e-ISSN 2549-7472
Edumatic: Jurnal Pendidikan Informatika
Vol. 9 No. 2, Agustus, 2025, **Hal. 532-541**
DOI: 10.29408/edumatic.v9i2.30714

Performance Evaluation of Naïve Bayes and SVM in Sentiment Analysis of Illegal Parking Attendants

Sandra Saputra [1,*] , Paradise [1] , Novanda Alim Setya Nugraha [1]

[1] Department of Informatics, Telkom University Purwokerto, Indonesia
[*] Correspondence: 21102235@ittelkom-pwt.ac.id

Received: 9 June 2025 | Revised: 19 June 2025 | Accepted: 11 July 2025 | Published: 15 August 2025

Abstract
The increase in the number of vehicles in Indonesia has led to high demand for parking spaces, which has triggered the emergence of illegal parking attendants. This phenomenon has elicited various public responses, particularly on social media platform X. This study analyzes public sentiment toward the presence of illegal parking attendants by comparing the performance of the Naïve Bayes and Support Vector Machine (SVM) algorithms. The data used consists of 1,484 Indonesian-language tweets collected via crawling techniques. The pre-processing stage included data cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. The data was then labeled with positive or negative sentiment using the InSet (Indonesia Sentiment Lexicon) approach and manually validated, before being divided into training and testing datasets. Feature extraction was performed using the TF-IDF method before being applied to the classification model. The evaluation results show that the SVM algorithm with a linear kernel approach produces the highest accuracy of 82%, outperforming Naïve Bayes: Gaussian 56%, Multinomial 74%, and Bernoulli 77%. These results are expected to contribute to the formulation of more organized and transparent parking policies, as well as demonstrate the importance of sentiment analysis as a tool to support data-driven decision making.

Keyword: illegal parking attendants; naïve bayes; social media x; svm; text classification

INTRODUCTION

Based on data from cnnindonesia.com, the growth in the number of motor vehicles in Indonesia, which reached 164 million units by August 2024, shows a significant increase, thereby driving up the demand for parking spaces to accommodate these vehicles (Intan et al., 2023). This increase has led to the emergence of parking attendants who are tasked with improving security and assisting consumers in managing their vehicle parking (Oszaer et al., 2023). According to Minister of Transportation Regulation No. 60 of 2021, anyone who wishes to work as a parking attendant must obtain an official permit issued by the local government. However, there are many individuals who engage in illegal activities, often found in shopping areas or even in places marked as free parking, causing public inconvenience (Wijaya et al., 2022). Poverty, economic pressure, and high unemployment rates are some of the main factors contributing to the emergence of illegal parking attendants in many places (Sedenel et al., 2022). Based on Presidential Regulation No. 112 of 2007, every commercial store is required to provide parking space, which is certainly very useful in preventing traffic congestion caused by shoppers or visitors (Mukti & Hasan, 2024). The presence of illegal parking attendants in various places has sparked various opinions among the public, making this topic widely discussed on social media (Verawati & Jaelani, 2024). This situation has led to social media, particularly the X platform, becoming a space where people often voice their experiences and opinions regarding the presence of illegal parking attendants.

Social media has become a platform for people to express their views and feelings on various topics openly without any restrictions (Nugraha et al., 2022). Especially on platform X, people often share stories about their daily activities and personal views, which serve as a means of interaction and expression in the public sphere (Safira & Hasan, 2023). Experiences related to commercial activities, especially parking, are often shared by users on the X platform, both positive and negative (Asro'i & Februariyanti, 2022). Many X social media users send tweets after interacting with illegal parking attendants, generating a variety of tweets on the social media platform. In such a situation, machine learning-based sentiment analysis becomes very important to systematically process public opinion, so that it can be used in effective sentiment classification modeling with algorithms such as Naïve Bayes and SVM.

Machine learning is a branch of artificial intelligence that focuses on developing algorithms that are able to learn patterns from data to make predictions or make decisions automatically (Husen et al., 2023). One of the widely used applications of machine learning is sentiment analysis, which aims to identify and classify emotions or opinions in text, either in the form of negative, neutral, or positive (Kholilullah et al., 2024). The purpose of this process is to gain a deeper understanding of public opinion or attitudes towards a particular issue, product or entity. This process involves pre-processing, weighting, and classification (Sanjaya et al., 2023). Using such techniques in this research is particularly relevant because it allows researchers to evaluate the public's views on the existence of illegal parking attendants through opinion data taken from X social media platform. This research uses the Naïve Bayes algorithm with three approaches, namely Gaussian, Multinomial, and Bernoulli, and the SVM algorithm with a linear kernel. Both algorithms are applied to classify data into positive and negative sentiment categories, and then compare their accuracy in recognizing sentiment patterns optimally.

One of the most commonly used classification algorithms in supervised learning is Naïve Bayes and SVM. Naïve Bayes is a probabilistic-based classification algorithm that applies the principle of Bayes' Theorem with the assumption of independence between features in the dataset (Setiawan & Suryono, 2024; Sudrajat, 2022), and has the advantage of handling large data through an efficient and lightweight computational process (Furqan & Nasir, 2024; Watratan et al., 2020; Zharifa & Ujianto, 2024). Meanwhile, SVM is a classification algorithm that works by finding the optimal hyperplane to separate data into different classes and is known to be effective in handling high-dimensional data, such as TF-IDF extracted text (Iskandar & Nataliani, 2021). Based on a number of previous studies showing varying results on the performance of the two algorithms, Naïve Bayes obtained 72% accuracy, exceeding SVM which only reached 66% (Samuel et al., 2023), while the Gaussian approach obtained an average accuracy of 97.48% on each data used (Ramadhan & Adhinata, 2022), Multinomial 98.89% (Yusran et al., 2024), and Bernoulli 93.06% (Raffi et al., 2023). Meanwhile, SVM in text classification provides better accuracy of 87.95% compared to Naïve Bayes of 81.65% (Sarimole & Kudrat, 2024), even reaching 75.28% when combined with a linear kernel and TF-IDF and Chi-Square features (Sutranggono & Imah, 2023).

This research aims to analyze public sentiment toward the presence of illegal parking attendants and evaluate the most accurate performance of the Naïve Bayes and SVM algorithms from social media X. By applying the TF-IDF method for feature extraction and testing several classification models. The results of this comparative analysis are expected to serve as a reference for selecting the appropriate classification model and provide data-driven insights for policymakers to develop a more systematic and transparent parking system.

METHOD

This research analyzes the sentiment of public tweets about illegal parking attendants on social media X using several Naïve Bayes algorithm approaches (Gaussian, Multinomial, and

Bernoulli) and SVM with a linear kernel through several stages. The initial stage of the research is data collection, which involves crawling Indonesian-language tweets from social media X users from 2024-07-30 to 2025-05-13. This process utilizes the tweet-harvest library to extract relevant data based on specific keywords “juru parkir”, “tukang parkir”, and “parkir liar”. The second stage is an important step in text data processing, namely pre-processing, which is the process of cleaning data so that it is more ready for use. The pre-processing stage is usually carried out through several steps, such as cleaning (removing symbols, URLs, emojis, and numbers), case folding (changing text to lowercase), word normalization (changing text to standard words), tokenization (breaking sentences into smaller words), stopwords removal (removing common words), this process is imported from the NLTK library, and the final stage is stemming (simplifying words to their base form), this process uses the Sastrawi library.

The next stage is data labeling done with the (Indonesia Sentiment Lexicon (InSet) lexicon dictionary-based approach, where each word or phrase in a sentence is analyzed to determine its sentiment value, whether positive or negative. From previous research, better results were obtained by dividing the training data by 80% which was used to build the model, and test data by 20% which was used to evaluate the quality of the model that had been made. In this research, the weighting is done using the TF-IDF method, TF-IDF gives higher weights to words that appear frequently in a document, this makes TF-IDF effective in identifying significant keywords and reducing the weight of general words that are less informative, so that the feature representation becomes more meaningful for the classification model. After the model of each method is developed, the next step is to test the performance of the model. The test is conducted using confusion matrix to calculate the accuracy, precision, recall, and F1-Score values.

RESULT AND DISCUSSION

Result

The dataset was retrieved using the crawling technique on X application with a total of 1620 data. The data retrieved uses the Indonesian language with the keywords “juru parkir”, “tukang parkir”, and “parkir liar”, with the time span of data collection starting from July 30, 2024 to May 13, 2025. The main columns taken from crawling data are tweet id, tweet text, created at, username, and language. Crawling data utilizes the python programming language, namely using tweet-harvest which is connected to the authentication token in the X application.

Table 1 shows the results of the pre-processing stages. Before the data labeling stage was carried out to remove duplicate data, the results obtained were 1,484 data. At this stage, researchers label the sentiment on the dataset using the InSet (Indonesia Sentiment Lexicon) lexicon-based approach. The labeled data is the result of the pre-processing process, especially the stemming column which contains the cleaned review text. This research utilizes a lexicon dictionary containing a list of positive and negative words to calculate the number of occurrences of each in each review. The sentiment score is calculated based on the difference between the number of positive and negative words, which then becomes the basis for determining the sentiment category. Reviews will be classified as positive reviews if the score is > 0, and as negative reviews if the score is < 0. From the data labeling results, 56.81% were classified as negative reviews with a total of 843 tweets, while positive reviews accounted for 43.19% with a total of 641 tweets.

The next step is to split the dataset in order to separate the data that will be used in the model training process. This study splits the dataset into two parts, namely 80% as training data and 20% as test data. The process of splitting the dataset into training data and test data uses the “train_test_split” function available in the scikit-learn library. Before the data splitting process, data with empty values in the “steming_data” column is removed to ensure that only clean data is used. The ‘steming_data’ column serves as a feature (x), while the “Sentiment” column is used as a label (y). The division is performed with a ratio of 80% for the training data (1,187 data) and 20% for the test data (297 data). The use of the parameter “random_state=42” aims to ensure consistent results each time the process is run, while “stratify=y” is used to maintain the balance of label distribution in both parts of the data.

Table 1. Pre-processing

Steps	Result
Full Text	Viral di medsos nih ada minimarket yang beri imbauan untuk tak memberikan uang ke juru parkir liar. #detikfinance #JuruParkirLiar #Sindir #JanganBayar https://t.co/8gd0Kj9N8N
Cleaning	Viral di medsos nih ada minimarket yang beri imbauan untuk tak memberikan uang ke juru parkir liar.
Case Folding	viral di medsos nih ada minimarket yang beri imbauan untuk tak memberikan uang ke juru parkir liar.
Normalization	viral di media sosial ini ada minimarket yang beri imbauan untuk tidak memberikan uang ke juru parkir liar.
Tokenization	['viral', 'di', 'media', 'sosial', 'ini', 'ada', 'minimarket', 'yang', 'beri', 'imbauan', 'untuk', 'tidak', 'memberikan', 'uang', 'ke', 'juru', 'parkir', 'liar']
Stopwords Removal	['viral', 'media', 'sosial', 'minimarket', 'beri', 'imbauan', 'tidak', 'memberikan', 'uang', 'juru', 'parkir', 'liar']
Stemming	['viral', 'media', 'sosial', 'minimarket', 'beri', 'imbau', 'tidak', 'beri', 'uang', 'juru', 'parkir', 'liar']

After the data splitting process, feature extraction was performed using the TF-IDF method with TfidfVectorizer from the scikit-learn library. The training data was processed using the fit_transform function, while the test data was transformed to match the training data schema. The extraction results showed that the training data had 1,187 rows and 3,884 columns, while the test data had 297 rows and 3,884 columns, forming a sparse matrix due to many zero values. Next, the classification model was tested using the Naïve Bayes algorithm (Gaussian, Multinomial, Bernoulli) and SVM with a linear kernel. Model performance was evaluated using a confusion matrix.

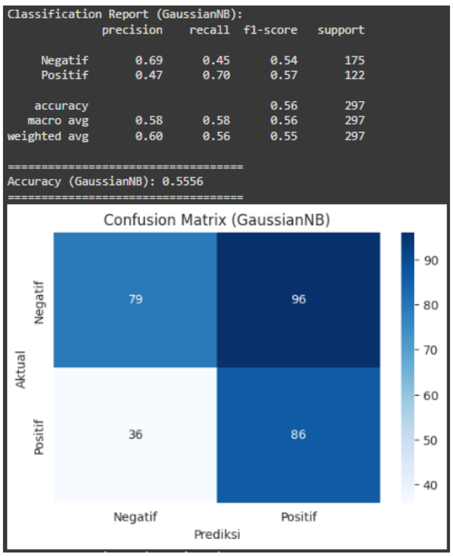

Figure 1. Confusion matrix gaussian naïve bayes

Based on Figure 1, the model with 56% accuracy is not yet able to classify data optimally. The confusion matrix shows that 86 positive data and 79 negative data were classified correctly, but there were prediction errors in 96 false positives and 36 false negatives. The precision and recall for the positive class are 0.47 and 0.70, respectively, while for the negative class they are 0.69 and 0.45. This indicates that the model is better at recognizing positive data in terms of recall, but still lacks precision in both classes.

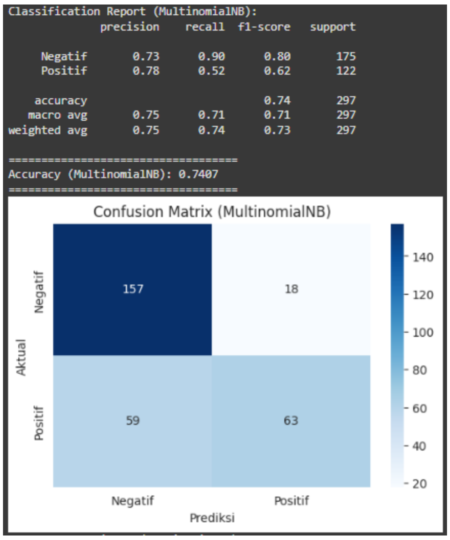

Figure 2. Confusion matrix multinomial naïve bayes

The confusion matrix results in Figure 2 show an accuracy of 74%, indicating that the model performs quite well. The model successfully classified 157 negative data and 63 positive data correctly, but there were still 18 false positives and 59 false negatives. The precision for the negative and positive classes is 0.73 and 0.78, respectively, while the recall is 0.90 for negative and 0.52 for positive. This indicates that the model is more reliable in recognizing negative data, but still less accurate for positive data.

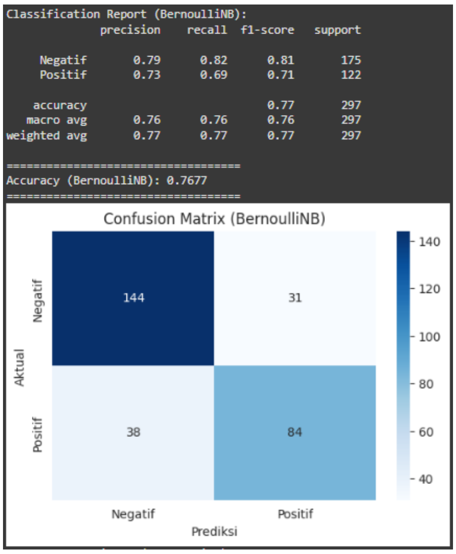

Figure 3. Confusion matrix bernoulli naïve bayes

Figure 3 shows the confusion matrix results of the Bernoulli approach, which achieved an accuracy of 77%. The model successfully classified 144 negative data and 84 positive data correctly, although there were still 31 false positives and 38 false negatives. The precision for the negative class was 0.79 and for the positive class was 0.73, while the recall was 0.82 and 0.69, respectively. These results show that the model is quite stable in distinguishing between the two classes, with better performance in recognizing negative data.

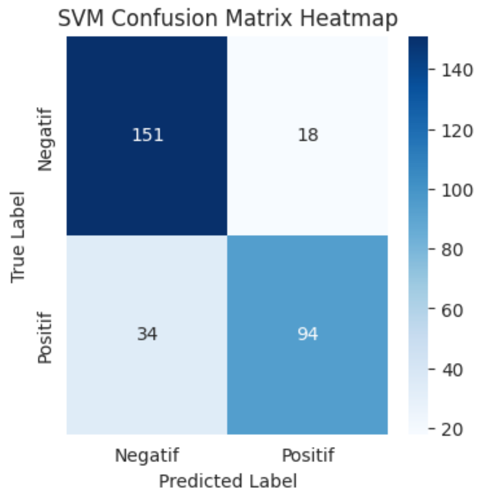

Figure 4. Confusion matrix svm

```
Accuracy: 0.8249158249158249

SVM Accuracy: 82.49%

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.82      0.89      0.85       169
     Positif       0.84      0.73      0.78       128

    accuracy                           0.82       297
   macro avg       0.83      0.81      0.82       297
weighted avg       0.83      0.82      0.82       297
```

Figure 5. Classification report svm

Based on Figures 4 and 5, the accuracy result obtained using linear kernel SVM is 82%. The precision value in the negative and positive classes is 0.82 and 0.84, and recall is 0.89 for negative and 0.73 for positive. A total of 151 negative data (True Negative) and 94 data (True Positive). However, there are still misclassifications of 18 negative data as (False Positive) and 34 positive data predicted as negative (False Negative). The fairly balanced f1-score values in both classes indicate that the SVM model has a consistent performance in recognizing both types of sentiment. These results show that SVM is one of the effective algorithms in the sentiment classification process for the dataset used in this study.

Based on the test results of the two algorithms, namely Naïve Bayes and Support Vector Machine in sentiment classification related to the existence of illegal parking attendants, it can be concluded that there is a significant difference in sentiment classification performance between the two algorithms. This difference emphasizes the importance of selecting the right classification model to ensure accurate interpretation of public opinion obtained from X social media data.

In the Naïve Bayes algorithm, the Gaussian approach showed the lowest performance with an accuracy of 56% and unbalanced precision and recall values, indicating a weakness in recognizing data patterns comprehensively as seen in table 2. Multinomial showed an improvement with an accuracy of 74% and better ability to identify negative classes, although it was still limited to positive classes. Meanwhile, Bernoulli provides the most optimal results among the three with an accuracy of 77% and a more balanced f1-score, reflecting stability in recognizing both sentiment classes. The SVM algorithm with a linear kernel shows the best performance with an accuracy of 82% and a better balance of precision, recall, and f1-score compared to the three Naïve Bayes approaches. The advantage of SVM is due to its suitability for the characteristics of TF-IDF extracted text data, which is high-dimensional and sparse. SVM is able to form an optimal separating hyperplane, while Gaussian is less suitable because it assumes a continuous and normal data distribution. It was concluded that SVM was superior in classifying sentiments related to the presence of illegal parking attendants.

Table 2. Algorithm comparison results

Accuracy		SVM	Gaussian Naïve Bayes	Multinomial Naïve Bayes	Bernoulli Naïve Bayes
		82%	56%	74%	77%
Precision	Negatif	82%	69%	73%	79%
	Positif	84%	47%	78%	73%
Recall	Negatif	89%	45%	90%	82%
	Positif	73%	70%	52%	69%
F1-Score	Negatif	85%	54%	80%	81%
	Positif	78%	57%	62%	71%

Discussion

This study examines public sentiment toward the presence of illegal parking attendants by comparing the performance of the Naïve Bayes and Support Vector Machine algorithms. The data used consisted of 1,484 tweets, which underwent pre-processing such as cleaning, tokenization, and normalization. The results showed that 57% of tweets (843) contained negative sentiment and 43% (641) contained positive sentiment, indicating a predominance of negative views among the public. The classification process used three variants of Naïve Bayes (Gaussian, Multinomial, Bernoulli) as well as SVM with a linear kernel. Evaluation was conducted based on accuracy, precision, recall, and F1-score to assess the effectiveness of each algorithm in classifying sentiment. Based on the evaluation results, it was found that:

Gaussian showed the lowest performance with an accuracy of 56%, indicating that this method is not suitable for sparse and categorical text data, as this algorithm is more suitable for continuous data. In contrast, Multinomial provides better results with an accuracy of 74%, because it calculates feature probabilities based on word frequency, making it more suitable for feature representations such as TF-IDF. However, this algorithm is still inaccurate in recognizing positive data. Bernoulli is the most effective Naïve Bayes approach in this study, with an accuracy of 77% and a balanced f1-score. This algorithm works based on the presence of words, making it suitable for simple sentiment analysis on short texts such as tweets. Additionally, Bernoulli is sensitive to rare but important features in classification. However, this model still experiences prediction errors, with 31 negative data points classified as positive (False Positive) and 38 positive data points classified as negative (False Negative).

The SVM with linear kernel recorded the best performance, with 82% accuracy and good precision-recall balance, demonstrating its effectiveness in processing high-dimensional text data. Compared to three Naïve Bayes models, SVM proved more accurate and efficient, especially in managing imbalanced data by forming optimal decision boundaries. This aligns with findings by Sutranggono & Imah (2023), who reported that SVM combined with TF-IDF and Chi-Square feature selection yielded the highest accuracy in sentiment analysis. However, despite its strong performance, the model still produced classification errors, with 34 positive data points misclassified as negative (False Negative) and 18 negative data points misclassified as positive (False Positive), resulting in a recall rate for positive sentiment of only 73%, indicating that approximately 27% of positive data points were not recognized accurately.

This research supports the findings of Sarimole & Kudrat (2024) that the SVM algorithm has higher accuracy than Naïve Bayes in sentiment analysis. However, unlike the research by Samuel et al. (2023) which mentions the superiority of Naïve Bayes, the results of this study show that SVM is more effective in managing high-dimensional and sparse TF-IDF data. This difference in results is thought to be influenced by the characteristics of the data and the preprocessing methods used. The uniqueness of this study lies in the use of a combination of TF-IDF and adaptive linear SVM kernel for social media data, as well as an in-depth analysis of three variants of Naïve Bayes: Gaussian, Multinomial, and Bernoulli.

The research results showing the dominance of negative sentiment in public tweets reflect public concerns about the existence of illegal parking attendants. These findings can provide valuable input for policymakers to establish a more orderly, transparent, and disruption-free parking system. However, this study has limitations as it only uses data from social media platform X, and therefore does not fully represent public opinion from other platforms or non-social media users.

CONCLUSION

Based on the results of the research that has been carried out, it can be concluded about the sentiment analysis of the existence of illegal parking attendants using Naïve Bayes and SVM algorithms. The evaluation results conclusively show that the SVM algorithm with a linear kernel, which achieved 82% accuracy, is significantly superior to all tested Naïve Bayes approaches (Gaussian 56%, Multinomial 74%, and Bernoulli 77%). In addition, data analysis revealed that public sentiment towards this phenomenon was dominated by negative responses at 57%. The superior performance of SVM has important implications. It indicates that the SVM approach, which works by finding the optimal separating hyperplane, is suitable for managing text data from social media which has high dimensional and sparse characteristics after being processed by TF-IDF. Based on its proven effectiveness and accuracy, this SVM model is a very suitable candidate to be implemented in an automated public opinion monitoring system. The predominance of negative sentiments found in this study is a data-driven justification for local governments and relevant authorities to reorganize the parking system more seriously, with the aim of creating order, transparency, and reducing public unrest.

REFERENCES

Asro'i, A., & Februariyanti, H. (2022). Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor. *Jurnal Khatulistiwa Informatika*, *10*(1), 17–24. https://doi.org/10.31294/jki.v10i1.12624

Furqan, M., & Nasir, A. F. A. (2024). Big data approach to sentiment analysis in machine learning-based microblogs: Perspectives of religious moderation public policy in indonesia. *Journal of Applied Engineering and Technological Science*, *5*(2), 955-965. https://doi.org/10.37385/jaets.v5i2.4498

Husen, R. A., Astuti, R., Marlia, L., Rahmaddeni, R., & Efrizoni, L. (2023). Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *3*(2), 211–218. https://doi.org/10.57152/malcom.v3i2.901

Intan, S. F., Permana, I., Salisah, F. N., Afdal, M., & Muttakin, F. (2023). Perbandingan Algoritma KNN, NBC, dan SVM: Analisis Sentimen Masyarakat Terhadap Perparkiran di Kota Pekanbaru. *JUSIFO (Jurnal Sistem Informasi)*, *9*(2), 85–96. https://doi.org/10.19109/jusifo.v9i2.21357

Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(6), 1120–1126. https://doi.org/10.29207/resti.v5i6.3588

Kholilullah, M., Martanto, M., & Hayati, U. (2024). Analisis Sentimen Pengguna Twitter(X) Tentang Piala Dunia Usia 17 Menggunakan Metode Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 392–398. https://doi.org/10.36040/jati.v8i1.8378

Mukti, A. T., & Hasan, F. N. (2024). Analisis Sentimen Warganet Terhadap Keberadaan Juru Parkir Liar Menggunakan Metode Naive Bayes Classifier. *Jurnal Media Informatika Budidarma*, *8*(1), 644–653. https://doi.org/10.30865/mib.v8i1.6982

Nugraha, S. N., Pebrianto, R., Latif, A., & Firdaus, M. R. (2022). Analisis Sentimen Twitter Terhadap Menteri Indonesia dengan Algoritma Support Vector Machine dan Naive Bayes. *Jurnal Teknik Elektro Dan Informatika*, *17*(1), 1–12. https://doi.org/10.30587/e-link.v17i1.3965

Oszaer, R. J., Nendissa, R. H., & Tita, H. M. Y. (2023). Penegakan Hukum Terhadap Juru Parkir Tidak Resmi Di Kota Ambon. *CAPITAN: Constitutional Law & Administrative Law Review*, *1*(1), 46–63. https://doi.org/10.47268/capitan.v1i1.9907

Raffi, M., Suharso, A., & Maulana, I. (2023). Analisis Sentimen Ulasan Aplikasi Binar Pada Google Play Store Menggunakan Algoritma Naïve Bayes. *INTECOMS: Journal of Information Technology and Computer Science*, *6*(1), 450–462. https://doi.org/10.31539/intecoms.v6i1.6117

Ramadhan, N. G., & Adhinata, F. D. (2022). Sentiment Analysis on Vaccine COVID-19 Using Word Count and Gaussian Naïve Bayes. *Indonesian Journal of Electrical Engineering and Computer Science*, *26*(3), 1765–1772. https://doi.org/10.11591/ijeecs.v26.i3.pp1765-1772

Safira, A., & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *ZONAsi: Jurnal Sistem Informasi*, *5*(1), 59-70. https://doi.org/10.31849/zn.v5i1.12856

Samuel, F. D., Atika, P. D., & Setiawati, S. (2023). Analisis Sentimen Masyarakat Terhadap Perkuliahan Daring di Twitter Menggunakan Algoritma Naive Bayes dan Support Vector Machine. *Journal of Students 'Research in Computer Science*, *4*(2), 261–272. https://doi.org/10.31599/6691v571

Sanjaya, T. P. R., Fauzi, A., & Masruriyah, A. F. N. (2023). Analisis Sentimen Ulasan pada E-commerce Shopee Menggunakan Algoritma Naive Bayes dan Support Vector Machine. *INFOTECH : Jurnal Informatika & Teknologi*, *4*(1), 16–26. https://doi.org/10.37373/infotech.v4i1.422

Sarimole, F. M., & Kudrat. (2024). Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine. *Jurnal Sains Dan Teknologi*, *5*(3), 783–790. https://doi.org/10.55338/saintek.v5i1.2702

Sedenel, A. F., Cheisviyanny, C., & Sari, V. F. (2022). Potensi Pendapatan Retribusi Parkir Dari Sudut Pandang Juru Parkir Liar di Kota Padang Tahun 2021. *Jurnal Eksplorasi Akuntansi (JEA)*, *4*(1), 74–92. https://doi.org/10.24036/jea.v4i1.493

Setiawan, A., & Suryono, R. R. (2024). Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika*, *8*(1), 183-192. https://doi.org/10.29408/edumatic.v8i1.25667

Sudrajat, A., Mulyani, N., & Marpaung, N. (2022). Sistem Pendukung Keputusan Penentuan Kelayakan Penangguhan Kredit Nasabah menggunakan Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika*, *6*(2), 205-214. https://doi.org/10.29408/edumatic.v6i2.6298

Sutranggono, A. N., & Imah, E. M. (2023). Tweets Emotions Analysis of Community Activities Restriction as COVID-19 Policy in Indonesia Using Support Vector Machine.

CommIT (Communication and Information Technology) Journal, *17*(1), 13–25. https://doi.org/10.21512/commit.v17i1.8189

Verawati, I., & Jaelani, S. N. (2024). Analisis Sentimen Pengguna Twitter Terhadap Bus Listrik Menggunakan Naïve Bayes. *Jurnal Media Informatika Budidarma*, *8*(2), 832–842. https://doi.org/10.30865/mib.v8i2.7030

Watratan, A. F., B, A. P., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, *1*(1), 7–14. https://doi.org/10.52158/jacost.v1i1.9

Wijaya, K. A. M., Dewi, A. A. S. L., & Suryani, L. P. (2022). Perijinan dan Tindak Pidana Terhadap Juru Parkir Liar di Kota Denpasar. *Jurnal Analogi Hukum*, *4*(3), 260–265.

Yusran, M., Siswanto, S., & Islamiyati, A. (2024). Comparison of Multinomial Naive Bayes and Bernoulli Naive Bayes on Sentiment Analysis of Kurikulum Merdeka with Query Expansion Ranking. *SISTEMASI: Jurnal Sistem Informasi*, *13*(1), 96–106. https://doi.org/10.32520/stmsi.v13i1.3187

Zharifa, A. H. A., & Ujianto, E. I. H. (2024). Analisis Sentimen Publik di Twitter Pasca Debat Kelima Pilpres 2024 dengan Naive Bayes. *Edumatic: Jurnal Pendidikan Informatika*, *8*(2), 754-763. https://doi.org/10.29408/edumatic.v8i2.28048