Eduvest – Journal of Universal Studies
Volume 6 Number 3, March, 2026
p- ISSN 2775-3735- e-ISSN 2775-3727

# Comparison of F1-Score Naive Bayes, Logistic Regression, K-Nearest Neighbors, and SVM for Sentiment Classification X in Police Institutions

**Robertos Hartanto Wijaya**
Universitas Kristen Satya Wacana, Indonesia
Email: 672021146@student.uksw.edu

***Keywords:***
*F1-Score, Machine Learning, SMOTE, Sentimen Analisis*

**ABSTRACT**

**Background:** Social media, especially platform X, is the main channel for the public to express their opinions on public institutions, including the police. Analysis of public sentiment on this platform can provide insight into police performance. This study aims to compare the performance of machine learning algorithms in the classification of negative sentiment towards policing, focusing on unbalanced social media data. **Objective:** This study aims to compare the performance of machine learning algorithms—Naive Bayes, Logistic Regression, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM)—in classifying negative sentiments towards policing on social media X, as well as overcoming data imbalances using the SMOTE method. **Method:** The dataset consisted of 1,274 Indonesian-language data collected by crawling, then processed using preprocessing techniques such as text cleaning, stopword removal, and TF-IDF feature extraction. Testing is conducted with and without the implementation of SMOTE for data balancing. Evaluate the model's performance using F1-Score. **Result:** Without SMOTE, all algorithms fail to recognize neutral classes. After the implementation of SMOTE, Logistic Regression showed the best performance with an F1-Score of 80.85%, followed by SVM, Naive Bayes, and KNN. The implementation of SMOTE significantly improves the model's ability to classify negative sentiments. **Conclusion:** The combination of Logistic Regression and SMOTE is the best approach to classifying public sentiment towards policing, which can help police agencies understand public sentiment more accurately.

## INTRODUCTION

Social media has become an unstoppable phenomenon. Social media has changed people's lives and the way people interact. Social media is usually used by a person as a medium of communication, a means of information and as a medium of entertainment for its users (Nasution & Hayaty, 2019). X is one of the social media that is microblogging as an interaction service. X has become one of the most well-known social media services in the world with over 200 million active users and more than 10.6 billion tweets generated (Nasution & Hayaty, 2019).

In the ever-evolving digital era, data is an important component in various sectors, including government agencies such as the police. Social media acts as a means of public communication that allows the public to express opinions, criticisms, and suggestions on the performance and policies of government institutions. In this context, the public can actively participate in supporting, observing, and criticizing public policies (Al Mustaqim et al., 2024). Therefore, the data generated from interactions on social media is not only relevant for the sake of communication, but also has value as a source of empirical data in analyzing public perceptions and sentiments on national strategic issues (Syahrohim et al., 2024).

Data generated from people's activities on social media has great potential to be analyzed systematically to understand the trends of public sentiment. Sentiment analysis is one of the widely used approaches to classify public opinion into positive, negative, or neutral categories (Permatasari et al., 2021). This approach is important because it can provide an objective picture of the level of trust, satisfaction, and public criticism of the performance of the police, which can ultimately be an input for policy evaluation and improvement of public services (Handika et al., 2024).

A number of previous studies have examined the use of machine learning algorithms in sentiment analysis. Research by Rangga Nasution shows that the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms have a competitive level of accuracy in the classification of Twitter sentiment (Matarat et al., n.d.). Meanwhile, a study by Y. Handika revealed that Naïve Bayes is effective in analyzing sentiment on police performance, although it still has limitations in handling unbalanced data. Another study by A. Sabir showed that Logistic Regression was able to provide stable performance in social media-based text classification. In addition, a comparative study by K. Matarat emphasized that no one algorithm is always superior, so a comprehensive evaluation of various methods is needed to determine the best model (Julizar & Sulaeman, n.d.).

Although various studies have been conducted, there are still some research gaps. First, most studies only compare algorithms without deeply considering the data imbalance problem that often occurs in social media datasets (Sabir et al., 2024). Second, research that specifically analyzes public sentiment towards police institutions in Indonesia is still limited, especially those that use a combination of resampling techniques such as the Synthetic Minority Over-Sampling Technique (SMOTE). Third, model performance evaluations often focus only on accuracy, whereas metrics such as F1-Score are more relevant in unbalanced data conditions (Bahtiar et al., 2023).

Based on this gap, this study offers novelty by combining the comparison of four machine learning algorithms Naïve Bayes, Logistic Regression, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) with a data balancing approach using the SMOTE method, as well as evaluating model performance using F1-Score metrics (Suandi et al., 2024). This approach is expected to provide a more comprehensive and accurate analysis in classifying public sentiment towards the police (Qi & Shabrina, 2023).

Based on this background, this study focuses on analyzing and comparing the performance of several machine learning algorithms in classifying public sentiment towards the police based on social media data (Cam et al., 2024). The algorithms used in this study include Naïve Bayes, Logistic Regression, K-Nearest Neighbors, and Support Vector Machine.

Model performance evaluation was carried out using the F1-Score metric to assess the balance between precision and recall, especially under unbalanced data conditions (Brownlee et al., 2020).

Through this comparison, it is hoped that a better understanding of how machine learning models can be used to analyze sentiment. The results of this study are also expected to be a reference for further research that wants to develop machine learning models for similar applications. The findings of this study are expected to assist developers and researchers in determining the development stage and the level of classification accuracy of a number of social media-based algorithms to obtain the right information and high accuracy values (Sulasno et al., 2022).

## METHOD

This research involves several stages, namely data collection, data preparation, and sentiment classification process using Naïve Bayes algorithms, Logistic Regression, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) to classify three classes of sentiment, namely positive, negative, and neutral. Furthermore, a comparison of prediction performance and evaluation of results were carried out to determine the algorithm with the most optimal performance, as shown in Figure 1. This research is focused on the classification of negative sentiment. This study uses Python version 3.11.4 and Jupyter Notebook as the development environment, as Python provides various libraries and frameworks that support data analysis and machine learning, such as scikit-learn, pandas, NumPy, and Matplotlib.

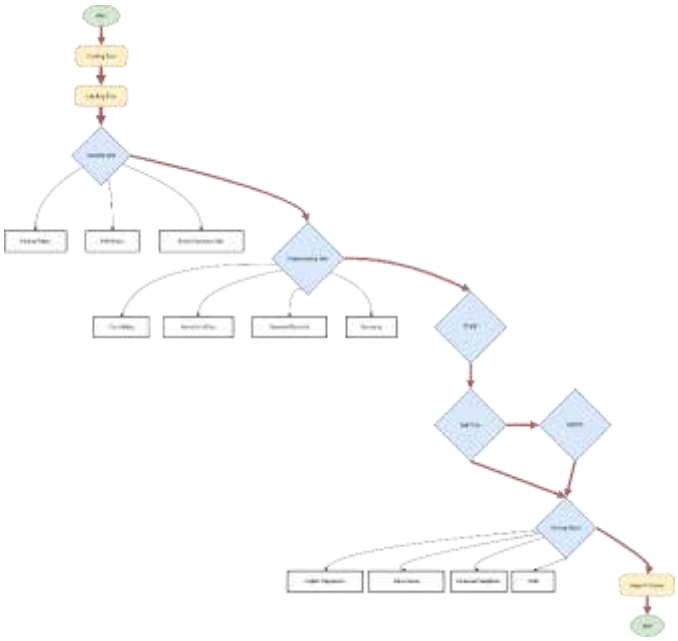


**Figure 1 Research Stages**

Source: Illustration of research flow by researcher (2025)

**Data Crawling**

Data *crawling* is a technique in *machine learning* that is used to collect data contained in a website. This method works by retrieving information with *keywords* that have been entered by the *user* (Dikiyanti et al., 2021).

**Data Cleaning**

Data *cleaning* is an important technique for identifying and correcting data errors (Iqbal et al., 2024). In this study, whose data was collected through *crawling* techniques, a lot of data may contain *noise*, therefore data cleaning was carried out to search for missing values, null values and duplicate values.

a. $\| \mathbf{x}_i - \mathbf{x}_j \|^2$adalah *testing.*

**Synthetic Minority Over-Sampling Technique (SMOTE)**

Terjadi ketika jumlah data pada kelas mayoritas jauh lebih besar dibanding kelas minoritas. Kondisi ini dapat menyebabkan model cenderung berpihak pada kelas mayoritas sehingga performa deteksi kelas minoritas menurun (Effendi et al., 2023). Salah satu pendekatan yang umum digunakan untuk mengatasi masalah ini adalah data-level resampling, khususnya metode oversampling seperti (Wang et al., 2021) *SMOTE* (Chen et al., 2024).

*SMOTE* works by creating synthetic samples in minority classes, rather than simply duplicating existing data. In general, for each minority sample, $\mathbf{x}_i$*SMOTE* will look for $k$*the nearest neighbors* of the minority class. Then *SMOTE* selects one of the neighbors at random and generates new data with linear interpolation as follows$\mathbf{x}_{zi}$ (Taskiran et al., 2025)

$$\mathbf{x}_{new} = \mathbf{x}_i + \lambda(\mathbf{x}_i{}^{(k)} - \mathbf{x}_i)$$

dengan keterangan sebagai berikut:

a. $\mathbf{x}_{new}$ : merupakan sampel sintetis baru yang dihasilkan dari proses *SMOTE*.
b. $\mathbf{x}_i$ : merupakan data yang dipilih dari kelas minoritas sebagai titik awal pembentukan sampel baru.
c. $\mathbf{x}_i{}^{(k)}$: adalah salah satu tetangga terdekat dari $x_i$berdasarkan nilai $k$-nearest neighbors.
d. λ : merupakan bilangan acak dalam rentang 0 hingga 1 yang mengatur proporsi jarak antara $x_i$dan tetangganya dalam proses interpolasi.

**Metrik Kinerja dan Evaluasi**

Penelitian ini akan menggunakan *F1-Score* sebagai metrik kinerja utama untuk mengevaluasi performa algoritma *Naïve Bayes*, *Logistic Regression*, dan *K-Nearest Neighbors*. Evaluasi akan mencakup analisis kelebihan dan kekurangan masing-masing algoritma dalam konteks analisis sentimen (Amri et al., 2025).

*F1-Score* dihitung menggunakan persamaan berikut (Rakasiwi et al., 2024):

$$F1\ SCORE = 2 \text{ x } \frac{\text{Precission x Recall}}{\text{Precission + Recall}}$$

Dengan precission dan recall didefinisikan sebagai berikut[29]:

$$Precission = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Remarks:

a. *True Positives (TP)* is the amount of data that is predicted to be positive and correct according to the actual class.
b. *False Positives (FP)* are the amount of data that is predicted to be positive but actually belongs to the negative class.
c. *False Negatives (FN)* are the amount of data that is predicted to be negative but actually belongs to the positive class.

Accuracy is the ratio of the number of correct positive predictions to all positive predictions generated by the model. Meanwhile, recall is the ratio between the number of correct positive predictions to all actual data that is included in the positive class.

## RESULT AND DISCUSSION

### Distribusi Kelas Dataset

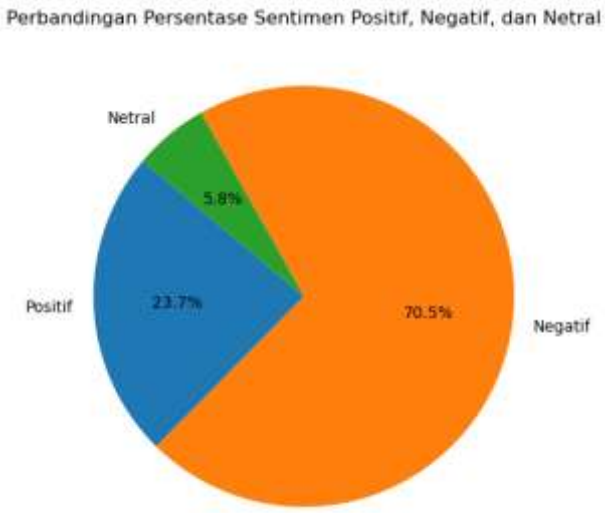
Data yang digunakan dalam penelitian ini diperoleh melalui proses crawling pada platform X dengan menggunakan kata kunci "Polisi OR Kepolisian OR Kinerja Polisi OR Kinerja Kepolisian OR Polri". Pengambilan data dilakukan pada rentang waktu 13 Maret 2024 hingga 7 September 2025, kemudian dilakukan proses penyaringan untuk memilih data yang menggunakan bahasa Indonesia (Manaf et al., 2023). Hasil proses crawling menghasilkan sebanyak 1.319 data, sebagaimana ditunjukkan pada Tabel 1. Jumlah data yang diperoleh relatif terbatas karena adanya kebijakan dan pembatasan akses dari platform X terhadap aktivitas pengambilan data. Selanjutnya, dilakukan proses pembersihan data untuk menghilangkan data duplikat, sehingga jumlah data akhir yang digunakan dalam penelitian ini menjadi 1.274 data.

**Tabel 1 Data Hasil Crawling**

| No | Text | Sentimen |
|---|---|---|
| 1 | @HerumawanPA @merapi_uncover @FXHarminanto icj itu lebih kuat dari akun X karena anggotanya lebih kritis dari adminnya. pernah inget kasus kapolres bantul dibully di ICJ? besoknya polda bikin himbauan/ acara2 cari simpati tapi tiap kali polisi salah tetep dibully | Negatif |
| 2 | Gini kah lemes tuh? Udah terlalu malas buat ladenin ya Allah i wish that this country had good cops yg do their work as they should. Masalah gini harusnya ditangani polisi.. tapi.. polisi di sini tuh.. | Negatif |

| | | |
|---|---|---|
| 3 | @persebayaupdate Kalo yang melanggar polisi seperti menembak gas air mata berani nuntut ga? | Negatif |
| 4 | Fix polisi sekarang kurang duit. Sampe kek gini. Mending ga usah pake polisiÃ‚Â²an. | Negatif |
| 5 | @mardigu024 Tidak Ada Berita Baik Dari Polisi. | Negatif |

Hasil analisis dari pengumpulan data pada platform X disajikan pada Gambar 1. Berdasarkan visualisasi diagram tersebut, distribusi persentase dibagi dari tiga kategori. Distribusi terbanyak menunjukkan pada sentiment negative dengan proporsi sebanyak 70,5%. Sentimen positif menyusul di urutan kedua dengan persentase sebanyak 23,7%. Sementara itu, sentiment netral menjadi sentiment dengan presentase terendah sebanyak 5,8% (Mantika et al., 2024).





*Gambar 2 Wordcloud sentimen*

**Source:** Data visualization using Python (WordCloud) by researchers (2025)

In Figure 2. Shows a visualization of the frequency of frequently appearing data presented in the form of WordCloud. The high frequency of keywords that often appear in this data is seen from the proportional font size. In this visualization, the keywords that appear most often are "POLICE", "POLRI", and "SOCIETY" (Thurnhofer-Hemsi et al., 2020). The dominance of the word "POLICE" indicates that this word is the main focus in this data. In addition, other words that appear with high frequency include 'work', 'law', 'individual', 'member', and 'people'. The appearance of the words 'person' and 'wrong' (with related phrases such as 'misconduct') gives a strong indication that the discussion that takes place is mostly related to controversial issues, violations, or criticism of the institution's performance (AminiMotlagh et al., 2022).

Data Preprocessing Results

The preprocessing stage includes cleaning mentions and URLs, normalizing words, removing punctuation, case folding, and removing stopwords and stemming. The preprocessing results show that sentence structure becomes simpler and focuses on meaningful words, thus improving the performance of the classification algorithm as seen in Table 2.

**Table 2 Preprocessing Results Data,**

| No | Original | Preprocessing |
|---|---|---|
| 1 | @HerumawanPA @merapi_uncover @FXHarminanto icj itu lebih kuat dari akun X karena anggotanya lebih kritis dari adminnya. pernah inget kasus kapolres bantul dibully di ICJ? besoknya polda bikin himbauan/ acara2 cari simpati tapi tiap kali polisi salah tetep dibully | icj itu lebih kuat dari akun kali karena anggota lebih kritis dari adminnya pernah ingat kasus kapolres bantul dibully di icj besok polda bikin himbauan acara cari simpati tapi tiap kali polisi salah tetap dibully |
| 2 | Gini kah lemes tuh? Udah terlalu malas buat ladenin ya Allah i wish that this country had good cops yg do their work as they should. Masalah gini harusnya ditangani polisi.. tapi.. polisi di sini tuh.. | begini kah lemes tuh sudah terlalu malas buat laden ya allah i wish that this country had good cops yang di their work as they should masalah begini harus tangan polisi tapi polisi di sini tuh |
| 3 | @persebayaupdate Kalo yang melanggar polisi seperti menembak gas air mata berani nuntut ga? | kalau yang langgar polisi seperti tembak gas air berani tuntut tidak |
| 4 | Fix polisi sekarang kurang duit. Sampe kek gini. Mending ga usah pake polisiÃ‚Â²an. | fix polisi sekarang kurang duit sampai kayak begini mending tidak usah pakai polisi |
| 5 | @mardigu024 Tidak Ada Berita Baik Dari Polisi. | begini kah lemes tuh sudah terlalu malas buat laden ya allah i wish that this country had good cops yang di their work as they should masalah begini harus tangan polisi tapi polisi di sini tuh |

**Source:** Results of data preprocessing by researchers (2025)

**Model Results without Using Smote.**

Initial testing was performed on the original dataset without class imbalance handling. The results of the performance evaluation of each algorithm are shown in Table 3. Based on the F1-Score score, Logistic Regression showed the best performance with a score of 76.93%, followed by SVM at 75.59%, KNN at 75.48%, and Naïve Bayes at 75.05%. Possible results (Wijati et al., 2024).

***Tabel 1 Akurasi Algoritma tanpa SMOTE***

| | Naïve Bayes | Logistic Regression | *KNN* | SVM |
|---|---|---|---|---|
| Accuracy | 0.8039 | 0.8157 | 0.7961 | 0.8078 |
| Precision | 0.7633 | 0.7564 | 0.7480 | 0.7657 |
| Recall | 0.8039 | 0.8157 | 0.7961 | 0.8078 |
| F1-Score | 0.7505 | **0.7693** | 0.7548 | 0.7559 |

**Source:** Data processing results using Python (2025)

Based on Figure 3, the initial dataset shows a significant class imbalance, with 894 negative sentiment data, 225 positive sentiment, and 155 neutral sentiment data. After applying the SMOTE technique with sampling_strategy = 'auto' to the data train, the amount of data in each class became balanced, which was 709 each. This shows that the auto strategy succeeds in oversampling the minority class while cleaning up data that has the potential to cause classification ambiguity.

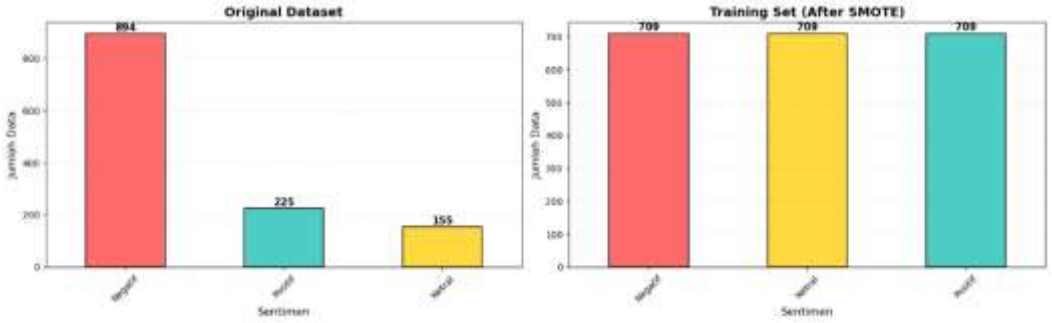



**Figure 4 Data distribution after balancing with SMOTE**

**Source:** Visualization of data distribution after SMOTE using Python by researchers (2025)

**Tabel 2 Akurasi SMOTE**

| | **Naïve Bayes** | **Logistic Regression** | ***KNN*** | **SVM** |
|---|---|---|---|---|
| Accuracy | 0.7255 | 0.8078 | 0.6863 | 0.8118 |
| Precision | 0.8181 | 0.8106 | 0.7758 | 0.7929 |
| Recall | 0.7255 | 0.8078 | 0.6863 | 0.8118 |
| F1-Score | 0.7627 | **0.8085** | 0.7240 | 0.7752 |

**Source:** Data processing results after the application of the SMOTE method by the researcher (2025)

After data balancing using the SMOTE method, there was a change in the performance of each algorithm seen in Table 4. Logistic Regression showed an improvement in performance with an F1-Score of 80.85%, which is the highest value compared to other algorithms. SVM

also experienced an increase with an F1-Score of 77.52%. Naïve Bayes experienced a smaller increase, while KNN still showed the lowest performance among the four algorithms. This performance improvement shows that SMOTE is effective in improving minority class representation, so that the model becomes more balanced in studying positive and neutral sentiment patterns. Algorithms such as Logistic Regression and SVM based on decision boundary formation tend to benefit more from a balanced class distribution, thus being able to significantly increase recall values and F1-Scores.

**Comparative Analysis of F1-Score by Class**

**Tabel 5 Perbandingan F1-Score per-kelas sebelum dan sesudah SMOTE**

| | Negatif | Netral | Positif |
|---|---|---|---|
| NB | 0.88 | 0.00 | 0.58 |
| LR | 0.89 | 0.00 | 0.65 |
| *KNN* | 0.87 | 0.00 | 0.63 |
| SVM | 0.88 | 0.00 | 0.60 |
| NB + SMOTE | 0.83 | 0.22 | 0.73 |
| **LR + SMOTE** | **0.89** | **0.27** | **0.74** |
| *KNN* + SMOTE | 0.80 | 0.16 | 0.69 |
| SVM + SMOTE | 0.89 | 0.08 | 0.66 |

Source: Results of machine learning model analysis by researchers (2025)

The results of the performance evaluation per class before and after the implementation of SMOTE are presented in Table 5. This evaluation was conducted to look in more detail at the model's ability to classify each category of sentiment, especially minority classes. Prior to the implementation of SMOTE, all algorithms showed an F1-Score value of 0.00 in the neutral class. This suggests that the model is not able to recognize the neutral class at all and tends to classify the entire data into the majority class, i.e. the negative class.

This condition indicates model bias due to the unbalanced distribution of classes in the dataset. After data balancing using SMOTE, there was an increase in the model's ability to recognize minority classes, especially in neutral and positive classes. Logistic Regression showed the most significant increase in the neutral class with an F1-score of 0.27, followed by Naïve Bayes of 0.22. However, the F1-Score in the neutral class is still lower than the negative and positive classes in the entire algorithm. This suggests that the classification of neutral sentiment is still a challenge, likely due to the characteristics of neutral texts that tend to be ambiguous and have similarities to both positive and negative classes.

## CONCLUSION

From the results of data collection and cleaning, 1,274 data in Indonesian were obtained which were used as research datasets. The results of the evaluation before the implementation of SMOTE showed that all algorithms failed to recognize neutral classes with an F1-Score value of 0.00, which indicates a model bias towards the majority class. The results of the test of the model without SMOTE showed that Logistic Regression provided the best performance with an F1-score of 76.93%, followed by SVM of 75.59%, KNN of 75.48%, and Naïve Bayes of 75.05%.

However, after applying SMOTE to the data train, there was an increase in the model's

ability to recognize minority classes, especially in neutral and positive classes. Logistic Regression showed the most significant improvement with the neutral class F1-Score of 0.27 and the overall F1-Score of 80.85%, followed by Naïve Bayes and KNN with moderate improvements, while SVM showed a relatively limited increase in the neutral class. These results suggest that data balancing plays an important role in improving the model's sensitivity to minority classes, although challenges in classifying neutral sentiment remain. Based on the overall results, the combination of Logistic Regression and SMOTE is the most optimal approach in this study to classify public sentiment towards policing on social media data, both in terms of overall performance and ability to recognize minority classes.

## REFERENCE

Al Mustaqim, D., Hakim, F. A., Atfalina, H., & Fatakh, A. (2024). Peran media sosial sebagai sarana partisipasi warganet dalam mewujudukan keadilan dan akuntabilitas penegakan hukum di Indonesia. *Journal of Multidisciplinary Research and Development*, *1*(1), 53–66.

AminiMotlagh, M., Shahhoseini, H., & Fatehi, N. (2022). A reliable sentiment analysis for classification of tweets in social networks. *Social Network Analysis and Mining*, *13*(1), 7.

Amri, M. R. A., Permana, E., Pachadria, P. A., & Fitri, S. (2025). Perbandingan Metode Naïve Bayes dan Random Forest dalam Memprediksi Penyakit Diabetes Melitus pada Klinik Citra Sejati. *Jurnal Teknologi Informasi Dan Multimedia*, *7*(4), 847–858.

Bahtiar, S. A. H., Dewa, C. K., & Luthfi, A. (2023). Comparison of Naïve Bayes and logistic regression in sentiment analysis on marketplace reviews using rating-based labeling. *Journal of Information Systems and Informatics*, *5*(3), 915–927.

Brownlee, J., Sanderson, M., Koshy, A., Cheremskoy, A., & Halfyard, J. (2020). Machine Learning Mastery With Python: Data Cleaning, Feature Selection, and Data Transforms in Python. *Machine Learning Mastery: Vermont, VIC, Australia*.

Cam, H., Cam, A. V., Demirel, U., & Ahmed, S. (2024). Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers. *Heliyon*, *10*(1).

Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. P. (2024). A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review*, *57*(6), 137.

Dikiyanti, T. D., Rukmi, A. M., & Irawan, M. I. (2021). Sentiment analysis and topic modeling of BPJS Kesehatan based on twitter crawling data using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation algorithm. *Journal of Physics: Conference Series*, *1821*(1), 12054.

Effendi, I. F., Utami, D. A., Rahmawati, R. A., Prasetyowibowo, R., & Isbandono, P. (2023). Twitter data sentiment analysis on the economic sector: Public response to government policies during the covid-19 pandemic in indonesia. *International Joint Conference on Arts and Humanities 2023 (IJCAH 2023)*, 472–491.

Handika, Y., Hanif, I. F., & Hasan, F. N. (2024). Analysis of Public Sentiment Towards POLRI's Performance using Naive Bayes and K-Nearest Neighbors. *IJID (International Journal on Informatics for Development)*, *13*(1), 386–399.

Iqbal, M., Afdal, M., & Novita, R. (2024). Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store: Implementation of Support Vector Machine Algorithm for Sentiment Analysis of Online Loan Application Review Data on Google Play Store. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(4), 1244–1252.

Julizar, A., & Sulaeman, M. K. (n.d.). *Evaluation of Logistic Regression and Random Forest*

*Algorithms for Hate Speech Identification*.

Manaf, S. A. R., Alamudi, A., & Fitrianto, A. (2023). Analisis Sentimen Tanggapan Masyarakat Pengguna Twitter terhadap Pembelajaran Tatap Muka. *Indonesian Journal of Statistics & Its Applications*, *7*(1).

Mantika, A. M., Triayudi, A., & Aldisa, R. T. (2024). Sentiment analysis on twitter using naïve Bayes and logistic regression for the 2024 presidential election. *SaNa: Journal of Blockchain, NFTs and Metaverse Technology*, *2*(1), 44–55.

Matarat, K., Mingmuang, C., & Charoenrat, W. (n.d.). A Comprehensive Performance Analysis of Supervised Machine Learning Techniques for Sentiment Analysis. *International Journal of Computer Applications*, *975*, 8887.

Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen twitter. *Jurnal Informatika*, *6*(2), 226–235.

Permatasari, P. A., Linawati, L., & Jasa, L. (2021). Survei Tentang Analisis Sentimen Pada Media Sosial. *Majalah Ilmiah Teknologi Elektro*, *20*(2), 177.

Qi, Y., & Shabrina, Z. (2023). Sentiment analysis using Twitter data: a comparative application of lexicon-and machine-learning-based approach. *Social Network Analysis and Mining*, *13*(1), 31.

Rakasiwi, R. K. M., Kurnianingsih, K., Suharjono, A., Enriko, I. K. A., & Kubota, N. (2024). Predicting Battery Storage of Residential PV Using Long Short-Term Memory. *JOIV: International Journal on Informatics Visualization*, *8*(1), 141–149.

Sabir, A., Ali, H. A., & Aljabery, M. A. (2024). ChatGPT tweets sentiment analysis using machine learning and data classification. *Informatica*, *48*(7).

Suandi, F., Anam, M. K., Firdaus, M. B., Fadli, S., Lathifah, L., Yumami, E., Saleh, A., & Hasibuan, A. Z. (2024). Enhancing Sentiment Analysis Performance Using SMOTE and Majority Voting in Machine Learning Algorithms. *7th International Conference on Applied Engineering (ICAE 2024)*, 126–138.

Sulasno, M. S., Amalia, H., & Paramarta, V. (2022). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Vaksin Covid-19 Pada Sosial Media Twitter Menggunakan Metode Naïve Bayes dan k-Nearest Neighbors. *TopUp Perbanas*, *1*(1).

Syahrohim, I., Saputra, S. D., Saputra, R. W., Pranatawijaya, V. H., & Priskila, R. (2024). Perbandingan analisis sentimen setelah pilpres 2024 di Twitter menggunakan algoritma machine learning. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(2). https://doi.org/10.23960/jitet.v12i2.4249

Taskiran, S. F., Turkoglu, B., Kaya, E., & Asuroglu, T. (2025). A comprehensive evaluation of oversampling techniques for enhancing text classification performance. *Scientific Reports*, *15*(1), 21631.

Thurnhofer-Hemsi, K., López-Rubio, E., Molina-Cabello, M. A., & Najarian, K. (2020). Radial basis function kernel optimization for support vector machine classifiers. *ArXiv Preprint ArXiv:2007.08233*.

Wang, L., Han, M., Li, X., Zhang, N., & Cheng, H. (2021). Review of classification methods on unbalanced data sets. *Ieee Access*, *9*, 64606–64628.

Wijati, D., Atika, P. D., Setiawati, S., & Rasim, R. (2024). Sentiment analysis of application reviews using the K-Nearest Neighbors (KNN) algorithm. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, *12*(1), 209–218.