Journal of Intelligent Decision Support System (IDSS) 7 (3) (2024), pp. 227-235

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Comparison of Naïve Bayes Classifier and Support Vector Machine for sentiment analysis on civil military relations conflict among Rohingya refugees as recommendation for defense policy making

Nanda Selviana Putri[1], Hondor Saragih[2], Aulia Khamas Heikhmakhtiar[3]

[1,2,3]*Informatika, Universitas Pertahanan, Bogor, Indonesia*

Article Info

Article history:

Received Jul 30, 2024
Revised Aug 14, 2024
Accepted Aug 15, 2024

Keywords:

Civil military conflict;
Naïve bayes classifier;
Rohingya refugees;
Sentiment analysis;
Support vector machine.

ABSTRACT

This research focuses on the evaluating the performance of various sentiment analysis techniques using the Naive Bayes Classifier and Support Vector Machine in identifying civil-military conflicts among Rohingya refugees. The goal is to assist leaders in formulating defense policies. This research uses text data from news sources on Twitter, with a total of 5018 data that have been processed to become clean data, then divided into 1004 test data and 4018 training data to be classified using the Support Vector Machine and Naive Bayes methods. This research analyzes the sentiment and polarity of public opinion related to the issues that occur in this situation. The results of the sentiment analysis from the two methods are then classified using the Support Vector Machine and Naive Bayes methods, and then compared to determine which method is more effective in capturing the complex dynamics of sentiment. The findings of this research indicate that the Support Vector Machine method has a higher accuracy in identifying sentiments related to the civil-military conflict among Rohingya refugees, with an accuracy of 87.95%, compared to the Naive Bayes Classifier with an accuracy of 85.16%. The analysis results in the form of frequently occurring words in the true positive word cloud, namely apology, human, angry, and solidarity, are handed over to experts to be formulated into recommendation sentences and can be used to assist in the formulation of policies for defense decision-makers in more effectively addressing the Rohingya refugee issue.

Corresponding Author:

Nanda Selviana Putri,
Informatika,
Universitas Pertahanan,
Kawasan IPSC Sentul, Sukahati, Citereup, Bogor, 16810, Jawa Barat, Indonesia.
Email: nandaselvi87@gmail.com

Introduction

The research presented in this thesis focuses on sentiment analysis using Naïve Bayes Classifier and Support Vector Machine to study the civil-military conflict involving Rohingya refugees. This analysis aims to provide recommendations for defense decision-making. The Rohingya refugee crisis is an ongoing humanitarian issue, with allegations of human rights violations and conflicts between civilian and military forces. Sentiment analysis of social media data related to this conflict can help uncover public perceptions and inform policymaking.

The Rohingya refugee crisis is a complex and ongoing humanitarian issue that has drawn significant global attention. The conflict between civilian and military forces, along with allegations of human rights violations, has led to a dire situation for the Rohingya people. Understanding public sentiment towards this crisis is crucial for informing defense decision-making and policymaking.

Previous studies have explored the use of machine learning techniques, such as Naïve Bayes Classifier and Support Vector Machine, for sentiment analysis in various domains (Harun et al., 2023). These methods have proven effective in classifying text data as positive, negative, or neutral, providing insights into public perceptions (Rahayu et al., 2022). However, the application of these techniques to the specific context of the Rohingya refugee crisis has not been extensively investigated.

This research aims to fill this gap by conducting a comparative sentiment analysis using Naïve Bayes Classifier and Support Vector Machine to study the public's sentiment towards the civil-military conflict involving the Rohingya refugees. The findings from this analysis will provide valuable recommendations for defense decision-makers to better address the humanitarian concerns and civil-military tensions surrounding this crisis.

The key difference between this research and previous studies is the focus on the Rohingya refugee conflict, which presents unique challenges and nuances compared to other application domains. By leveraging the strengths of both Naïve Bayes Classifier and Support Vector Machine, this study aims to identify the most suitable approach for sentiment analysis in the context of the Rohingya crisis, ultimately contributing to more informed and effective defense policy formulation.

The findings from this comparative sentiment analysis study can then inform recommendations for defense decision-makers to better address the civil-military tensions and humanitarian concerns surrounding the Rohingya refugee situation. The insights gained can support more informed and effective defense policy formulation.

Method

This research is included in quantitative research. Quantitative research is a systematic approach that uses numerical data to study social issues and phenomena. This research involves the quantification and measurement of variables to understand the causal relationship in order to obtain a result from the sample (Haradhan, 2020). In this study, the authors collected data from Twitter, performed pre-processing on the data, and applied SVM and Naïve Bayes classification algorithms to analyze sentiment, then the classification results of each method were compared using a confusion matrix to determine which algorithm has the best accuracy results.

Quantitative research is more focused on in-depth understanding and interpretation of specific phenomena or contexts through descriptive data such as observation or content analysis. Qualitative research will focus more on understanding the perceptions and meanings found in the text or data collected (Yam & Taufik, 2021).

a. Data crawling

Data collection used in this research is derived from Twitter user reviews or opinions. Data collection is carried out through a crawling process using the tweet-harvest library on the Google Colab platform, and the results are saved in a .csv (Comma Separated Value) format for further analysis. The data crawling is conducted within a 3-month period from January 1, 2024, to March 1, 2024. The attributes used in the tweet dataset include the creation date, id, content, quote, reply, retweet, like, language, username of the creator, image URL, in reply, location, conversation ID, username, and tweet URL.

b. Pre-processing

The pre-processing stage involves labeling the tweets as positive, negative, or neutral, then tokenizing the text, filtering the tokens by removing stopwords, punctuation, and other irrelevant elements, applying stemming or lemmatization to reduce the dimensionality of the text data, and finally splitting the labeled dataset into training and testing subsets to be used for building and evaluating the sentiment analysis model. The dataset in this study has quite a large variation, so a 20:80 split can be sufficient to train the model and cover all the diversity provided, as stated in the research by Agung Nugroho and Agit Amrullah (Validation et al., 2023).

c. K-fold cross validation

The author validates the training and test data using K-Fold Cross Validation. The training data is divided into 10 folds, and the model is trained on the training data and evaluated on the validation data in each fold. The accuracy results on the validation and test data are then analyzed (Nada et al., 2023). The fold accuracy values range from 0.843 to 0.880, indicating the consistency of the model's performance. The test accuracy values range from 0.854 to 0.858, which is close to the average fold accuracy, suggesting that the model can generalize well to new data (Narkhede, 2019).

The ROC AUC value of around 0.857 shows that the model has good discriminative ability in differentiating between positive and negative classes. These results support the validity of the model and the quality of the dataset used in this research. Overall, the validation process demonstrates that the developed model is capable of providing accurate predictions for new data and does not suffer from overfitting. The results of k-fold processing can be seen in table 1 below

Table 1. K-fold results

Fold	Fold accuracy	Test accuracy
1	0.858	0.858
2	0.843	0.858
3	0.851	0.856
4	0.843	0.855
5	0.850	0.857
6	0.870	0.857
7	0.850	0.854
8	0.854	0.855
9	0.880	0.858
10	0.843	0.858
Average	0.854	0.857

Based on table 1, the authors can conclude that the dataset model is valid with the AUC (Area Under the Curve) and ROC (Receiver Operating Characteristic) parameters. The fold accuracy values range from 0.843 to 0.880, with an average of around 0.854, indicating consistent model performance during the K-Fold Cross Validation process, where the model can maintain a reasonably high and stable accuracy for each validation data fold. The test accuracy values range from 0.854 to 0.858, with an average of around 0.857, which is very close to the average fold accuracy, indicating that the model can generalize well on new (test) data. The AUC ROC value obtained from this data is around 0.857, where a value of 1 indicates a perfect model, and a value of 0.5 indicates a model no better than random guessing (Carrington et al., 2023). The result of 0.857 shows that the model has good discriminative ability in distinguishing positive and negative classes.

d. Classification

After the data goes through preprocessing and folding test, it will be classified using the SVM and Naïve Bayes methods. The classification stage using SVM and Naïve Bayes involves the use of trained models to predict the appropriate labels or categories for new data (Nugraha, 2022). For the SVM part, the authors implemented the SVM algorithm using the Python programming language and the sklearn library. Based on the program that was run, the SVM accuracy result is 0.8794820717131474 or 87.95%, indicating that the SVM model can correctly classify 87.95% of the total samples. The test results will be represented through a confusion matrix to show the model's performance on each class, not just the overall performance.

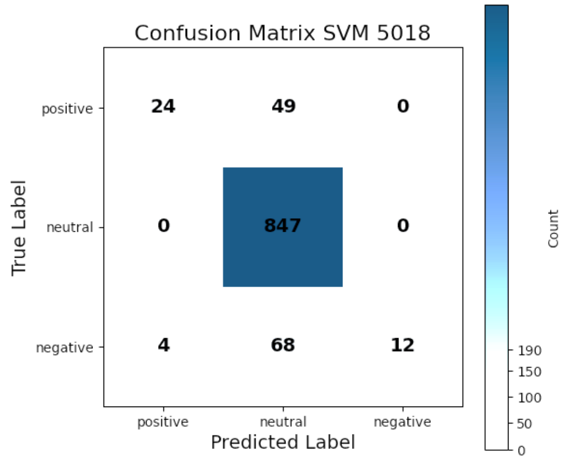

Figure. 1. The image shows a confusion matrix for a machine learning model using SVM method with 5018 data, displaying the predicted labels (positive, neutral, and negative) on the x-axis, the true labels on the y-axis.

Table 2. Classification report from SVM

	Precision	Recall	F1-score
Negative	0.86	0.33	0.48
Neutral	0.88	1.00	0.94
Positive	1.00	0.14	0.25
Accuration			0.88

Precision represents the ratio of correct predictions to total predictions for a class, recall represents the ratio of correct predictions to total actual samples for a class, and the f1-score is the average of the two (Yacouby & Axman, 2020).

Next step, authors classified the data using the Naïve Bayes method and obtained an accuracy value of 0.851593625498008 or 85.15%. The confusion matrix shows that the positive class has 3 correct predictions and 70 incorrect predictions, the neutral class has 847 correct predictions and 0 incorrect predictions, and the negative class has 5 correct predictions and 79 incorrect predictions. The classification report shows that the model has very good performance on the neutral class, but faces challenges on the negative and positive classes, resulting in low F1-scores (Fikri et al., 2020).

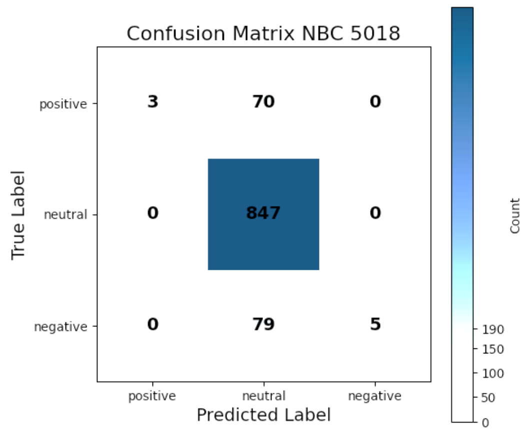

Figure 2. The image shows a confusion matrix for a machine learning model using Naïve Bayes method with 5018 data, displaying the predicted labels (positive, neutral, and negative) on the x-axis, the true labels on the y-axis (Malsi & Jatikusumo, 2022).

Table 3. Classification report from Naive Bayes

	Precision	Recall	F1-score
Negative	1.00	0.04	0.08
Neutral	0.85	1.00	0.92
Positive	1.00	0.06	0.11
Accuration			0.85

Table 3 show that the model correctly classifies 85% of the total samples. The model has very good performance on the neutral class, but there are constraints on the negative and positive classes, resulting in a low f1-score (Sriyano & Setiawan, 2021).

e. Evaluation

After obtaining the classification accuracy results, the next step is to compare the accuracy results of the two methods used.

Table 4. Comparison of classification results

Model	Accuration (%)	Precision	Recall	F1-score
SVM	87.95%	0.89	0.88	0.84
NBC	85.16%	0.87	0.85	0.79

Table 4 show that the overall classification accuracy using the SVM method is greater than the Naïve Bayes method. From these average results, it can be stated that the SVM method is the best method to be used in this research.

Table 5. Comparison of the number of sentiments

	Total sentiment		
	True positive	True neutral	True negative
SVM	24	847	12
NBC	3	847	5

Tabe 5 show a comparison of the classification of the number of sentiments that fall into the true positive, true neutral, and true negative categories by SVM and Naïve Bayes. It can be seen that the number of sentiments predicted by SVM is greater than those predicted by the Naïve Bayes method.

Results and Discussions

The dataset in this study is a collection of texts collected from Twitter using the crawling method, amounting to 5018 data. The data taken are only tweets in Indonesian with the words “militer Indonesia Rohingya”, “masyarakat Rohingya”, “kerugian Rohingya”, “mendukung Rohingya”, and “sipil mililter Rohingya” in the range from January 1, 2024 to March 1, 2024. The data was randomly selected from all Twitter social media users. The training data used in the classification consists of 4014 tweets and the test data is 1004 tweets.

Table 6. Data before and after going through pre-processing

Preliminary Data	Clean Data
Semua akun yg selama ini memberi disinformasi soal pengungsi Rohingya harus ditelusuri siapa penggeraknya. Akun anonim berjumlah besar + influencer yg selama ini ga pernah bicara isu kemanusiaan tetiba muncul hanya di isu Rohingya. Semoga kalian dapat balasannya di akhirat.	semua akun lama beri informasi soal pengungsi rohingya telusuri siapa penggerak akun anonim jumlah besar influencer lama pernah bicara manusia tetiba muncul hanya rohingya moga kalian dapat balasan akhirat

The differences between the data taken using the crawling method and the data that has gone through 4 filtering processes such as tokenization, token filtering, stemming, and stop word removal are shown. It is evident that the tweets that have undergone pre-processing are more concise and only retain the core words without affixes, which facilitates the classification process.

Based on table 4, each model achieves more than 85%, and the analysis shows that the most common sentiment is neutral sentiment. In this study, the accuracy level of the SVM method is higher at 87.95%, while Naïve Bayes is 85.16%. The explanation for each report is :

1. Accuration

The test results show that the SVM algorithm achieves an accuracy of 87.95%, while Naïve Bayes achieves an accuracy of 85.16%. This indicates that SVM has better classification capabilities than Naïve Bayes in the context of sentiment analysis of data related to the conflict among Rohingya refugees. This significant difference in accuracy can be caused by SVM's stronger ability to handle complex and nonlinear data, as well as SVM's superiority in carrying out optimal class separation (Valero-Carreras et al., 2023).

2. Precision

The precision value for SVM is 89%, while Naïve Bayes reaches 87%. A higher precision value on SVM shows that this algorithm has a better ability to identify positive instances accurately (Mehta & Pandya, 2020). This indicates that SVM can be more effective in predicting instances that are truly relevant as positive sentiment compared to Naïve Bayes (Riyadi et al., 2022).

3. Recall

The SVM algorithm has a recall value of 88%, while Naïve Bayes reaches 85%. The higher recall value on SVM shows that this algorithm can find complex patterns and flexibility in parameter settings (Patel et al., 2023).

4. F1-score

Based on calculations, SVM obtained an F1-score of 84%, while Naïve Bayes reached 79%. A higher F1-score value for SVM indicates that this algorithm has better overall performance, with the ability to combine superior precision and recall (Santoso et al., 2023).

The following figures are the results of word cloud processing for each sentiment :

Figure 3. Word cloud TP SVM

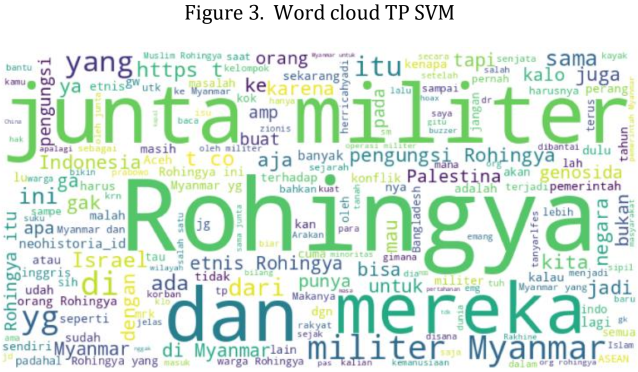

Figure 4. Word cloud TNt SVM

Figure 5. Word cloud TN SVM

Overall, the research results show that the Support Vector Machine (SVM) algorithm has superior performance compared to the Naive Bayes Classifier (NBC) in conducting sentiment analysis on data related to military-civil relations conflicts among Rohingya refugees. This significant difference in performance can have important implications for decision makers in the defense sector.

This research is in line with previous studies by Dinar Ajeng (Kristiyanti et al., 2019), Uning (U. Lestari et al., 2022), Dwi Normawati and Prayogi (Normawati & Prayogi, 2021), Mujaddid Izzul Fikri (Fikri et al., 2020), as well as research by Kusmawati (Kusumawati et al., 2019) because the results of this study add evidence supporting the findings in previous research related to the effectiveness of the Support Vector Machine method in analyzing sentiments related to conflicts. In line with previous studies, this research shows that Support Vector Machine has higher accuracy compared to Naive Bayes Classifier in identifying complex public sentiment dynamics on issues. These findings strengthen the recommendation to use the Support Vector Machine approach to assist in formulating defense policies that are responsive to the Rohingya refugee issue.

The findings of the comparative sentiment analysis study between the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) algorithms show that the SVM algorithm outperforms NBC in terms of accuracy, precision, recall, and F1-score. The higher accuracy of 87.95% achieved by SVM indicates its superior classification capabilities compared to the 85.16% accuracy of NBC in the context of this specific application domain. The difference in performance can be attributed to SVM's stronger ability to handle complex and nonlinear data, as well as its superiority in performing optimal class separation.

These findings have important implications for defense decision-makers in formulating policies to address the Rohingya refugee crisis. The superior performance of the SVM algorithm suggests that it should be the preferred approach for sentiment analysis in this domain, as it can provide more accurate and reliable insights to support informed policymaking. By leveraging the insights gained from this comparative study, defense decision-makers can better understand the public's sentiment towards the Rohingya crisis and develop more responsive and effective policies to address the underlying issues, which can help in mitigating the civil-military conflict and promoting a more harmonious resolution to this complex humanitarian crisis.

Conclusions

The findings of this comparative sentiment analysis study demonstrate the superior performance of the SVM algorithm over the NBC in the context of analyzing Twitter data related to the civil-military relations issue and the humanitarian crisis of Rohingya refugees. The SVM algorithm achieved notably higher accuracy of 87.95%, precision of 89%, recall of 88%, and F1-score of 84% compared to the NBC method that only has 85.16% of accuracy, indicating its stronger capabilities in sentiment classification. The higher accuracy of SVM can be attributed to its ability to effectively handle complex and nonlinear data patterns, as well as its superiority in performing optimal class separation. This suggests that the SVM approach is better equipped to capture the nuanced and multifaceted sentiments expressed by the public on social media regarding the Rohingya crisis and civil-military tensions. The superior precision and recall of SVM further reinforce its effectiveness in accurately identifying positive and negative sentiment instances, which is crucial for policymakers to gain a comprehensive understanding of public perceptions. Sentiment classification of Twitter data related to the civil-military relations issue and the humanitarian crisis of Rohingya refugees can be performed using SVM and NBC algorithms. These algorithms can be utilized to analyze the polarity of sentiments, categorizing them as positive, negative, or neutral. Given the higher accuracy of SVM compared to NBC, the SVM algorithm can be considered a more suitable method for aiding leadership in policy formulation for defense matters.

References

Carrington, A. M., Manuel, D. G., Fieguth, P. W., Ramsay, T., Osmani, V., Wernly, B., Bennett, C., Hawken, S., Magwood, O., Sheikh, Y., Mcinnes, M., & Holzinger, A. (2023). Deep ROC Analysis and AUC as Balanced Average Accuracy, for Improved Classifier Selection, Audit and Explanation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(1), 329–341. https://doi.org/10.1109/TPAMI.2022.3145392

Elsayed, F. E. (2020). Social Media Role in Relieving the Rohingya Humanitarian Crisis. *New Media and Mass Communication*, *87*, 28–48. https://doi.org/10.7176/nmmc/87-04

Fikri, M. I., Sabrila, T. S., Azhar, Y., & Malang, U. M. (2020). Comparison of the Naïve Bayes Method and Support Vector Machine on Twitter Sentiment Analysis. *SMATIKA Jurnal: STIKI Informatika Jurnal*, *10*(2), 71–76.

Haradhan, M. (2020). Quantitative Research: A Successful Investigation in Natural and Social Sciences. *Journal of Economic Development, Environment and People*, *9*(4), 52–79. https://mpra.ub.uni-muenchen.de/105149/

Harun, N. A., Huspi, S. H., & A. Iahad, N. (2023). Question Classification for Helpdesk Support Forum Using Support Vector Machine and Naïve Bayes Algorithm. *International Journal of Innovative Computing*, *13*(1), 37–45. https://doi.org/10.11113/ijic.v13n1.388

Kristiyanti, D. A., Umam, A. H., Wahyudi, M., Amin, R., & Marlinda, L. (2019). Comparison of SVM Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter. *2018 6th International Conference on Cyber and IT Service Management, CITSM 2018*, *Citsm 2018*, 1–6. https://doi.org/10.1109/CITSM.2018.8674352

Kusumawati, R., D'Arofah, A., & Pramana, P. A. (2019). Comparison Performance of Naive Bayes Classifier and Support Vector Machine Algorithm for Twitter's Classification of Tokopedia Services. *Journal of Physics: Conference Series*, *1320*(1), 1–11. https://doi.org/10.1088/1742-6596/1320/1/012016

Lestari, M. I., & Anggraeni, D. (2021). Analisis dampak sentimen masyarakat selama pandemi covid-19 terhadap kurs rupiah (Studi kasus pandemi covid-19 di Indonesia). *Jurnal EMBA*, *9*(1), 1–14.

Lestari, U., Romadhani, T., Suraya, S., & Fatkhiyah, E. (2022). Sentiment Analysis for Extracting Student Opinion Data on Higher Education Services Using the Naive Bayes Classifier and Support Vector Machine Methods (Case Study Akprind Institute of Science and Technology Yogyakarta). *Jurnal TAM (Technology Acceptance Model)*, *13*(1), 51. https://doi.org/10.56327/jurnaltam.v13i1.1220

Malsi, E., & Jatikusumo, D. (2022). Analisis Sentimen Terhadap Ulasan Aplikasi FLIP.ID Menggunakan Klasifikasi Naïve Bayes. *Jurnal Informatika dan Teknologi Informasi*, *18*(1), 1–11.

Mehta, P., & Pandya, S. (2020). A review on sentiment analysis methodologies, practices and applications. *International Journal of Scientific and Technology Research*, *9*(2), 601–609.

Nada, D. D., Soehardjoepri, S., & Atok, R. M. (2023). Perbandingan Analisis Sentimen Mengenai BPJS pada Media Sosial Twitter Menggunakan Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). *Jurnal Sains dan Seni ITS*, *11*(6). https://doi.org/10.12962/j23373520.v11i6.96330

Narkhede, S. (2019). *Understanding AUC - ROC Curve*. 6–11.

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *5*(2), 697–711.

Nugraha, A. F. (2022). Naïve Bayes dan Support Vector Machine Berbasis PSO untuk Seleksi Fitur pada Sentiment Analysis. *Innovation in Research of Informatics (INNOVATICS)*, *4*(2), 56–61. https://doi.org/10.37058/innovatics.v4i2.5291

Patel, T. S., Patel, D. P., Sanyal, M., & Shrivastav, P. S. (2023). *Prediction of Heart Disease and Survivability using Support Vector Machine and Naive Bayes Algorithm*. *M*. http://dx.doi.org/10.1101/2023.06.09.543776%0Ahttps://syndication.highwire.org/content/doi/10.1101/2023.06.09.543776

Rahayu, A. S., Fauzi, A., & Rahmat, R. (2022). Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify. *Jurnal Sistem Komputer dan Informatika (JSON)*, *4*(2), 349. https://doi.org/10.30865/json.v4i2.5398

Riyadi, S., Siregar, M. M., Margolang, K. fadhli F., & Andriani, K. (2022). Analysis of Svm and Naive Bayes Algorithm in Classification of Nad Loans in Save and Loan Cooperatives. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, *8*(3), 261–270. https://doi.org/10.33330/jurteksi.v8i3.1483

Santoso, I., Oktora, S. I., Muchlisoh, S., & Pasaribu, E. (2023). Social Network Analysis untuk Identifikasi Pengguna Twitter Berpengaruh pada Topik Bencana Gempa dan Tsunami di Indonesia. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, *9*(1), 115. https://doi.org/10.26418/jp.v9i1.62211

Saputra, A., Subing, M., & Pratama, R. (2023). Perbandingan Metode Naïve Bayes Classifier Dan Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Mengenai Piala Dunia Fifa 2022. *Teknomatika*, *13*(01), 22–31.

Sriyano, C. S., & Setiawan, E. B. (2021). Pendeteksian Berita Hoax Menggunakan Naive Bayes Multinomial Pada Twitter dengan Fitur Pembobotan TF-IDF. *e-Proceeding of Engineering : Vol.8, No.2*, *8*(2), 3396–3405.

Valero-Carreras, D., Alcaraz, J., & Landete, M. (2023). Comparing two SVM models through different metrics based on the confusion matrix. *Computers and Operations Research*, *152*(April 2022), 106131. https://doi.org/10.1016/j.cor.2022.106131

Validation, C., Galih, K. S. P., Galih, K. S. P., Validation, C., & Kunci, K. (2023). *1 1 , 2**. *5*(2), 294–300.

Yacouby, R., & Axman, D. (2020). *Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models*. 79–91. https://doi.org/10.18653/v1/2020.eval4nlp-1.9

Yam, J. H., & Taufik, R. (2021). Hipotesis Penelitian Kuantitatif. *Perspektif : Jurnal Ilmu Administrasi*, *3*(2), 96–102. https://doi.org/10.33592/perspektif.v3i2.1540