

Tweet Sentiment Analysis with Support Vector Machine (SVM) Algorithm for PT. XYZ Digital Strategy

Nazilatul Azza¹, Nur Hadian²✉

^{1,2} Information Technology, Faculty of Science and Technology, Institut Teknologi dan Sains Nahdlatul Ulama Pekalongan, Indonesia

✉ **Corresponding Author** : Nur Hadian (e-mail: nurhadian97@gmail.com)

Article Information	ABSTRACT
Article History Received : August 16, 2025 Revised : September 02, 2025 Published : October 04, 2025	The development of the digital era has encouraged the use of social media as a strategic tool for promotion and information dissemination, including in the publishing industry. PT. XYZ utilizes Twitter to raise awareness of literacy issues. This study analyzes public sentiment towards literacy through tweets with the hashtags #literasi, #literasiindonesia, and #penerbitbuku using the Support Vector Machine (SVM) algorithm. The research stages include scraping, preprocessing, labeling, TF-IDF vectorization, and evaluation with a confusion matrix. From the 466 tweets analyzed, a balanced distribution of positive and negative sentiments was obtained. The model produced an accuracy of 47.87%, precision of 47%, recall of 57%, and an F1-score of 51%. These results are lower than those of Afrillia et al. (2022), who achieved an accuracy of 70.8%, and Widyanto et al. (2023), who obtained 80.41% with an RBF kernel. These differences confirm the limitations of SVM on small datasets and informal language. This study contributes by showing the potential and limitations of SVM in analyzing literacy on social media. These results also emphasize the need for further research with larger datasets and advanced methods such as ensemble learning and deep learning (LSTM, BERT).
Keywords: <i>Sentiment Analysis;</i> <i>Support Vector Machine;</i> <i>Literacy;</i> <i>Social Media;</i> <i>Digital Strategy.</i>	

INTRODUCTION

Technological developments have brought significant changes to marketing strategies in business circles (Ahadiyah, 2024). In the modern business world, companies are required to be able to utilize information technology effectively to increase competitiveness (Edwin Kiky Aprianto, 2021). In this context, effective information technology management can have an important role in the business world, by providing added value to the business through an accountability framework process that is focused on achieving strategic goals in the use of information technology (Hadian et al., 2023). This progress also encourages the birth of various new ideas that change marketing strategies to be more adaptive in the digital era.

Twitter is a social media platform that focuses on short, real-time communication. With more than 140 million active users and around 400 million tweets per day (Zusrotun et al., 2022). Twitter is a text-based social media platform that gives users the freedom to create and share posts in the form of text, and has great potential as a source of data for sentiment analysis (Falenchia & Sumardijjati, 2023). This potential is further strengthened by the use of hashtags on Twitter which function as a grouping of topics and can be an indicator of public conversation trends (Azmi et al., 2021). Twitter social media is often widely used by companies for sentiment

analysis by utilizing hashtags or hashtags that function to group content and facilitate searches. The more users who use a hashtag, the hashtag will become a trend or trending topic on twitter (Umar et al., 2022). Hashtags on Twitter have several functions, including facilitating the search and presentation of information, as well as marking topics that are currently being discussed or trending (Budiman & Widiksono, 2018). As part of the text data that develops on Twitter, hashtags generally reflect issues or topics that are receiving attention from the public (Umar et al., 2022). This information has the potential to be evaluated by various industries, including publishing companies, to reach a wider audience.

Although social media has great potential, not all companies can utilize it optimally. PT. xyz, a publishing company, has started to utilize social media for literacy promotion and education, but the effectiveness of this strategy has never been analyzed in depth. The lack of data analysis causes the digital marketing strategy that is carried out to be not optimal, from the lack of data analysis this study seeks to reveal how public perceptions of literacy issues based on data related to literacy hashtags in tweets while evaluating the extent to which the SVM algorithm is able to classify literacy sentiment on social media.

Sentiment analysis is a subset of text classification that focuses on natural language processing (NLP), computational linguistics, and text mining. This technique is used to analyze opinions, judgments, and expressions in text (Mailoa, 2021). In hashtag or sentiment analysis, there are several algorithms that can be used, one of which is Support Vector Machine (SVM). SVM is one of the algorithm models derived from statistical learning theory used for classification and regression tasks sometimes used in research fields, such as text mining (Hendrawan & Kusniyati, 2024). In addition, this algorithm has a higher accuracy rate than other algorithms.

In previous research on the Comparison of Naïve Bayes and SVM Sentiment Analysis of the Health Bill on Twitter, the results of this study indicate that the SVM algorithm shows superior accuracy results (0.87) compared to Multinomial Naïve Bayes (0.82), and produces more precision, recall, and F1-score (Widyanto et al., 2023). Previous research discussed the comparison of SVM and Naïve Bayes Algorithms on Sentiment Analysis of Thesis Obligation Removal Policy, this study used two text classification algorithms and a supervised learning approach, with the comparison results obtained that the SVM algorithm became a superior algorithm compared to Naïve Bayes (Yunita & Kamayani, 2023). Similar research was conducted on Twitter Sentiment Analysis Regarding the Implementation of Permendikbudristek Number 30 of 2021 Using TextBlob and Support Vector Machine by producing an accuracy of 70.8% with 35.4%, recall 50% and F1-score of 41.4%, this study shows the dominance of positive public opinion towards the policy (Afrillia et al., 2022). Although sentiment analysis research on Twitter has been conducted extensively, there has been no study that specifically highlights literacy sentiment in Indonesia in the context of publishing companies' digital strategies. Most previous studies have focused on political issues, public policy, education, or entertainment, so that the aspect of literacy has received little attention.

With this trend and sentiment analysis, PT Xyz is expected to develop a more effective and adaptive digital marketing strategy. In addition, the utilization of social media as a means of promotion is expected to increase the visibility of publishing companies and contribute to introducing literacy to a wider community, especially the younger generation who are more active in various social media. In line with this, the main objective of this study is to map public perceptions of literacy issues on Twitter while evaluating the extent to which the Support Vector Machine (SVM) algorithm can be used as a tool in formulating digital strategies for publishing companies.

RESEARCH METHODS

In this research utilizing the Support Vector Machine (SVM) algorithm to analyze the opinion or sentiment, the steps applied in the research are shown in **Figure 1** which refers to the approach (Umar et al., 2022) which is the basis for designing this research method.

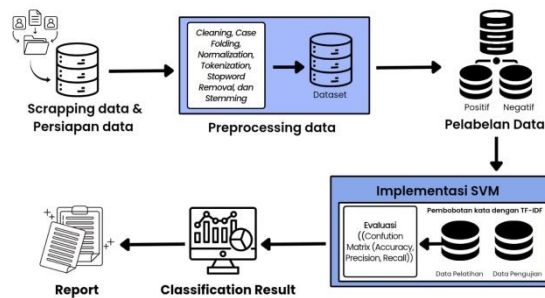


Figure 1. Research steps

The explanation of the research steps used in this study, as shown in **Figure 1**, is as follows:

Scrapping Data

The data retrieval process uses crawling techniques, namely with the API from twitter where the process uses a library in the python code, namely the tweety library. This research takes hashtag data and opinions from several topics such as #literasi, #literasiindonesia, and #penerbitbuku on twitter, then the data will go through an export process into a csv file, so that a csv file is obtained with the names literacy, literasiindonesia, and book publishers.

Data Preprocessing

This step is carried out before entering the classification process, the collection of tweet data that has been collected must go through a preprocessing stage to transform data that was originally unstructured or contained a lot of noise, into structured data so that the analysis process becomes easier (Husada & Paramita, 2021). Starting with the data cleaning process (the step of deleting or filling in empty or noisy data), case folding (converting data or standardizing capital letters to lowercase and removing various symbols), normalization (the process of converting text data into standard words according to the KBBI), tokenizing (separating each word in a sentence into tokens), filtering or stopword removal (removing words that are not relevant to sentiment analysis), and stemming (removing prefixes or suffixes from words that contain conjunctions, prepositions, or pronouns to form base words that comply with the KBBI) (Darwis et al., 2020).

Dataset Labeling

In this step, labels will be given to the data that has been carried out in the data cleaning and preprocessing steps. Full text data will be given positive meanings that are categorized as positive sentiments, while negative words are categorized as negative sentiments. In labeling full_text data, we use a lexicon dictionary as a reference to ensure accuracy in labeling. Lexicon dictionary or lexicon can refer to a collection of vocabulary, a simple dictionary, a list of terms arranged alphabetically along with their meanings, language units that have meanings and their uses, as well as certain keywords (Ningsih & Indrawati, 2023).

SVM Model

The analysis stage was performed using the Support Vector Machine (SVM) method with a linear kernel, where the parameters $C=1.0$, $\text{gamma}=\text{'scale'}$, and $\text{decision_function_shape}=\text{'ovr'}$ remained at their default values from scikit-learn. The linear kernel was selected to form a hyperplane that separates positive and negative sentiments linearly. The model was evaluated using a confusion matrix to assess classification performance through accuracy, precision, recall, and F1-score metrics, while the SVM model function formula is shown in Equation (1) (Munandar et al., 2023).

$$f(x) = W^T X + b \quad (1)$$

In the classification function using SVM is represented by $f(x)$. The function $f(x)$ is used to predict the target value y based on the input feature x , and W represents the weight vector, X stands for the input feature vector, and b is the bias. So based on the statement of Vapnik and Cortes (1995), a formula equation is obtained with the following formulation:

$$[(WT.xi) + b] \geq 1 \text{ untuk } y_i = 1 \quad (2)$$

$$[(WT.xi) + b] \leq -1 \text{ untuk } y_i = -1 \quad (3)$$

Evaluation

The step used to perform the evaluation tests the performance of the method in the Support Vector Machine (SVM) algorithm. In this evaluation using the confusion matrix instrument then used to calculate sensitivity, specificity, and accuracy. sensitivity is used to measure how much the amount of data that is successfully classified correctly (True Positive) compared to the total amount of data included in the positive class (Ramlan et al., 2023).

RESULTS AND DISCUSSION

This research report collects data using the Twitter API with a data crawling technique created using Python. In collecting tweet data, it uses hashtags that are frequently used on Instagram feeds by PT. xyz. The completion of the data collection process is shown in Figure 2, which displays the sentiment data resulting from hashtag scraping on February 12, 2025. The collected data consists of 213 tweets with the hashtag #literasi, 44 tweets with the hashtag #penerbitbuku, and 209 tweets with the hashtag #literasiindonesia, for a total of 466 raw tweets.

	conversation_id_str	created_at	favorite_count	full_text	id_str
0	1896421847429992500	Mon Mar 03 04:48:16 +0000 2025	1	Si Humas Polres Malinau melaksanakan kegiatan ...	1896421847429992500 https://pbs.twimg.com/media/GIF
1	1896220427414126701	Mon Mar 03 04:35:54 +0000 2025	1	@skalalikeur bener literasi bindo dan rasa nas...	1896419238065369165
2	1895592933530984878	Mon Mar 03 04:33:50 +0000 2025	0	@lovepsygom96 @9uly101gB2 @lowwwselfesteem @a...	1896418717573238865
3	1896416300165136570	Mon Mar 03 04:24:14 +0000 2025	0	oon bgt minim literasi togocek aku yauda la ya...	1896416300165136570 https://pbs.twimg.com/media/GIF
4	1896415814993301686	Mon Mar 03 04:22:18 +0000	0	Kasat Lantas Polres Tanggamus Melaksanakan	1896415814993301686 https://pbs.twimg.com/media/GIF

Figure 2. Crowlling of Merged Data

Scraping Data

The data in **Figure 2** is the result of scrapping using data crwaling techniques that have been merged into one, then selected from some of the data above, full_text data is taken to be used as data to be analyzed.

Data Preprocessing

In this step, the preprocessing process is carried out so that the data can be processed by the system. Here are some stages in the data processing process:

a. Cleaning Data

From the selected columns as shown in **Table 1**, the processed data found missing values in the image_url, in_reply_to_screen_name, and location categories. These missing values were then filled in with the labels "No image," "Not a reply," and "Unknown." This was done so as not to lose any existing data. In addition, punctuation cleaning was also performed, such as removing numbers, removing unnecessary punctuation, and removing excess spaces in the data.

Table 1. Changes after Data Cleanning

Before Clean_text	After Clean_text
@skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi	skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi

b. Case Folding

Table 2 shows an example of the results of the steps used to convert all characters in full_text to lowercase, so that the data format is more consistent for further processing.

Table 2. Sample Case Folding Data

Before Case folding	After Case folding
skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi	skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi

c. Word Normalization

Table 3 shows a sample of normalized data. This step was performed using a standard word file taken from Kaggle to adjust non-standard words to their standard form in the full_text data.

Table 3. Sample Data Normalization

Before Normalization	After Normalization
skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi	skalalikeur benar literasi bindo dan rasa nasionalismenya kena efisiensi

d. Tokenizing

Table 4 shows the results of breaking down the data into tokens (word units) used in sentiment analysis.

Table 4. Tokenize Data Sample

Before Tokenizing	After Tokenizing
skalalikeur benar literasi bindo dan rasa nasionalismenya kena efisiensi	['skalalikeur', 'benar', 'literasi', 'bindo', 'dan', 'rasa', 'nasionalismenya', 'kena', 'efisiensi']

e. Stopword Removal (Filtering)

Table 5 shows the removal of common words that do not have a clear meaning or do not contribute to the sentiment analysis process.

Table 5. Sample Filtering Data

Before Filtering	After Filtering
['skalalikeur', 'benar', 'literasi', 'bindo', 'dan', 'rasa', 'nasionalismenya', 'kena', 'efisiensi']	['skalalikeur', 'literasi', 'bindo', 'nasionalismenya', 'kena', 'efisiensi']

f. Stemming

Table 6 shows the results of changing affixed words into root words, so that the meaning is more consistent in the analysis.

Table 6. Sample Data Stemming

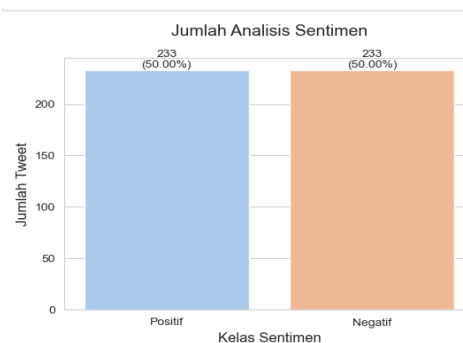
Before Stemming	After Stemming
['skalalikeur', 'literasi', 'bindo', 'nasionalismenya', 'kena', 'efisiensi']	['skalalikeur', 'literasi', 'bindo', 'nasionalisme', 'kena', 'efisiensi']

Data Labeling

After getting clean data with 466 data points, the data was labeled positive and negative to be used as data labeling. From these results, the labeled dataset was then stored in the system for further sentiment analysis. The data labeling can be seen in **Table 7** and differences in labeling data can be seen in **Figure 3**.

Table 7. Labeling Data Sample

No	Before Data Labeling	After Data Labeling
1	skalalikeur bener literasi bindo dan rasa nasionalismenya kena efisiensi	Negatif
2	lovepaygorn96 9Uy10j1gB2 lowwwselfesteem arunikaasha Literasi dulu	Positif

**Figure 3.** Overall Sentiment Label Distribution

Based on the results of the data labeling that has been collected, the data is balanced between positive sentiment (233 data points) and negative sentiment (233 data points). This may indicate that public opinion on the topics of literacy, literasiindonesia, and book publishers tends to be diverse and not dominated by any particular sentiment.

Modelling SVM

Before implementing the Support Vector Machine (SVM) algorithm model, the labeled data was first divided into two parts with a data ratio of 4:1, namely test data (372 or 79.83% of the data) and training data (94 or 20.17% of the data). Then continued with the text classification process at this stage, using the Support Vector Machine method, this part of the data will be tested using a ratio of 0.2, which means 20% of the data used for testing data and 80% used for training data. After data division, continued with the word weighting method carried out using the TF-IDF method with the help of the TfidfVectorizer module in **Figure 4** shows the detailed results of weighting (Safitri et al., 2023).

```

Matriks Vektorisasi untuk Data Latih:
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 ...
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]]

Sebagian kecil Matriks Vektorisasi untuk Data Latih:
[[0.      0.      0.      ... 0.      0.      0.      ]
 [0.      0.      0.      ... 0.      0.      0.      ]
 [0.      0.      0.      ... 0.      0.      0.      ]
 [0.      0.      0.34824013 ... 0.      0.      0.      ]
 [0.      0.      0.      ... 0.      0.      0.      ]]

```

Figure 4. TF-IDF Vectorization of Training Data and Test Data

Evaluation

In this step, all processes in SVM that have been carried out will be displayed as the results of the confusion matrix evaluation addressed in Figure 5 and Figure 6 below:

```
#Menampilkan confusion matrik dalam angka
cm_svm = confusion_matrix(y_test, y_pred_svm)
print("SVM Confusion Matrix:")
print(cm_svm)
```

SVM Confusion Matrix:
[[26 20]
[29 19]]

SVM Accuracy: 0.4787234042553192

SVM Accuracy: 47.87%

SVM Classification Report:

	precision	recall	f1-score	support
Negatif	0.47	0.57	0.51	46
Positif	0.49	0.40	0.44	48
accuracy			0.48	94
macro avg	0.48	0.48	0.48	94
weighted avg	0.48	0.48	0.47	94

Figure 5. Accuracy of Testing and Training Data

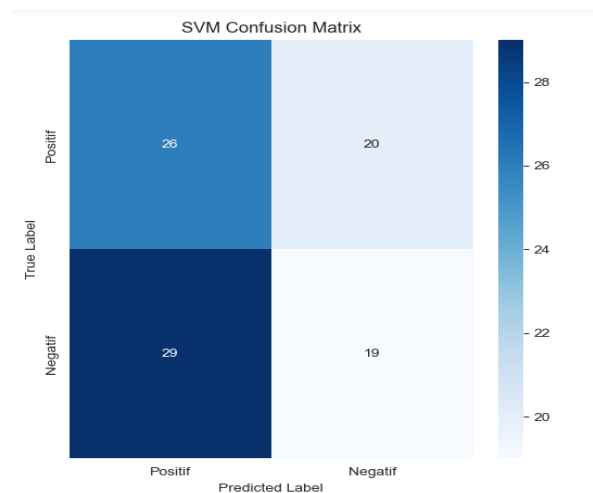
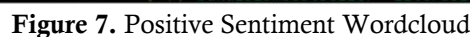


Figure 6. Visualization of Testing and Training Data Accuracy

Based on the confusion matrix on the test and training data, it shows 47.87% accuracy, 47% precision, 57% recall, and 51% F1-score. The accuracy of this study is relatively low compared to similar studies that use SVM in Twitter sentiment analysis. Study (Afrillia et al., 2022) on Permendikbudristek Number 30 of 2021 achieved an accuracy of 70.8% with a precision of 35.4%, recall of 50%, and an F1-score of 41.4%, while study (Husada & Paramita, 2021) on airline sentiment obtained an average accuracy of 80.41%, with the best results on the RBF kernel (precision of 84.33%, recall of 84.67%, and F1-score of 84.50%). This difference is likely due to the limitations of this study's dataset, which only consisted of 466 tweets, the presence of informal language variations, and the lack of hyperparameter optimization. In fact, hyperparameter optimization is important for obtaining the most suitable machine learning parameters (Khalid & Javaid, 2020). One technique that is widely used in this process is Grid Search (GS), which allows for the systematic selection of optimal parameter combinations (Joshua Agung Nurcahyo & Theopilus Bayu Sasongko, 2023). Nonetheless, the balanced distribution of positive and negative sentiments shows that public opinion on literacy is still

From the data processing results of twitter user tweets on literacy, Indonesian literacy, and publishers as sentiment responses, both positive and negative sentiments. **Figure 7** shows the visualization of the occurrence of words from positive sentiment using wordcloud.

[illegible]

Furthermore, based on **Figure 8**, the word with the largest size shows the highest frequency of word occurrence in the tweet data that has negative sentiment. The words that appear most frequently in this word cloud include 'literacy', 'literasiindonesia', 'indonesia', 'read', and 'book'. The appearance of these words shows that tweets with negative sentiments discuss literacy and books, but with a less positive context, for example related to the lack of literacy or people's complaints.

The word cloud visualization shows that the word “literacy” dominates both positive and negative sentiments. In positive sentiment, words such as “literacy,” “Indonesia,” “read,” and “book” appear in the context of promoting reading and literacy activities, while in negative

sentiment, the same words are used more in the context of complaints about low reading interest, limited access, or criticism of national literacy conditions. The dominance of the word “literacy” on both sides confirms that this issue is a major topic in the public sphere, with perceptions ranging from support to criticism. Therefore, a more in-depth analysis is needed, for example through topic modeling or sentiment intensity analysis, to understand the contextual meaning of each sentiment more comprehensively.

CONCLUSION

Based on an analysis of 466 tweets with the hashtags #literacy, #literacyindonesia, and #bookpublishing using the Support Vector Machine (SVM) algorithm, the results show that public perception is evenly divided between positive and negative sentiments. The SVM model with a linear kernel produced an accuracy of 47.87%, precision of 47%, recall of 57%, and an F1-score of 51%, confirming the limitations of SVM in analyzing literacy sentiment with a small dataset. The main contribution of this study is to provide an initial overview of public opinion regarding literacy on social media while opening up opportunities for further research with larger datasets and the application of hybrid algorithms. For future development, it is recommended to compare with other more adaptive methods such as ensemble learning (Random Forest, Gradient Boosting) and deep learning (LSTM, BERT), so that accuracy increases and sentiment interpretation becomes more representative.

REFERENCES

- Afrillia, Y., Rosnita, L., & Siska, D. (2022). Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine. *G-Tech: Jurnal Teknologi Terapan*, 6(2), 387–394. <https://doi.org/10.33379/gtech.v6i2.1778>
- Ahadiyah, F. N. (2024). Perkembangan Teknologi Infomasi Terhadap Peningkatan Bisnis Online. *INTERDISIPLIN: Journal of Qualitative and Quantitative Research*, 1(1), 41–49. <https://doi.org/10.61166/interdisiplin.v1i1.5>
- Azmi, N. A., Fathani, A. T., Sadayi, D. P., Fitriani, I., & Adiyaksa, M. R. (2021). Social Media Network Analysis (SNA): Identifikasi Komunikasi dan Penyebaran Informasi Melalui Media Sosial Twitter. *Jurnal Media Informatika Budidarma*, 5(4), 1422. <https://doi.org/10.30865/mib.v5i4.3257>
- Budiman, A. A., & Widiksono, S. (2018). *Aplikasi Pengolahan Data Untuk Menganalisa Penggunaan Hashtag Pada Twitter*. 8(2). <http://www.twitter.com>
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Edutic - Scientific Journal of Informatics Education*, 7(1), 1–11. <https://doi.org/10.21107/edutic.v7i1.8779>
- Edwin Kiky Aprianto, N. (2021). Peran Teknologi Informasi dan Komunikasi dalam Bisnis. *International Journal Administration, Business and Organization (IJABO)* |, 2(1), 1–7. <https://ijabo.a3i.or.id>
- Falenchia, M., & Sumardijati, S. (2023). Literasi Digital Remaja dalam Permainan Roleplay di Media Sosial Twitter. *Jiip - Jurnal Ilmiah Ilmu Pendidikan*, 6(6), 4064–4072. <https://doi.org/10.54371/jiip.v6i6.1585>
- Hadian, N., Hakim, M., & Fanani, M. R. (2023). Implementasi Model Service-Oriented Architecture (SOA) dalam Perancangan Sistem Informasi UMKM. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(3), 311–318. <https://doi.org/10.47233/jteksis.v5i3.879>
- Hendrawan, G. N., & Kusniyati, H. (2024). *CESS Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen Kendaraan Listrik di Media Sosial Twitter Evaluating the Performance of Naive*

- Bayes and SVM in Sentiment Analysis of Electric Vehicles on Twitter Social Media*. 9(1), 299–313. <https://doi.org/10.24114/cess.v9i1.54086>
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18–26. <https://doi.org/10.34148/teknika.v10i1.311>
- Joshua Agung Nurcahyo, & Theopilus Bayu Sasongko. (2023). Hyperparameter Tuning Algoritma Supervised Learning untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras. *The Indonesian Journal of Computer Science*, 12(3), 1351–1365. <https://doi.org/10.33022/ijcs.v12i3.3254>
- Khalid, R., & Javaid, N. (2020). A survey on hyperparameters optimization algorithms of forecasting models in smart grid. *Sustainable Cities and Society*, 61, 102275. <https://doi.org/10.1016/j.scs.2020.102275>
- Mailoa, F. F. (2021). Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia. *Journal of Information Systems for Public Health*, 6(1), 44. <https://doi.org/10.22146/jisph.44455>
- Munandar, A. A., Farikhin, F., & Widodo, C. E. (2023). Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM. *JOINTECS (Journal of Information Technology and Computer Science)*, 8(2), 77. <https://doi.org/10.31328/jointecs.v8i2.4747>
- Ningsih, A., & Indrawati, D. (2023). Penggunaan Leksikon Bahasa Indonesia Baku pada Lembaga dalam Pemartabatan Bahasa Indonesia di Ruang Publik (Studi Perkumpulan Ruang Belajar Aqil). *Sapala*, 10(2). <https://ejournal.unesa.ac.id/index.php/jurnal-sapala/article/view/54569>
- Ramlan, R., Satyahadewi, N., & Andani, W. (2023). Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM. *Jambura Journal of Mathematics*, 5(2), 431–445. <https://doi.org/10.34312/jjom.v5i2.20860>
- Safitri, T., Umaidah, Y., & Maulana, I. (2023). Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine. *Journal of Applied Informatics and Computing*, 7(1), 28–35. <https://doi.org/10.30871/jaic.v7i1.5039>
- Umar, R., Sunardi, S., & Ardiansyah, M. N. (2022). Penerapan Algoritma Support Vector Machine Pada Analisis Sentimen Hashtag Twitter. *JURIKOM (Jurnal Riset Komputer)*, 9(5), 1607. <https://doi.org/10.30865/jurikom.v9i5.4877>
- Widyanto, T., Ristiana, I., & Wibowo, A. (2023). Komparasi Naïve Bayes dan SVM Analisis Sentimen RUU Kesehatan di Twitter. *SINTECH (Science and Information Technology) Journal*, 6(3), 147–161. <https://doi.org/10.31598/sintechjournal.v6i3.1433>
- Yunita, R., & Kamayani, M. (2023). Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Penghapusan Kewajiban Skripsi. *Indonesian Journal of Computer Science*, 12(5), 2879–2890. <https://doi.org/10.33022/ijcs.v12i5.3415>
- Zusrotun, O. P., Murti, A. C., & Fiati, R. (2022). Analisis Sentimen Terhadap Belajar Online pada Media Sosial Twitter Menggunakan Algoritma Naive Bayes. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 11(3), 310–319. <https://doi.org/10.23887/janapati.v11i3.49160>