ISSN 2301-8984 (Print), 2809-1884 (Online)
Published by Institute of Computer Science

Comparison of random forest and SVM methods in sentiment analysis about electric cars in Indonesia

Indra Pratistha[1], Adi Panca Saputra Iskandar[2], Eugenius Gene Rangga Lanang[3], Ni Wayan Jeri Kusuma Dewi[4]
[1,2,3,4]Informatics, Institut Bisnis dan Teknologi Indonesia, Indonesia

ARTICLE INFO

Article history:

Received Jul 1, 2025
Revised Jul 9, 2025
Accepted Jul 15, 2025

Keywords:

Electric Vehicles;
Random Forest Method;
Sentiment Analysis;
SVM Method.

ABSTRACT

This study examined public sentiment toward electric vehicles (EVs) in Indonesia, where the adoption of EVs reached 28,188 registered units in 2023. The research analyzed user-generated content from the social media platform X (formerly known as Twitter), collecting 1,507 tweets that underwent preprocessing, including text normalization and sentiment labeling. Two machine learning models, Random Forest and Support Vector Machine (SVM), were implemented to classify the tweets into positive and negative sentiments. Each model was evaluated under three experimental scenarios with varying training dataset sizes. The results indicated that the SVM model achieved the best performance in the third scenario, with an accuracy of 81.3%, precision of 88%, and recall of 91%. In comparison, Random Forest achieved its highest results in the same scenario, with an accuracy of 77%, precision of 91%, and recall of 81%. These findings demonstrated that SVM outperformed Random Forest in terms of overall balance between accuracy and recall, making it the more effective model for sentiment classification in this context.

Corresponding Author:

Indra Pratistha,
Informatics,
Institut Bisnis dan Teknologi Indonesia,
Jl. Tukad Pakerisan No.97, Panjer, Denpasar Selatan, Kota Denpasar, Bali 80225, Indonesia
Email: indra.pratistha@instiki.ac.id

1. INTRODUCTION

The rapid development of information and communication technology has significantly influenced various sectors, including the automotive industry. One of the major innovations is the emergence of electric vehicles (EVs), which are considered environmentally friendly alternatives to traditional fossil-fueled cars. In Indonesia, the adoption of electric cars has been growing steadily. According to data from the Ministry of Transportation of the Republic of Indonesia, the number of registered electric vehicles reached 28,188 units in 2023, indicating increasing public interest and government support for green transportation solutions (Seraphina & Gunawan, 2025).

Despite the benefits offered by electric vehicles, such as zero emissions, lower operational costs, and reduced dependence on fossil fuels, their adoption has sparked a wide range of public opinions. Social media platforms, particularly X (formerly known as Twitter), have become essential channels where individuals express their views, share experiences, and discuss new technologies (Tarigan & Yusupa, 2024). As a result, analyzing user-generated content on these platforms can provide valuable insights into public sentiment toward electric cars.

The urgency of this study is not merely driven by global technological trends, but by specific and pressing challenges within Indonesia's automotive landscape. Issues such as the high initial cost of EVs, limited charging infrastructure, uncertainty regarding battery performance, and public skepticism toward government policies on EV incentives remain significant barriers to widespread adoption. These concerns are frequently voiced in public forums, especially on social media, making it essential to understand how the public perceives these issues. Without addressing these perception gaps, the government's roadmap toward electric vehicle adoption may encounter resistance or fail to align with societal expectations.

Sentiment analysis, also known as opinion mining, is a subfield of Natural Language Processing (NLP) that focuses on extracting subjective information from text. According to Liu (2012), sentiment analysis involves classifying opinions expressed in a piece of text as positive, negative, or neutral, and is widely applied in domains such as customer feedback, political analysis, and social behavior studies. It serves as a powerful tool for uncovering the underlying attitudes of users on social media (Ahsan, 2021).

In this context, machine learning algorithms play a central role in automating sentiment classification. Support Vector Machine (SVM) is a supervised learning algorithm introduced by Vapnik (1995) that is particularly effective in high-dimensional spaces and is commonly used in text classification tasks due to its ability to find the optimal separating hyperplane. Meanwhile, Random Forest (RF), introduced by Breiman (2001), is an ensemble method that builds multiple decision trees and aggregates their predictions for improved accuracy and reduced overfitting.

This research aims to evaluate the effectiveness of both models by comparing their accuracy, precision, and recall across different experimental scenarios. By identifying the more suitable algorithm for sentiment classification in this context, the findings aim to contribute to the development of better analytical tools for policymakers, automotive industry stakeholders, and researchers interested in public perception analysis.

2. RESEARCH METHOD

This study is an experimental quantitative study that uses a supervised machine learning approach to conduct a sentiment analysis of public opinion regarding electric cars in Indonesia (Sivaranjani, 2021).

Data Collection

The dataset used in this study was collected from the social media platform X (formerly known as Twitter) using keyword-based crawling with the term "mobil listrik" (electric car). Tweets were gathered from September 19, 2022, to August 6, 2023, resulting in 1,507 unique tweets. This dataset was processed to ensure relevance, removing retweets, duplicates, and irrelevant content.

While data collected from social media platforms like X allows real-time access to spontaneous public opinion, it is important to acknowledge that it may not fully represent the broader Indonesian population. Twitter users tend to be younger, urban-based, and more tech-savvy, which can skew the sentiment distribution compared to that of rural or less digitally active populations. Therefore, the results of this study primarily reflect the perspectives of the digital community actively engaging with EV discourse online, rather than a comprehensive cross-section of all social segments. Nonetheless, since this group often includes early adopters and influencers of public opinion, their sentiment can be considered a valuable indicator of emerging trends in public perception.

Data Preprocessing

Before classification, the raw tweets underwent several preprocessing steps to enhance the quality of textual data: a) Case Folding: Converted all text to lowercase for consistency; b) Normalization: Transformed informal or slang words into standard Indonesian; c) Stopword Removal: Removed common Indonesian stopwords using a predefined stopword list; d) Tokenization: Split text into individual tokens (words); e) Stemming: Reduced words to their root form using the Sastrawi stemmer for Bahasa Indonesia.

Sentiment Labeling

Sentiment labeling was performed using the IndoBERT-base-p1 model, a pre-trained BERT model optimized for Indonesian language tasks (Tan, 2021). The model classified tweets into three sentiment categories: positive, negative, and neutral. However, for this study, only positive and negative classes were retained, and neutral data (6 tweets) were excluded to improve classification clarity.

Feature Extraction

The term frequency-inverse document frequency (TF-IDF) method was applied to convert textual data into numerical features suitable for machine learning models. This technique assigns weights to each token based on its frequency in a tweet relative to its occurrence across the corpus, thus emphasizing informative words.

Classification Models

This study compared two supervised machine learning algorithms: a) Support Vector Machine (SVM): A linear classifier that seeks to find the optimal hyperplane separating two classes (positive and negative sentiments); b) Random Forest (RF): An ensemble learning method that builds multiple decision trees and merges their outputs through majority voting for classification tasks; c) Both models were implemented using the Scikit-learn library in Python, with default hyperparameters and "*class_weight='balanced'*" to address the class imbalance.

Experimental Design

Three experimental scenarios were designed to evaluate model performance under varying training set sizes: a) Scenario 1: 600 training samples (300 positive, 300 negative); b) Scenario 2: 800 training samples (400 positive, 400 negative); c) Scenario 3: 1,000 training samples (500 positive, 500 negative)

Evaluation Metrics

The models were evaluated using standard classification metrics derived from the confusion matrix: a) Accuracy: Proportion of correctly predicted instances overall predictions; b) Precision: Proportion of true positives among all predicted positives; c) Recall: Proportion of true positives among all actual positives; d) F1-Score: Harmonic mean of precision and recall, indicating overall balance.

3. RESULTS AND DISCUSSIONS

The process of sentiment classification using the Support Vector Machine (SVM) method is illustrated in Figure 1. The flowchart outlines the sequential steps involved in transforming raw tweet data into predictive sentiment outcomes. Initially, tweet texts are converted into numerical feature vectors using the TF-IDF (Term Frequency–Inverse Document Frequency) technique, which emphasizes important terms while minimizing common, less informative words. The dataset is then subjected to 10-fold cross-validation to ensure robust evaluation and minimize bias.

The data is divided into training and testing subsets, where the training data is used to build the SVM model. During the training phase, the algorithm learns to distinguish between positive and negative sentiments based on labeled data. Once the model is trained, it is applied to the test data to generate predictions. This systematic process ensures that the SVM model is both accurate and generalizable for classifying sentiment in Indonesian-language tweets.

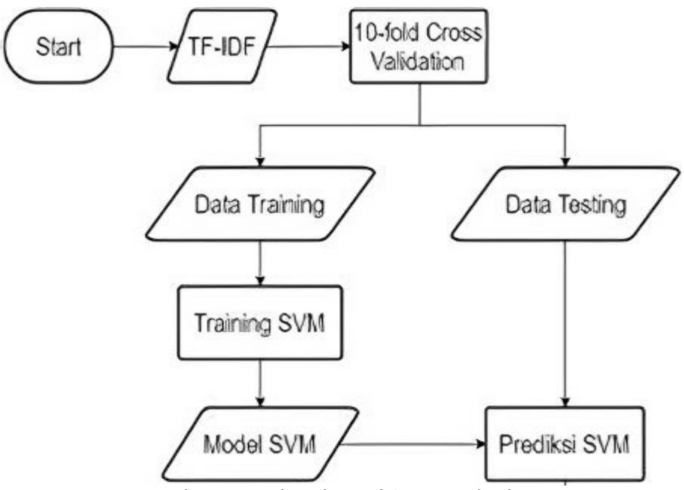

Figure 1. Flowchart of SVM method

This section presents the experimental results of sentiment classification using Support Vector Machine (SVM) and Random Forest (RF) models. Three training scenarios were conducted to evaluate the performance of each model with varying dataset sizes: 600, 800, and 1,000 training samples. Evaluation metrics used include accuracy, precision, recall, and F1-score.

Scenario 1: 600 Training Samples

In the first scenario, both models were trained using 600 labeled tweets (300 positive and 300 negative) and tested on 214 unseen tweets.

Table 1. Scenario 1: 600 training samples

Model	Accuracy	Precision (POS)	Recall (POS)	F1-Score (POS)
SVM	77.0%	33%	26%	29%
Random Forest	70.0%	29%	79%	42%

The results in Table 1 reveal differing strengths between the two models (Alabi, 2021). The Support Vector Machine (SVM) achieved a higher accuracy (77.0%) compared to the Random Forest (70.0%), indicating better overall classification correctness on the test set. However, when analyzing the recall, which reflects the model's ability to detect actual positive sentiments, Random Forest significantly outperformed SVM, with 79% recall versus SVM's 26%.

This suggests that Random Forest was more sensitive to identifying positive sentiment tweets, an advantage in contexts where missing positive opinions may be costly. However, this came at the expense of lower precision (29%), meaning many tweets predicted as positive were actually negative. In contrast, SVM showed higher precision (33%), but its low recall implies it missed a large portion of positive tweets. The F1-Score, which balances precision and recall, reflects this trade-off: Random Forest scored 42%, while SVM scored 29%. Despite SVM having higher accuracy, its lower recall dragged down the F1 score. This indicates that Random Forest may be preferable in applications that prioritize capturing as many positive sentiments as possible, while SVM may be more conservative in predicting positive cases (Amjad, 2022; Tang, 2020).

Scenario 2: 800 Training Samples

In this scenario, models were trained on 800 samples (400 positive and 400 negative). The results are summarized below:

Table 2. Scenario 2: 800 training samples

Model	Accuracy	Precision (POS)	Recall (POS)	F1-Score (POS)
SVM	78.0%	43%	81%	56%
Random Forest	72.0%	40%	69%	51%

In the second scenario, both models were trained with 800 labeled tweets (400 positive and 400 negative), and their performance improved significantly compared to the previous scenario with 600 samples (Liang, 2020).

The Support Vector Machine (SVM) again demonstrated higher accuracy (78.0%) than the Random Forest (72.0%), indicating a more consistent classification across the expanded dataset (Sarkar, 2023). More significantly, SVM achieved a substantial gain in recall (81%), suggesting it successfully identified the majority of actual positive tweets. This marks a dramatic improvement from its 26% recall in Scenario 1, likely due to the increased volume and diversity of training data. In terms of precision, SVM scored 43%, slightly higher than Random Forest's 40%. While neither model achieved very high precision, the balance between precision and recall allowed SVM to attain a better F1 score (56%) compared to Random Forest's 51%. These results indicate that SVM benefits more from increased training data, especially in learning patterns that distinguish positive sentiments. Random Forest also showed improvement in all metrics, but its performance gains were relatively smaller.

Scenario 3: 1,000 Training Samples

In the final scenario, 1,000 training tweets (500 positive, 500 negative) were used, and the test set remained the same.

Table 3. Scenario 3: 1000 training samples

Model	Accuracy	Precision (POS)	Recall (POS)	F1-Score (POS)
SVM	81.3%	88%	91%	89%
Random Forest	77.0%	91%	81%	86%

In the third and final scenario, both models were trained with the largest dataset of 1,000 tweets (500 positive and 500 negative) (Olisah, 2022). This scenario yielded the highest performance across all metrics for both models, clearly demonstrating the positive impact of increasing training data on classification accuracy and robustness.

The Support Vector Machine (SVM) achieved its best results here, with an impressive accuracy of 81.3%, precision of 88%, recall of 91%, and F1-score of 89%. These figures suggest that SVM not only classified tweets accurately overall but also effectively identified positive sentiment while maintaining a low false positive rate. Random Forest, although slightly behind SVM in overall accuracy (77.0%), showed strong precision (91%), even surpassing SVM in this metric. This means that among all the tweets classified as positive, the vast majority were indeed positive. However, its recall (81%) was lower than SVM's, indicating that it missed more actual positive tweets (Li, 2020; Rani, 2023; Tetila, 2020). The F1-score, which harmonizes precision and recall, was also strong for both models, 86% for Random Forest and 89% for SVM. These high scores confirm that both models are highly capable in this scenario, though SVM holds a slight overall edge, particularly in capturing a higher proportion of true positive sentiments.

Comparative Analysis

To further evaluate the real-world applicability of both models, a comparative prediction test was conducted using 300 unseen tweets. The goal was to observe the distribution of sentiment predictions by each model when applied to the same test data (Bressan, 2020; Fang, 2020). The Random Forest model classified 283 tweets as negative and 17 as positive, indicating a strong bias toward negative sentiment yet still showing some sensitivity to positive expressions. This result is visualized in Figure 2, which illustrates the number of tweets assigned to each class. The X-axis represents the sentiment labels Negative and Positive, while the Y-axis denotes the number of predictions per class (Nayarisseri, 2021; Singh, 2022). The use of distinct bar colors enhances visual clarity, and numeric labels above each bar indicate the exact count.

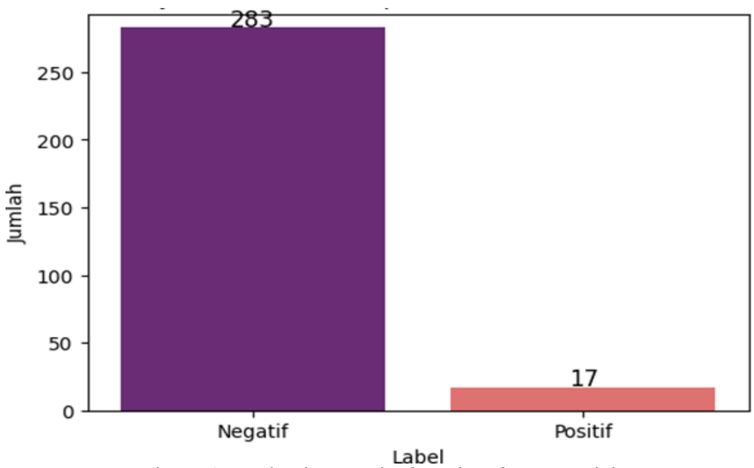

Figure 2. Evaluation result of random forest model

Conversely, the SVM model exhibited even stronger sensitivity toward negative sentiment. It predicted 298 tweets as negative and only 2 tweets as positive, as shown in Figure 3. This suggests that while SVM is more conservative in assigning positive sentiment, it may also underrepresent minority sentiment classes if the data distribution is heavily imbalanced. The high recall observed in earlier scenarios aligns with this outcome, as the model tends to group tweets more confidently into the dominant class (Guo, 2021; Mohajane, 2021; Yazdinejad, 2023).

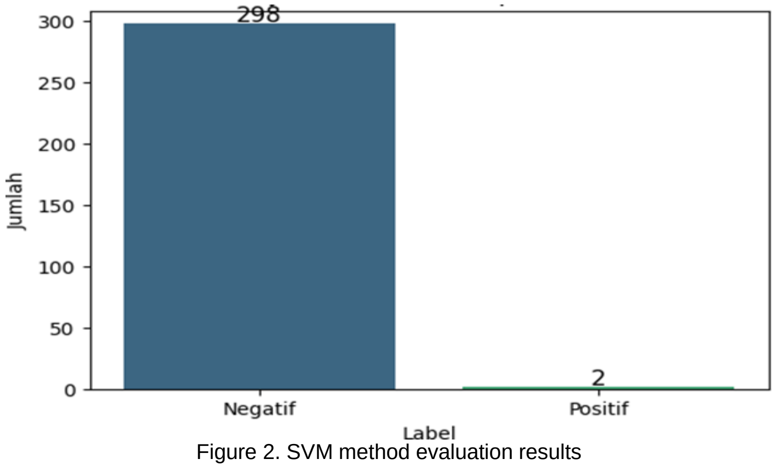

Figure 2. SVM method evaluation results

These results reflect not only the statistical metrics but also reveal the behavioral tendencies of the models. While Random Forest retains a moderate capacity to detect both classes, SVM prioritizes accuracy in the dominant class, potentially at the cost of missing minority-class instances. This is crucial for applications where identifying minority sentiments (e.g., positive feedback in a mostly negative corpus) is strategically important (Dutta, 2020).

Discussions

The comparative results of the three experimental scenarios demonstrate a clear trend in model performance as the size of the training dataset increases. Both Support Vector Machine (SVM) and Random Forest (RF) showed improvement in accuracy, precision, recall, and F1-score as more training data was used. However, the rate and consistency of improvement varied between the two models (Kim, 2020; Prasad, 2020).

In Scenario 1 with 600 training samples, SVM slightly outperformed RF in accuracy (77% vs. 70%) but struggled in terms of recall, indicating its conservative nature in predicting positive sentiments. On the other hand, Random Forest achieved a high recall (79%) but suffered from lower precision (29%), showing a tendency to over-classify positive sentiments (Fitriyani, 2020; Ishaq, 2021; Perez-Diaz, 2020). In Scenario 2, with 800 training samples, both models improved. SVM exhibited a better balance between precision and recall (43% and 81%), while RF also improved but still lagged slightly behind in the F1 score. This suggests that SVM is more effective at leveraging additional training data, possibly due to its strength in handling high-dimensional text features with a clear decision boundary.

Scenario 3 marked a significant leap in performance. SVM reached its peak performance, with an accuracy of 81.3%, precision of 88%, recall of 91%, and F1-score of 89%. In contrast, Random Forest also showed strong results, with a slightly higher precision (91%) but lower recall (81%). This difference indicates that while Random Forest is highly accurate when it predicts a tweet as positive, it misses more actual positive instances than SVM (Ghorbani, 2020; Liu, 2021).

From the final evaluation of unseen test data (visualized in Figures 4.20 and 4.21), Random Forest predicted 283 tweets as negative and 17 as positive, while SVM predicted 298 as negative and only 2 as positive. These results reflect SVM's stronger confidence in classifying tweets as negative, possibly due to class imbalance during training or more conservative thresholding. The bar charts presented earlier also emphasize this disparity in sentiment distribution, highlighting the trade-off between sensitivity and precision.

Overall, while both models are capable of sentiment classification, SVM consistently demonstrated better generalization and balance across evaluation metrics, making it more suitable for sentiment analysis tasks where balanced detection of both sentiment classes is critical (Alzamzami, 2020; Saleh, 2021; Shahriar, 2021). Meanwhile, Random Forest may be preferable in scenarios where precision for positive predictions is prioritized over recall.

4. CONCLUSION

This study set out to analyze public sentiment toward electric vehicles in Indonesia using two machine learning methods, Support Vector Machine (SVM) and Random Forest (RF), based on tweet data collected from the social media platform X. As described in the Introduction, the primary objective was to determine which model achieves better performance in classifying sentiments into positive and negative categories, using varying sizes of training data. The results presented in the Results and Discussion section confirm that SVM outperformed Random Forest across all three experimental scenarios, particularly when the size of the training dataset was increased to 1,000 samples. In that final scenario, SVM achieved the highest performance with 81.3% accuracy, 88% precision, 91% recall, and an F1-score of 89%, whereas Random Forest achieved slightly lower results despite a strong precision score.

These findings demonstrate that SVM is more effective and stable in handling sentiment classification tasks for Indonesian-language tweets, especially when data is properly preprocessed and the feature space is well-structured using TF-IDF. The conservative nature of SVM contributes to its lower false-positive rate, making it more reliable in balanced classification tasks. In alignment with the research objectives, the experiment proves that the choice of model significantly influences the quality of sentiment analysis results. The compatibility between the problem formulation and experimental outcomes confirms that the SVM model is more suitable for public sentiment monitoring regarding electric car adoption in Indonesia.

Beyond the technical comparison, this study makes a practical contribution to advancing sentiment analysis in the Indonesian context. It demonstrates that traditional machine learning models such as SVM—when carefully applied to Indonesian-language data with appropriate preprocessing—can achieve high levels of accuracy and generalizability. This supports the adoption of localized sentiment analysis tools that are cost-effective, interpretable, and implementable without requiring large-scale computational resources. Additionally, this research highlights the need to consider linguistic and cultural nuances unique to Indonesian users, thereby encouraging further development of sentiment models tailored specifically to Bahasa Indonesia and local digital behaviors.

This research opens up several avenues for further study. First, future work could explore deep learning models such as LSTM or transformers (e.g., IndoBERT fine-tuning) to improve sentiment detection on more complex or ambiguous expressions. Second, expanding the dataset to include other platforms, such as Instagram or news comments, could offer a more comprehensive view of public opinion. Third, the use of multiclass classification, including neutral sentiment, which was omitted in this study, may provide richer insights into nuanced public perception. Practically, the findings can be used by automotive companies, environmental policymakers, and marketing analysts in Indonesia to monitor sentiment trends and develop responsive strategies. With the growing demand for cleaner transportation, sentiment analysis models like SVM can serve as reliable tools to assess public readiness and acceptance of new technologies.

REFERENCES

Ahsan, M. M. (2021). Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. *Technologies*, *9*(3). https://doi.org/10.3390/technologies9030052

Alabi, R. O. (2021). Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. In *Artificial Intelligence in Medicine* (Vol. 115). https://doi.org/10.1016/j.artmed.2021.102060

Alzamzami, F. (2020). Light Gradient Boosting Machine for General Sentiment Classification on Short Texts: A Comparative Evaluation. *IEEE Access*, *8*, 101840–101858. https://doi.org/10.1109/ACCESS.2020.2997330

Amjad, M. (2022). Prediction of Pile Bearing Capacity Using XGBoost Algorithm: Modeling and Performance Evaluation. *Applied Sciences Switzerland*, *12*(4). https://doi.org/10.3390/app12042126

Bressan, T. S. (2020). Evaluation of machine learning methods for lithology classification using geophysical data. *Computers and Geosciences*, *139*. https://doi.org/10.1016/j.cageo.2020.104475

Dutta, A. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, *159*. https://doi.org/10.1016/j.eswa.2020.113408

Fang, Z. (2020). Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Computers and Geosciences*, *139*. https://doi.org/10.1016/j.cageo.2020.104470

Fitriyani, N. L. (2020). HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System. *IEEE Access*, *8*, 133034–133050. https://doi.org/10.1109/ACCESS.2020.3010511

Ghorbani, R. (2020). Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques. *IEEE Access*, *8*, 67899–67911.

https://doi.org/10.1109/ACCESS.2020.2986809
Guo, Y. (2021). Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, *120*. https://doi.org/10.1016/j.ecolind.2020.106935
Ishaq, A. (2021). Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques. *IEEE Access*, *9*, 39707–39716. https://doi.org/10.1109/ACCESS.2021.3064084
Kim, Y. (2020). Gap-filling approaches for eddy covariance methane fluxes: A comparison of three machine learning algorithms and a traditional method with principal component analysis. *Global Change Biology*, *26*(3), 1499–1518. https://doi.org/10.1111/gcb.14845
Li, Y. (2020). Do we really need deep CNN for plant diseases identification? *Computers and Electronics in Agriculture*, *178*. https://doi.org/10.1016/j.compag.2020.105803
Liang, Y. C. (2020). Machine learning-based prediction of air quality. *Applied Sciences Switzerland*, *10*(24), 1–17. https://doi.org/10.3390/app10249151
Liu, L. (2021). Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning. *IEEE Access*, *9*, 7550–7563. https://doi.org/10.1109/ACCESS.2020.3048198
Mohajane, M. (2021). Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area. *Ecological Indicators*, *129*. https://doi.org/10.1016/j.ecolind.2021.107869
Nayarisseri, A. (2021). Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery. *Current Drug Targets*, *22*(6), 631–655. https://doi.org/10.2174/1389450122999210104205732
Olisah, C. C. (2022). Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. *Computer Methods and Programs in Biomedicine*, *220*. https://doi.org/10.1016/j.cmpb.2022.106773
Perez-Diaz, J. A. (2020). A Flexible SDN-Based Architecture for Identifying and Mitigating Low-Rate DDoS Attacks Using Machine Learning. *IEEE Access*, *8*, 155859–155872. https://doi.org/10.1109/ACCESS.2020.3019330
Prasad, P. (2020). Application of machine learning techniques in groundwater potential mapping along the west coast of India. *Giscience and Remote Sensing*, 735–752. https://doi.org/10.1080/15481603.2020.1794104
Rani, P. (2023). Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. *Computers and Electrical Engineering*, *105*. https://doi.org/10.1016/j.compeleceng.2022.108543
Rasyidi, M. A. (2021). Classification of handwritten javanese script using random forest algorithm. *Bulletin of Electrical Engineering and Informatics*, *10*(3), 1308–1315. https://doi.org/10.11591/eei.v10i3.3036
Saleh, H. (2021). OPCNN-FAKE: Optimized Convolutional Neural Network for Fake News Detection. *IEEE Access*, *9*, 129471–129489. https://doi.org/10.1109/ACCESS.2021.3112806
Sarkar, C. (2023). Leaf disease detection using machine learning and deep learning: Review and challenges. In *Applied Soft Computing* (Vol. 145). https://doi.org/10.1016/j.asoc.2023.110534
Seraphina, Y. A., & Gunawan, P. H. (2025). Sentiment Analysis of Public Responses Regarding The Use of Electric Cars in Indonesia with Support Vector Machine and Random Forest Methods. *The Indonesian Journal of Computer Science*, *14*(1), 510–519. https://doi.org/10.33022/ijcs.v14i1.4649
Shahriar, S. (2021). Prediction of EV charging behavior using machine learning. *IEEE Access*, *9*, 111576–111586. https://doi.org/10.1109/ACCESS.2021.3103119
Singh, V. (2022). A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. *Diagnostics*, *12*(1). https://doi.org/10.3390/diagnostics12010116
Sivaranjani, S. (2021). Diabetes Prediction using Machine Learning Algorithms with Feature Selection and Dimensionality Reduction. In *2021 7th International Conference on Advanced Computing and Communication Systems Icaccs 2021* (pp. 141–146). https://doi.org/10.1109/ICACCS51430.2021.9441935
Sri, G. A., Prameswari, A., Rakib, M., & Aris, V. (2024). Sentiment Analysis Of Electric Vehicles On Twitter Using Machine Learning. *INNOVATIVE: Journal Of Social Science Research*, *4*(4), 837–845.
Tan, L. (2021). Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods. *Agriengineering*, *3*(3), 542–558. https://doi.org/10.3390/agriengineering3030035
Tang, J. (2020). Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review. *Analytic Methods in Accident Research*, *27*. https://doi.org/10.1016/j.amar.2020.100123
Tarigan, V. T., & Yusupa, A. (2024). Perbandingan Algoritma Maching Learning dalam Analisis Sentimen Mobil Listrik di Indonesia pada Media Sosial Twitter/X. *Jurnal Informatika Polinema*, *10*(4), 479–490. https://doi.org/10.33795/jip.v10i4.5130
Tetila, E. C. (2020). Detection and classification of soybean pests using deep learning with UAV images. *Computers and Electronics in Agriculture*, *179*. https://doi.org/10.1016/j.compag.2020.105836
Yazdinejad, A. (2023). An ensemble deep learning model for cyber threat hunting in industrial internet of things. *Digital Communications and Networks*, *9*(1), 101–110. https://doi.org/10.1016/j.dcan.2022.09.008