most popular public service applications in the country. In the realm of food delivery, Gojek has established collaborative relationships with approximately 400,000 merchants, spanning over 370 cities across Indonesia. However, an analysis of positive sentiment reviews for Gojek reveals that it is the lowest compared to similar apps like Grab and Maxim, at only around 20.21%. This finding suggests that user perceptions of these applications vary, with each application possessing its own unique strengths.

As the number of users increases, a variety of reviews and responses emerge on platforms such as Google Playstore, reflecting the diverse experiences and perceptions of users regarding the services provided. Given the wide range of services and the highly diverse user base, it is to be expected that each individual customer would have a different experience when using Gojek services. This phenomenon gives rise to a myriad of responses, including praise, criticism, and suggestions, which are reflected in user reviews on digital platforms such as the Google Play Store. The Google Play Store user reviews constitute a substantial data set, offering insights into

user satisfaction, grievances, and expectations concerning the Gojek application. A thorough analysis of these reviews can yield profound insights into aspects of the service that require enhancement or preservation. However, the substantial volume and heterogeneity of the review data necessitate an efficient and accurate analytical approach. A paucity of studies has been conducted on the application of NLP to the Gojek platform through the lens of model comparison. In such cases, the role of Natural Language Processing (NLP) becomes crucial. Natural Language Processing facilitates the processing and analysis of large-scale text data, thereby enabling the identification of patterns, sentiments, and topics that emerge in user reviews. By employing techniques such as sentiment analysis and topic modeling, researchers can categorize reviews based on their polarity (positive or negative) and identify the aspects of the service that are frequently discussed by users.

The author will employ data analysis platforms such as Orange to facilitate the analysis of data. Orange is an open-source software program that provides a visual interface for data analysis and machine learning. It allows users to build analysis workflows without the need to write extensive code. The incorporation of diverse widgets for data processing, feature extraction, and model training enables Orange to facilitate the sentiment analysis process, ensuring its efficient execution. As demonstrated in the study by Hasmadi et al. (2024), the integration of the Naïve Bayes algorithm with the Orange platform attained an accuracy of 87.4% in evaluating the sentiment towards the quality of Gojek driver services. Furthermore, a range of machine learning algorithms have been employed in the sentiment analysis of Gojek application reviews. Contrary to the approaches of preceding studies, the present research utilizes balanced data and the Orange visual platform, a system that enables system replication by individuals lacking programming expertise. Classification of reviews based on sentiment has been achieved through the utilization of various machine learning methodologies, including Naïve Bayes, K-Nearest Neighbors (KNN), Neural Network, Decision Tree, and Random Forest algorithms.

In this context, it is imperative to categorize these reviews into classifications that represent users' attitudes or assessments of the service, such as positive and negative sentiments. This classification system serves as a foundation for comprehending the extent to which users experience satisfaction or dissatisfaction with the various features of the application. The application of labels based on sentiment polarity facilitates the interpretation of substantial data sets, thereby enabling companies to expeditiously identify prevailing issues and aspects of the service that are appreciated. That is to say, sentiment labels function not only as a classification tool but also as a direct reflection of diverse user experiences. Consequently, acquiring a profound comprehension of these sentiment patterns constitutes a strategic imperative for ensuring service quality and for responding to user needs in a more adaptive manner.

The objective of this study is to assess the efficacy of machine learning algorithms in the classification of user sentiment toward the Gojek application, as evidenced by reviews from the Google Play Store.

2. Method

The objective of this study is to ascertain the sentiment expressed in reviews of the Gojek mobile application on the Google Play store. The methodology employed involves the utilization of Natural Language Processing (NLP) techniques, processed through Orange software version 3.39.0 using Windows 11 Home Single Language Edition with 4.00 GB of RAM and an AMD Ryzen 3 3250U processor with Radeon Graphics. The data utilized in this study was obtained from Kaggle.com. To ensure the analysis is systematic and can produce accurate sentiment predictions, the process is carried out through four main stages: data collection and label conversion, text preprocessing, feature extraction using TF-IDF, and testing machine learning algorithms, as shown in Figures 1. The subsequent section delineates the progression of each stage.

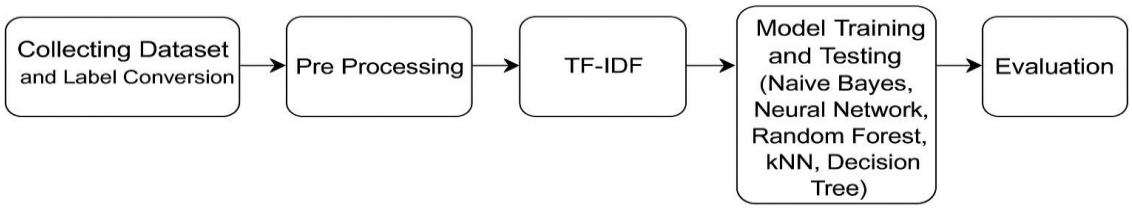

Figure 1. Machine learning modelling

2.1 Collecting Dataset and Label Conversion

The initial phase of this research entailed the aggregation of data in the form of user reviews of the Gojek application from Google Playstore in June 2025, which was obtained from Kaggle. The data set under consideration consists of 3,615 reviews, which were subsequently divided into two distinct categories: training data comprising 2,892 reviews and testing data comprising 723 reviews, with an 80:20 ratio. Each review is assigned a rating from 1 to 5. For the purpose of classification, a label conversion method was employed. In this method, ratings 1 and 2 were designated as negative sentiment (label 0), while ratings 4 and 5 were classified as positive sentiment (label 1). The neutral category, designated as "3," was eliminated and excluded from the dataset to enhance the classification's focus. A thorough examination of the aggregate testing data reveals that 362 reviews are classified as positive, while 361 reviews are designated as negative, suggesting a balanced distribution of responses. The converted review data according to the ratings can be seen in Table 1.

Table 1. Testing data

Review	Rank	Converted Label
tiap minggu ada promo diskon, mantap lah, bagus bnget buat mesen makan, ojek atau apapun itu	5	1
mempermudah untuk transportasi, makan, bahkan GO-SEND bintang 5	5	1
saya senang dengan aplikasi gojek yg sangat membantu perjalanan saya	5	1
Aplikasi Gila. Harga penanganan dan pengiriman lebih mahal daripada harga makanannya. Padahal jarak gak jauh	1	0
sangat kecewa, tidak bisa cancel, driver lama penjemputan, ongkir semakin mahal, CS menangani masalah dengan sangat buruk untuk sebuah masalah yang simpel	1	0

2.2 Pre Processing

Subsequent to the collection and labeling of the data, the subsequent step is to pre-process the review text. This stage is pivotal for the elimination of extraneous elements and the standardization of the data structure, thereby ensuring optimal processing by the algorithm. The process is comprised of three primary components: tokenization, case transformation, and stopword filtering. The following explanations are provided below:

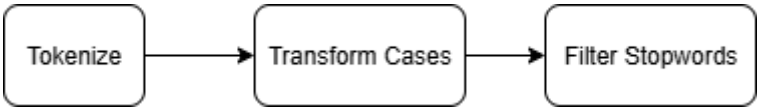

Figure 2. Pre-processing stages

2.2.1 Tokenize

Tokenization is defined as the process of dividing text or sentences into smaller units called tokens (Rimba, 2019). These tokens are defined as words, short phrases, or symbols that are regarded as meaningful units in the context of text analysis. For instance, the user review sentence "sangat membantu sekali dan tetap harus ditingkatkan lagi" will be segmented into several tokens as follows: The following words were utilized: 'sangat', 'membantu', 'sekali', 'dan', 'tetap', 'harus', 'ditingkatkan', 'lagi'. This process enables the analysis of each word on an individual basis, considering both its frequency and its association with a specific sentiment label.

2.2.2 Transform Cases

At this stage, all letters in the text are converted to lowercase (Cahyaningtyas & Widyasari, 2021). The objective is to maintain uniformity in word processing, as in raw text form, the same word can be written with uppercase or lowercase letters and be interpreted differently by the computer. For instance, the words "Sangat" and "sangat" would technically be recognized as two different entities if no transformation were performed.

2.2.3 Filter Stopword

The stopword filter is a linguistic tool that can remove common words that often appear in sentences but do not contribute to the specific meaning of the text in the context of sentiment analysis (Wardhani et al., 2024). Examples of such words include “dan”, “yang”, “itu”, “harus”, “akan”, and the like. It is widely accepted that these words play a negligible role in the identification of the polarity of a review.

2.3 TF-IDF

Subsequent to the data cleaning process, each review was converted into a numerical representation using the TF-IDF (Term Frequency–Inverse Document Frequency) method. This method is employed to calculate the importance weight of a word in a document relative to all documents in the dataset (Istiqamah and Rijal, 2024). Words that are frequently present in one review but infrequent in other reviews will have a high TF-IDF value. These words are considered more meaningful by the algorithm. This representation enables the system to identify and evaluate keywords that contribute to positive or negative sentiment. This calculation can be performed using formula (1), where *t* represents a word (term) and TF indicates the frequency of occurrence of *t* in a document. Concurrently, *N* signifies the total number of documents, and *df* denotes the number of documents containing term *t*. The Inverse Document Frequency (IDF) formula is employed to balance the weight of a word based on its frequency in the corpus (Husain et al., 2024), which can be elucidated as follows:

$$\text{TF-IDF}\ (t) = TF * \log \frac{N}{df} \quad (1)$$

2.4 Model Testing and Training

The final step in this process is to test five machine learning models to determine which model is most effective in sentiment classification. The models employed in this study include neural networks, naïve Bayes, random forests, decision trees, and k-nearest neighbors (kNN). The training process utilizes training data, and subsequently, the model is evaluated using the prepared testing data.

2.4.1 Naïve Bayes

Naïve Bayes is a probabilistic classification method based on Bayes' theorem, assuming that each feature in the data is independent of one another. This approach is considered effective for text data processing because it can quickly classify labeled documents even with limited training data (Wahyuni, 2022). In the context of sentiment analysis, Naïve Bayes can work by predicting the probability of a category based on word patterns that appear in the text, making it a relevant method for analyzing user reviews on platforms such as the Google Play Store (Agustina, 2023).

2.4.2 Neural Network

In the domain of Natural Language Processing (NLP), Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been employed for text classification for an extended period (Tatchum et al., 2024). A subfield of machine learning that utilizes artificial neural networks with multiple layers (deep neural networks) to automatically learn data representations is also referred to as deep learning (Sarker, 2021). Neural network models and deep learning algorithms have demonstrated the capacity to extract complex features directly from raw data through intensive training processes. This capacity renders neural network algorithms efficacious in the management of voluminous and unstructured data, a salient example being the review texts of the Gojek application. In the context of sentiment analysis, Neural Networks facilitate the modeling of intricate relationships inherent in textual data, thereby facilitating the identification of patterns that accurately reflect user opinions or emotions.

2.4.3 Random Forest

The Random Forest algorithm is a machine learning algorithm that is characterized by its ability to produce multiple decision trees. Random Forest is a machine learning method that constructs multiple decision trees during the training phase. It then generates class predictions based on the majority vote of these trees. Decision trees are constructed by first identifying the root node and subsequently terminating with multiple leaf nodes to yield the desired outcome, leveraging training data (Alita and Isnain, 2020). This algorithm's primary strength lies in its capacity to minimize overfitting, thereby enhancing the accuracy of predictions (Maulita and Wahid, 2024). In the domain of sentiment analysis, Random Forest has demonstrated efficacy in managing substantial and intricate datasets, a notable example being the Gojek app review data.

2.4.4 K-Nearest Neighbors (KNN)

KNN is a non-parametric classification algorithm that determines the class of a data point based on the majority class of its k nearest neighbors in the feature space. In the context of sentiment analysis of Gojek application reviews, KNN has been employed to identify sentiment patterns based on the similarity between reviews (Muttaqin and Kharisudin, 2021). The KNN algorithm was selected due to its capacity to calculate the proximity between new cases (test data) and old cases (training data) based on the matching weights of a number of existing features. Subsequent to the calculation of distance, the distance to the training data that is closest is considered to have similarity (Rahayu et al., 2022).

2.4.5 Decision Tree

The Decision Tree algorithm is a prediction model that utilizes a tree structure to identify and resolve problems by evaluating various factors within the problem's scope (Asshiddiqi and Lhaksmana, 2020). This model is frequently employed in machine learning contexts, wherein a diagram commences with a single node that subsequently bifurcates into multiple branches, with each branch in turn giving rise to another branch (Syafrianto, 2022).

2.5 Evaluation

The results of the data analysis conducted using the aforementioned five approaches were compiled into a confusion matrix, which included True Positive (TP) and True Negative (TN); False Positive (FP) and False Negative (FN). Subsequent to the collection of the data, a thorough analysis was conducted to ensure the accuracy and precision of the results. Subsequently, accuracy values were generated, including precision, recall, and F1-Score, to evaluate the performance of the machine learning model. The accuracy values were derived using the following formula (2):

According to Fahmi and Pratiwi (2024), the accuracy results are measured by the ratio of correct predictions to the total number of correct and accurate predictions.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

Precision is defined as the ratio of correct positive predictions (true positives) to the total number of positive predictions (true positives + false positives).

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Recall is defined as the ratio of correct positive predictions (true positives) to the total number of correct positive data (true positives + false negatives).

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

The F1-Score is a metric that provides an overview of the harmonic mean of precision and recall, offering a comprehensive evaluation of the equilibrium between these two crucial metrics.

$$F1 = \frac{2\ x\ Recall\ x\ Precision}{Recall + Precision} \quad (5)$$

The Matthews Correlation Coefficient (MCC) is a metric used to evaluate the performance of binary classification models.

$$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

3. Results And Discussion

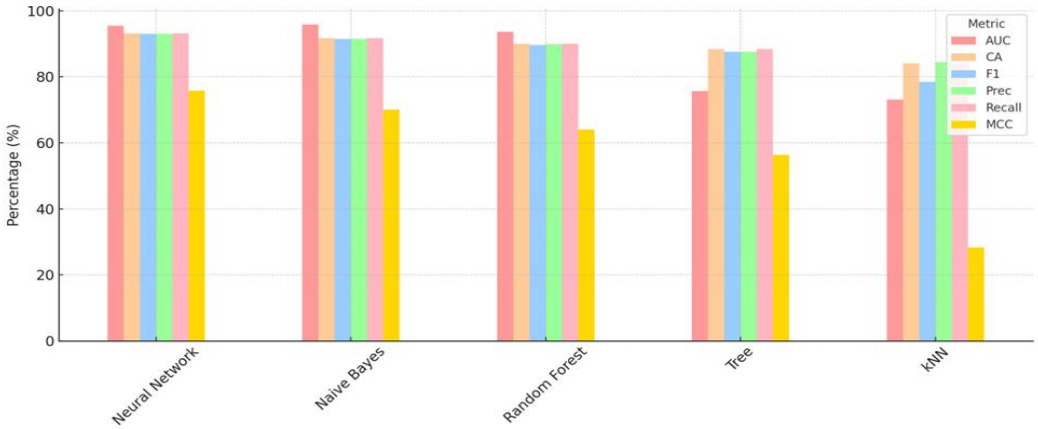

Figure 3. Performance comparison of five classification models

Table 2. Evaluation Matrix

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	95,50%	93,20%	93,00%	93,00%	93,20%	75,80%
Naive Bayes	95,80%	91,70%	91,40%	91,40%	91,70%	70,10%
Random Forest	93,60%	90,04%	89,60%	89,90%	90,04%	64,00%
Tree	75,70%	88,40%	87,60%	87,60%	88,40%	56,30%
kNN	73,10%	84,00%	78,50%	84,40%	84,00%	28,30%

The Confusion Matrix table above illustrates the performance of five machine learning algorithms in classifying user sentiment reviews of the Gojek application based on the metrics Area Under the Curve (AUC), accuracy, precision, recall, F1-score, and Matthew's Correlation Coefficient (MCC). The Neural Network model exhibited the optimal overall performance, with an accuracy of 93.20%, precision of 93.00%, and recall of 93.20%. This indicates that the model effectively identified the majority of reviews, demonstrating an equitable balance between detecting positive and negative reviews. The F1-score of 93.00% substantiates these findings, underscoring the model's precision and reliability in prediction. The MCC of 75.80% serves as a robust indicator of the model's classification system's significant alignment with the true labels.

Concurrently, the Naïve Bayes model demonstrates a high degree of similarity in performance with the Neural Network, exhibiting an accuracy of 91.70%, a precision of 91.40%, and a recall of 91.70%. This suggests that over 91% of the positive review predictions made by Naïve Bayes accurately reflect the actual positive sentiment. This precision figure is of particular importance in the context of app reviews, as it prevents misclassification errors that could mislead developers. For example, negative or neutral reviews may be mistakenly identified as positive reviews (Fahmi and Pratiwi, 2024). This model operates with a high degree of precision, thereby reducing the probability of false positives and enhancing the confidence that reviews categorized as positive genuinely reflect positive user experiences.

Subsequent to this, Random Forest demonstrated an accuracy of 90.04% and a precision of 89.90%, suggesting its capacity to accurately predict sentiment. However, the model's MCC of 64.00% indicates that its consistency falls short of the performance standards set by the previous

two models. The Decision Tree model exhibits marginally diminished performance metrics, with an accuracy of 88.40%, an F1-score of 87.60%, and an MCC of 56.30%. While these results remain positive, they suggest that Decision Tree may not possess the same degree of reliability as previous models in accurately mapping the relationship between features and labels. The k-Nearest Neighbor (kNN) model demonstrated the least efficacy among the five models evaluated, exhibiting an accuracy of 84.00%, an F1-score of 78.50%, and the lowest Micro-Calibrated (MCC) of 28.30%. These findings suggest that this model exhibits reduced effectiveness in consistently detecting sentiment. Despite its high precision rate of 84.40%, its capacity to capture the broader patterns in the data is comparatively deficient, as evidenced by its lower recall and F1-score metrics. The k-Nearest Neighbor (kNN) model demonstrates that it does not necessitate extensive model training processes, as is characteristic of Neural Networks. This characteristic renders it suitable for lightweight systems, as it merely stores data samples, resulting in a relatively lightweight internal structure (Liu G. et al., 2022).

Overall, the application of neural networks in the sentiment analysis of Gojek application user reviews has been demonstrated to be the most optimal approach. The efficacy of neural networks has been demonstrated to exceed that of other algorithms, a finding that aligns with the observations reported by Sarker (2021). Furthermore, Mienye & Swart (2024) have demonstrated that neural network models exhibit superior outcomes due to their capacity to discern intricate patterns. However, these results differ from the study by Muttaqin & Kharisudin (2021), which identified SVM as the most accurate algorithm. This discrepancy can be attributed to the utilization of data balancing methods and analogous features. This model facilitates enhanced user response comprehension for developers, thereby mitigating the risk of decision-making informed by misclassified data. In the context of developing service-based applications like Gojek, the selection of an algorithm with high precision and recall is imperative to ensure that the user experience and expectations are authentically reflected in data-driven product development strategies.

4. Conclusions

The findings of the research endeavor indicate that the implementation of Natural Language Processing (NLP) through the Orange platform has demonstrated efficacy in the identification of sentiment in user reviews of the Gojek application. Among the five machine learning algorithms examined, the Neural Network model demonstrated the most optimal performance, exhibiting an accuracy rate of 93.20%. Additionally, it exhibited elevated and balanced precision and recall values in the classification of positive and negative reviews. In practice, this research makes a significant contribution to the development of an automatic sentiment analysis system. This system can be applied by application developers to efficiently understand user responses. These findings can also be utilized in strategic decision-making related to improving application services based on user review data.

However, this study is not without its limitations. Firstly, the data utilized in this study is derived exclusively from a single platform, the Google Play Store. Secondly, the data is from June 2025. Thirdly, the study employs classical algorithms that do not yet incorporate deep learning-based models. Consequently, further research is recommended to employ multi-platform datasets to enhance data representation, as well as to integrate more advanced natural language processing (NLP) models, such as Bidirectional Encoder Representations from Transformers (BERT) or Long Short-Term Memory (LSTM), to obtain more accurate and contextual classification results. In the future, this system can be developed into an automated dashboard that will help application developers monitor user satisfaction in real time.

Acknowledgements

We would like to express our gratitude to the individuals who have contributed to the development of this scientific article, including Mr. Hendro Margono, S. Sos., M. Sc., Ph.D., who

served as the lecturer for the Data Science course. His contributions, including the provision of material and knowledge, have been instrumental in the development of this article.

References

Agustina, N., Citra, D. H., Purnama, W., Nisa, C., & Kurnia, A. R. (2022). Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store: The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews On Google Play Store. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 2(1), 47-54.

Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, *8*(2), 50-58.

Asshiddiqi, M. F., & Lhaksmana, K. M. (2020). Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI. *eProceedings of Engineering*, *7*(3).

Cahyaningtyas, C., Nataliani, Y., & Widiasari, I. R. (2021). Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE. *Aiti*, *18*(2), 173-184.

Fahmi, F., & Pratiwi, A. M. (2024). Identifying Sentiment in User Reviews of Get Contact Application using Natural Language Processing. *Proceedings - 2024 International of Seminar on Application for Technology of Information and Communication: Smart And Emerging Technology for a Better Life, ISemantic 2024*, 428–433. https://doi.org/10.1109/iSemantic63362.2024.10762453

Fitra Ramadhan, F. (2021). Chatbot pada E-Commerce berbasis Android dengan Pendekatan Natural Language Processing. *JCSE Journal of Computer Science an Engineering*, *2*(1), 27–39. http://icsejournal.com/index.php/

Husain, N. P., Sukirman, S., & SAJIAH, S. (2024). Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine. *Journal of System and Computer Engineering (JSCE)*, *5*(1), 91–102. https://doi.org/10.61628/jsce.v5i1.1105

Istiqamah, N., & Rijal, M. (2024). Klasifikasi Ulasan Konsumen Menggunakan Random Forest dan SMOTE. *Journal of System and Computer Engineering (JSCE)*, *5*(1), 66–77. https://doi.org/10.61628/jsce.v5i1.1061

Katadata. 2019. Persaingan Ketat Gojek dan Grab Menjadi SuperApp. https://katadata.co.id/telaah/2019/04/16/persaingan-ketat-gojek-dan-grab-menjadi-superapp.

Katadata. 2025. Gojek, Grab, dan Maxim, Mana yang Punya Sentimen Paling Positif? https://databoks.katadata.co.id/transportasi-logistik/statistik/682c48bcd8785/gojek-grab-dan-maxim-mana-yang-punya-sentimen-paling-positif

Liu, G., Zhao, H., Fan, F., Liu, G., Xu, Q., & Nazir, S. (2022). An Enhanced Intrusion Detection Model Based on Improved kNN in WSNs. Sensors, 22(4), 1407. https://doi.org/10.3390/s22041407

Maulita, I., & Wahid, A. (2024). Prediksi Magnitudo Gempa Menggunakan Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron Berdasarkan Data Kedalaman dan Geolokasi (Predicting Earthquake Magnitude Using Random Forest, Support Vector Regression, XGBoost, LightGBM, and Multi-Layer Perceptron Based on Depth and Geolocation Data). *Jurnal Pendidikan Dan Teknologi Indonesia*, *4*, 221-232.

Mienye, I. D., & Swart, T. G. (2024). A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications. Information, 15(12), 755. https://doi.org/10.3390/info15120755

Muttaqin, M. N., Kharisudin, I. 2021. Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor. UNNES Journal of Mathematics. 10(2):22-27.

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 J. Pendidik. Inform*, *6*(1), 98-106.

Rahman, R. A., Pranatawijaya, V. H., & Sari, N. N. K. (2023). Analisis Sentimen Berbasis Aspek pada Ulasan Aplikasi Gojek. KONSTELASI: Konvergensi Teknologi dan Sistem Informasi, 4(1). https://doi.org/10.24002/konstelasi.v4i1.8922

Rimba, Chory Nuzulul. (2019). Analisis Sentimen pada tingkat kepuasan pengguna layanan data seluler menggunakan algoritma Support Vector Machine(SVM).

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, *2*(6), 1-20.

Syafrianto, A. (2022). Perbandingan Algoritma Naïve Bayes dan Decision Tree Pada Sentimen Analisis. *The Indonesian Journal of Computer Science Research*, *1*(2).

Tatchum, G. W., Nzeko'o, A. J. N., Makembe, F. S., & Djam, X. Y. (2024). Class-Oriented Text Vectorization for Text Classification: Case Study of Job Offer Classification. *Journal of Computer Science and Engineering (JCSE)*, *5*(2), 116–136.

Wahyuni, W. (2022). Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naive Bayes. Jurnal Informatika Ekonomi Bisnis, 148-153.

Wardhani, D., Astuti, R., & Saputra, D. D. (2024). Optimasi Feature Selection Text Mining: Stemming Dan Stopword Untuk Sentimen Analisis Aplikasi SatuSehat. *Innovative: Journal Of Social Science Research*, *4*(1), 7537-7548.