Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 1 Iss. 1 July 2023, pp: 1-9
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Sentiment Analysis of Towards Electric Cars using Naive Bayes Classifier and Support Vector Machine Algorithm

Suryani[1*], Muhammad Fauzi Fayyad[2], Daffa Takratama Savra[3], Viki Kurniawan[4], Baihaqi Hilmi Estanto[5]

[1,2,3,4]Department of Information Systems Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[5]Department of Biomedical Engineering, Faculty of Technology, Pamukkale University, Turkey

E-Mail: [1]12050320389@students.uin-suska.ac.id,
[2]12050314864@students.uin-suska.ac.id, [3]12050312110@students.uin-suska.ac.id,
[4]12050313603@students.uin-suska.ac.id, [5]bestanto19@posta.pau.edu.tr

Received March 11th 2023; Revised May 24th 2023; Accepted Jun 18th 2023
Corresponding Author: Suryani

Abstract

The use of non-renewable energy sources causes a reduction in fossil fuel resources, and greenhouse gas emissions. Based on the 2020 Climate Transparency Report, G20 member countries are trying to minimize gas emissions according to the target of the Nationally Determined Contribution (NDC), that the transportation sector contributes 27% of air pollution. The solution to reduce greenhouse gas emissions is to start using electric cars. The change from conventional transportation to electric transportation is expected to reduce carbon emissions and dependency on fossil fuels. However, the transition from conventional transportation to electric transportation raises pros and cons for the people of Indonesia. Social media Twitter is a forum for sharing opinions. Twitter users can express opinions on a matter. This study uses the sentiment analysis method to determine public opinion on the use of electric cars in Indonesia. Sentiment classification was performed using the NBC and SVM Algorithms. The results of this study indicate the use of two different algorithms, namely the Naive Bayes Classifier and SVM with the highest accuracy in Naive Bayes with k = 2 and k = 9 is 88%, while the highest accuracy in SVM with k = 9 and k = 10 is 90%. Thus, SVM has better capabilities than Naive Bayes in this study.

Keywords**:** Electric Cars, Naïve Bayes Classifier, Sentiment Analysis, Support Vector Machine, Twitter

1. INTRODUCTION

Non-renewable energy sources are shrinking very fast, this inspires researchers to search for various renewable energy sources with strategic planned control to replace non-renewable energy sources [1]. Frequently discussed topics relate to the global energy deficit, contaminated ecology, and increasing temperatures around the world. Another issue of concern today is the depletion of fossil fuel resources, and greenhouse gas emissions [2]. Greenhouse gases are the largest contributor to global warming caused by human activities [3]. One example is the use of fossil fuels (gasoline) in the field of transportation, both private and public vehicles [4]. According to Climate Transparency Report 2020 that the transportation sector contributes 27% of air pollution, G20 member countries seek to minimize gas emissions in accordance with the target of the Nationally Determined Contribution (NDC) [3]. A solution that can be done to reduce greenhouse gas emissions by starting the use of electric vehicles [2]. Indonesia's participation in efforts to reduce greenhouse gas emissions and world climate change is a commitment of the country's government. Presidential Regulation No. 55 of 2019 is a regulation on the use of electric vehicles as road transportation powered by batteries [3]. One of the most popular electric vehicles is the electric car. Electric cars are motorized vehicles that are environmentally friendly because they are powered by batteries and are able to reduce air pollution [5]. By changing conventional transportation to electric transportation, it is expected to reduce carbon emissions and dependence on fossil fuels [1].

However, the transition from conventional transportation to electric transportation has created pros and cons for the people of Indonesia. People can express their opinions in various ways, one of which is through social media. Social media has become a space to express oneself, give opinions and share ideas on various topics [6]–[9]. Among the various social media platforms, Twitter has experienced widespread deployment with large user adoption and rapid growth in communication volume [10]. This social media has a total of 7.2 million users spread across of Indonesia [11]. Twitter users can share and read short messages of no more than

140 characters known as 'Tweets', to provide perspective, express themselves and convey opinions on a matter [9], [12]. Opinions expressed on a discussion can be in favor of or against something. The opinions of twitter users are mainly used to determine whether the opinions are positive or negative using sentiment analysis. Sentiment analysis is the process of analyzing opinion data or views in the form of text to produce a conclusion based on these opinions [13]–[18]. Sentiment analysis will usually classify twitter messages into 3 classes, namely positive, negative and neutral [9], [19]. This analysis process requires a technology known as machine learning to classify the sentiment of twitter users. With the help of classification algorithms, sentiment analysis can be more easily completed.

Classification algorithms that are always used for sentiment analysis are Naïve Bayes Classifier and Support Vector Machine (SVM). Based on research by Riyadi et al (2020) on the Sentiment of Indonesian Residents towards electric vehicles, the Naïve Bayes algorithm is superior to Decision Tree and KKN with 94% accuracy [3]. The same research was also done by Kristiyanti et al (2018) who conducted research on the topic of comparing the Naïve Bayes and SVM algorithms for sentiment analysis of candidates for governor of West Java, the results of this study Naïve Bayes algorithm obtained 94% accuracy as the best accuracy [20]. Other research was also conducted by Aldisa and Maulana (2022) with the topic of comparing the Naïve Bayes, Decision Tree and SVM algorithms for analyzing the sentiment of the Indonesian people towards the Covid-19 booster vaccination, the results of the Naïve Bayes algorithm were superior to the Decision Tree and SVM algorithms with 83.81% accuracy [21]. The latest research was conducted by Undamayanti et al (2022) with the topic of applying the Naïve Bayes algorithm with Particle Swarm Optimization optimizer for analyzing student sentiment towards the MBKM program, the results of the Naïve Bayes algorithm modeling PSO optimization obtained an accuracy of 71. 96% [22]. And the latest research conducted by Santoso et al (2022) related to sentiment analysis of electric cars using the SVM algorithm with PSO optimization, obtained the accuracy of the PSO optimization SVM algorithm of 86.07% [5].

Referring to prior research, it is known that SVM and Naïve Bayes have good abilities in solving the problem of twitter user sentiment classification. Therefore, this research is an experiment to conduct a comparison between SVM and Naïve Bayes to find out the opinions of twitter users on the use of electric cars in Indonesia, whether they tend to be neutral, positive or negative.

2. MATERIALS AND METHODS

This research analyzes tweets with the keyword electric car. The stages carried out from collecting Twitter data, then performing several preprocessing stages to produce data that can be used in the next process. The research stages can be seen in Figure 1.

2.1 Sentiment Analysis

Sentiment analysis is one of the research branches of text mining that refers to the broad field of natural language processing and linguistic computing [23]. The focus of sentiment analysis is to analyze the opinions, comments and emotions of a person on a particular topic, product or service written in a reading document [18]. The basic task in sentiment analysis is to categorize readings according to polarity as positive, negative or neutral sentiment [24]. Sentiment analysis can also be used to analyze the emotional feelings expressed in reading documents whether disappointed, happy or angry [23]. In the business world, sentiment analysis is used to apply the analysis of consumer comments on a product or service owned by a company or organization, the results of the analysis then become one of the bases for decision making [18], [25].

2.2 Naive Bayes Classifiers

Bayes' theorem is the basis of the Naïve Bayes algorithm and has the function of finding the probability of each class to accurately predict the data [26]. This algorithm is widely used in solving high-dimensional data classification problems such as text because it has good accuracy and has been tested with a number of other classification algorithms [21]. The equation of Bayes' theorem is shown in equation 1.

$$P(C|X) = \frac{P(C|X).P(C)}{P(X)} \quad (1)$$

Description:
X : Data with unknown class
C : Hypothesized data X is a specific class
$P(C|X)$: Probability of hypothesis C based on condition X (posterior probability)
$P(C)$: Probability of hypothesis C (prior probability)
$P(X|C)$: Probability of X based on the condition in hypothesis C (likelihood)
$P(X)$: Probability of X (predictor prior probability)

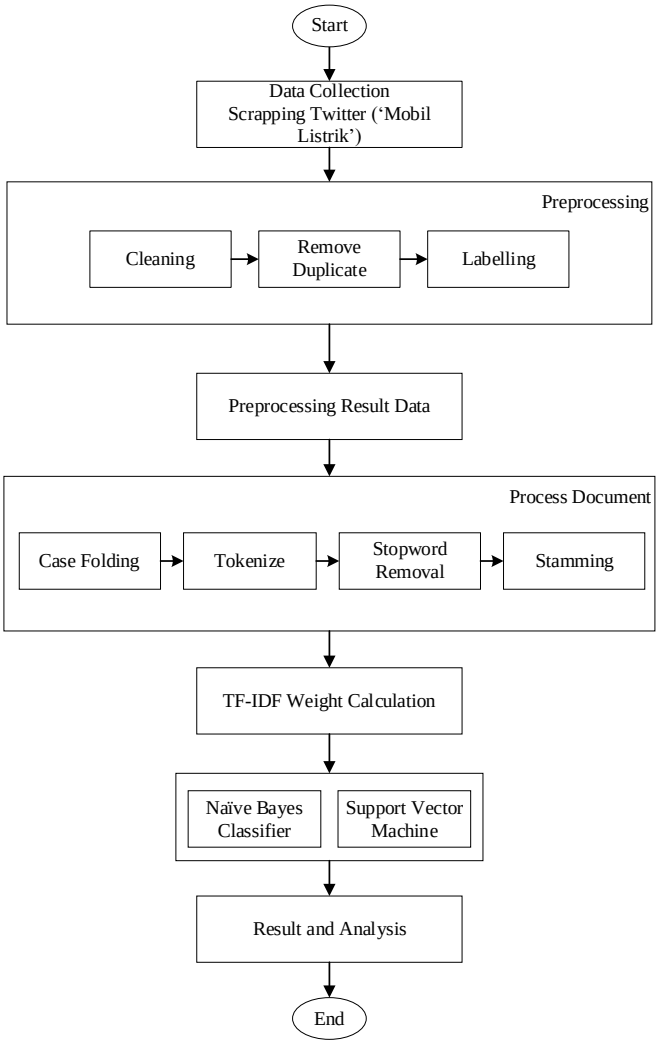

Image 1. Research Methodology

2.3 Support Vector Machine (SVM)

SVM is a supervised learning method which performs pattern recognition and data analysis. Classification and regression analysis problems can be solved using this algorithm [27]. In addition, SVM is known to have a good ability to solve sentiment analysis problems, because it has advantages in determining distance, namely the computational process becomes faster [5], [19]. The equation in SVM is shown in equation 2 or 3.

$$f(x) = w.x + b \tag{2}$$

or

$$f(x) = \sum_{i=1}^{m} a_i y_i K(x, x_i) + b \tag{3}$$

Description:

w : the hyperplane parameter sought (the perpendicular line between the hyperplane line and the support vector point)
x : Support Vector Machine input data points
ai : weight value of each data point
$K(x, xi)$: kernel function
b : hyperplane parameter to be searched (bias value)

In making decisions, SVM uses a kernel function K(xi, xd) [28][30]. In this research, sentiment classification uses the Radial Basis Function kernel equation shown in equation 4.

$$K(x_i, x) = \exp(-\gamma|x_i - x|^2), \gamma > 0 \quad (4)$$

2.4 Term Frequency-Inverse Document Frequency (TF-IDF) Weighting

TF-IDF is a stage of feature extraction process, which is giving weight to each word contained in a tweet that has previously been preprocessed [29]. Giving weight with TF-IDF aims to analyze how important a word represents a sentence [28]. This weighting is done based on the frequency of occurrence of words in a document [29], and can be found based on equation 5.

$$TF - IDF\,(d, t) = TD(d, t) * IDF\,(t) \quad (5)$$

where:

$$TF\,(d, t) = \frac{\text{number of t words in document d}}{\text{total words in document d}} \quad (6)$$

$$IDF(t) = \frac{\text{total documents}}{\text{number of documents containing word t}} \quad (7)$$

and:

t : word
d : document

3. RESULTS AND ANALYSIS

3.1 Data Collection

The initial stage in this research is collecting data. The data used as the object of research are Indonesian tweets with the keyword 'electric car', which have previously been collected by the crawling method using the Twitter API. The data collected amounted to 1000 data during the 3rd week of December 2022. The scrapped data has three attributes, namely datetime, username and content. In this study only needs one attribute, namely content, which contains Twitter user comments on the use of electric cars in Indonesia. The data collected is shown in Table 1.

Table 1. Crawling Result Data

No	Content
1	Oo jadi desain mobil listrik yg kubilang bentuknya mirip alien itu disebut sebagai desain futuristik
2	@Rofialimajid wkwkwkwk kalo mau serba listrik kaya mobil listrik atau kompor listrik mah minimal banget nih ya jangan ada mati lampu lagi wkwk nyusahin rakyat aja
3	Wii hebat nab Ayo menabung biar bisa beli mobil listrik wkwk https://t.co/9PqMV1yHbV
...	...
1000	Ya strategi transportasi ramah lingkungannya kan mobil listrik lol

The next stage is preprocessing which consists of cleaning and labeling to determine the sentiment of each data. In the cleaning process, meaningless parts of the data such as URL addresses, hashtags, usernames, emojis and special characters will be removed. Similarly, the duplication of data, so that the data that initially amounted to 1000 data only left 680 data. Labeling is done automatically using the TextBlob library, the results of the sentiment are shown in Table 2.

Table 2. Sentiment Identification Results

No	Content	Sentiment
1	Oo jadi desain mobil listrik yg kubilang bentuknya mirip alien itu disebut sebagai desain futuristik	Negative
2	wkwkwkwk kalo mau serba listrik kaya mobil listrik atau kompor listrik mah minimal banget nih ya jangan ada mati lampu lagi wkwk nyusahin rakyat aja	Negative
3	Wii hebat nab Ayo menabung biar bisa beli mobil listrik wkwk	Neutral
...	...	...
4	Ya strategi transportasi ramah lingkungannya kan mobil listrik lol	Positive

Based on Table 3, the TextBlob library successfully identifies and classifies data into three sentiment groups, namely positive, neutral and negative. The percentage of each sentiment is shown in Figure 2.

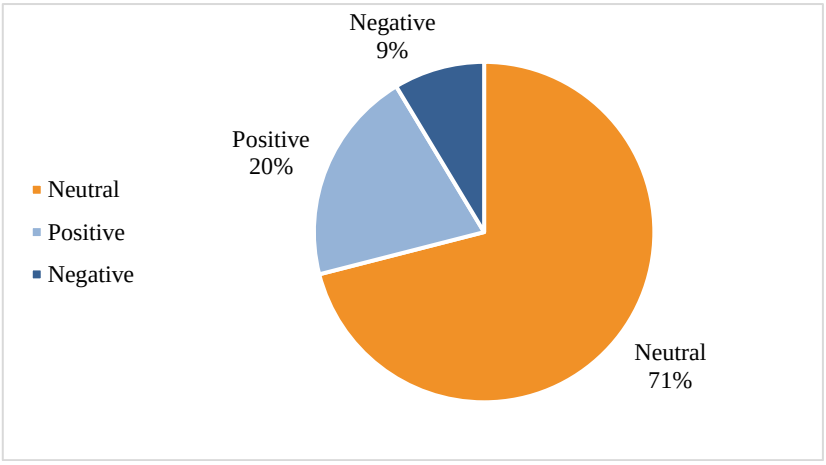

Image 2. Sentiment Percentage

3.2 Process Document

The first stage of preprocessing produces clean data that is already labeled. Furthermore, the data is used in the advanced preprocessing stage or proces document which consists of case folding, tokenization, stopword removal, stemming and weight calculation using TF-IDF. Case folding is the initial stage of the document processing which aims to change each word or text into a uniform form, namely lowercase. The next process is tokenization which aims to break a sentence into tokens based on its constituent syllables. Tokenization will produce a new data dictionary that is not related to the research topic. For that, it is necessary to do filtering utilizing stopword removal to clean up irrelevant words.

The next stage is stemming to return the word to its original form or base word by removing prefixes and affixes. Based on the stemming results, the next step is to calculate the word weight using TF-IDF. Weighting is done based on the frequency of occurrence of words in each document. This process aims to analyze the words that appear most often in conversations around the keyword 'electric car'. The results of the calculation of word weights using TF-IDF are shown in Table 3.

Table 3. TF-IDF Weight Calculation

No.	Battery	Energy	...	Vehicle	Electricity	Motor	Friendly	...	Subsidies
1	0	0	...	0	1	0	0	...	0
2	0	0	...	0	0.306	0	0	...	0
3	0	0	...	0	0.090	0	0	...	0
...	...	...	...	...	...	...	...	...	...
680	0	0	....	0	0	0	0	...	0

Based on the results of calculating the weights of each word using TF-IDF, a word cloud visualization is then performed to find the words that appear most often in conversations related to 'electric cars'. The wordcloud visualization is shown in Figure 3.

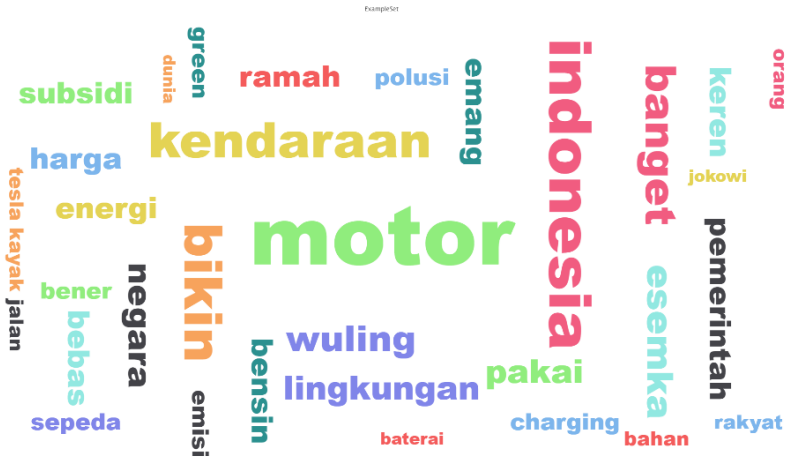

Image 3. Electric Car Wordcloud

Based on the word cloud visualization in Figure 3, it is known that words such as motorcycle, vehicle, wuling, environment, subsidy, esemka and others, have a high frequency of occurrence and contribute to the conversation around the keyword 'electric car'.

3.3 Naive Bayes Classifier and Support Vector Machine (SVM) Classification

Based on the sentiment analysis results shown in Figure 2, of the 680 Indonesian tweet data that has previously been preprocessed and documents processing, there are 139 data that express positive sentiments, 485 data express neutral sentiments, and 59 data express negative sentiments. The results of this analysis need to be re-evaluated to see the accuracy of sentiment classification. Therefore, it is necessary to calculate accuracy.

In this research, the evaluation of sentiment analysis results is carried out using two different algorithms, namely Naive Bayes Classifier and Support Vector Machine. These two algorithms were chosen based on certain considerations that refer to previous research on similar topics. The accuracy of an algorithm is influenced by several things, one of which is data sharing [23]. In this study, the data division technique used is K-Fold Cross Validation with a value of k = 10. The results of the accuracy test on the Naive Bayes and SVM algorithms are shown in Figure 4.

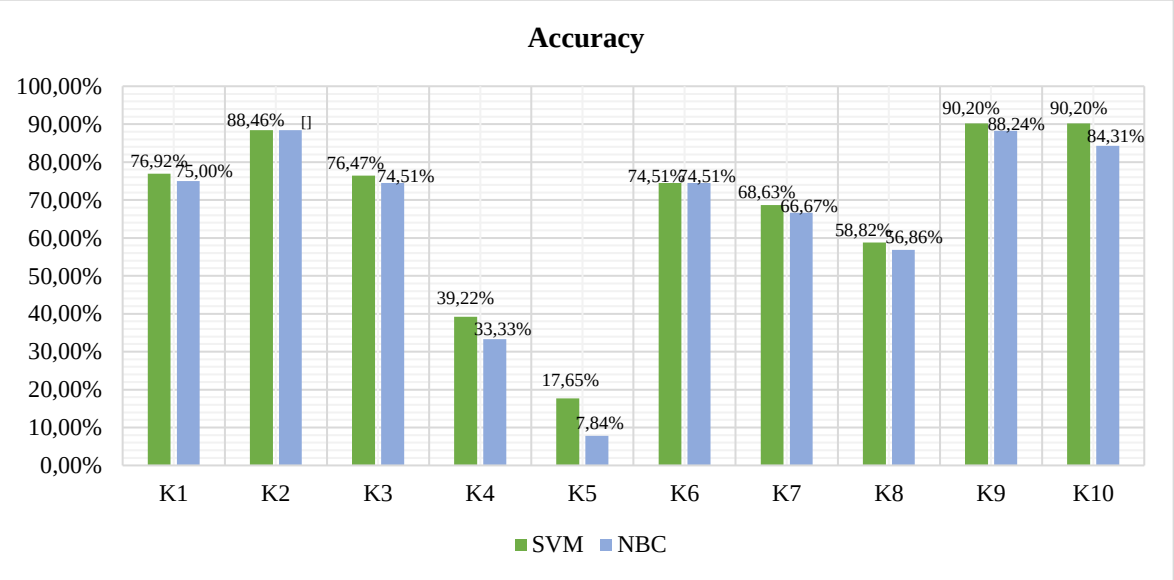

Figure 4. Accuracy value

Evaluation of sentiment analysis results using k = 10 data division in the Naive Bayes and SVM algorithms produces different accuracy values, as shown in Figure 4. In the Naive Bayes algorithm, the highest accuracy is at the value of k = 2, and k = 9, which is 88%. While in the SVM algorithm, the highest accuracy value is at the value of k = 9, and k = 10, which is 90%. At the value of k = 2, both algorithms have the same accuracy value, which is 88%. Based on these accuracy results, it can be concluded that both algorithms have a good ability to solve problems related to sentiment analysis of electric cars. This assumption is reinforced by the highest accuracy difference between Naive Bayes and SVM of only 2%.

In the next evaluation, evaluation is done based on the calculation of precision, recall and F1-score values in both algorithms. The calculation results on Naive Bayes are shown in Table 4.

Table 4. Precision, Recall, F1-score NBC results

Experiment	Precision	Recall	F1-score
K1	60.42%	31.71%	29.21%
K2	96.15%	33.33%	31.29%
K3	59.18%	32.48%	28.79%
K4	42.75%	39.29%	27.78%
K5	66.67%	41.18%	12.70%
K6	59.18%	31.67%	28.46%
K7	57.45%	29.82%	26.67%
K8	85.62%	33.33%	24.17%
K9	63.33%	32.61%	31.25%
K10	46.74%	44.79%	45.74%

Based on the results of the precision calculation, the best value in the Naive Bayes algorithm is obtained in the k=2 value experiment, which is 96.15%, while the best recall calculation is the k=10 value, which is 44.79%, and the best F1-score experiment at k=10 value, which is 45.74%. Conversely, in the SVM algorithm, the results of the calculation of precision, recall and F1-score values are shown in Table 5.

Table 5. SVM Precision, Recall, F1-score Results

Experiment	Precision	Recall	F1-score
K1	76.92%	76.92%	76.92%
K2	88.46%	88.46%	88.46%
K3	76.47%	76.47%	76.47%
K4	39.22%	39.22%	39.22%
K5	17.65%	17.65%	17.65%
K6	74.51%	74.51%	74.51%
K7	68.63%	68.63%	68.63%
K8	58.82%	58.82%	58.82%
K9	90.20%	90.20%	90.20%
K10	90.20%	90.20%	90.20%

In the SVM algorithm, the highest precision, recall and F1-score values are obtained in the k=9 value experiment, and the k=10 value is 90.20%. While the lowest value is obtained in the k = 5 value experiment, which is 17.65%.

4. CONCLUSIONS

Based on the results of the research that has been conducted, several findings are obtained. The first finding states that the sentiment of Twitter users towards the use of electric cars in Indonesia tends to be neutral during the 3rd week of December 2022. Of the 1000 data collected, the remaining 680 data can be used in the calculation. Of the 680 data, 20% or 139 data expressed positive sentiment, 71% or 485 data expressed neutral sentiment, and 9% or 59 data expressed negative sentiment. The results of this sentiment analysis are then re-evaluated using two different algorithms, namely Naive Bayes Classifier and SVM. The highest accuracy in Naive Bayes is 88%, while the highest accuracy in SVM is 90%. Thus, SVM has a better ability than Naive Bayes in this study. Furthermore, based on the calculation of word weights using TF-IDF, we managed to get the second finding which is the frequency of words that appear most often in the conversation around the keyword 'electric car'. Based on the weighting, it is known that words such as motorcycle, vehicle, wuling, environment, emission, subsidy, and esemka are words with a high frequency of occurrence.

REFERENCES

[1] S. A. A. Rizvi, A. Xin, A. Masood, S. Iqbal, M. U. Jan, and H. ur Rehman, "Electric Vehicles and their Impacts on Integration into Power Grid: A Review," in *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, IEEE, 2018, pp. 1–6.

[2] A. Mohammad, R. Zamora, and T. T. Lie, "Integration of electric vehicles in the distribution network: A review of PV based electric vehicle modelling," *Energies (Basel)*, vol. 13, no. 17, Sep. 2020, doi: 10.3390/en13174541.

[3] A. F. Riyadi, F. R. Rahman, M. A. Nofa Pratama, M. K. Khafidli, and H. Patria, "Assessment of Social Sentiment Toward Electric Vehicle Technology: Empirical Evidence in Indonesia," *EXPERT: Jurnal Manajemen Sistem Informasi dan Teknologi*, vol. 11, no. 2, p. 141, Dec. 2021, doi: 10.36448/expert.v11i2.2171.

[4] J. I. Mikayilov, S. Mukhtarov, H. Dinçer, S. Yüksel, and R. Aydın, "Elasticity analysis of fossil energy sources for sustainable economies: A case of gasoline consumption in Turkey," *Energies (Basel)*, vol. 13, no. 3, 2020, doi: 10.3390/en13030731.

[5] A. Santoso, A. Nugroho, and A. S. Sunge, "Sentiment Analysis About Electric Cars With Support Vector Machine Method and Feature Selection Particle Swarm Optimization," *Journal of Practical Computer Science*, vol. 2, no. 1, pp. 24–31, 2022.

[6] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 765–772. doi: 10.1016/j.procs.2019.11.181.

[7] M. K. Anam, B. N. Pikir, and M. B. Firdaus, "Applications of Na¨ive Bayes Classifier, K-Nearest Neighbor and Decision Tree to Analyze Sentiment on Netizen and Government Interaction," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, pp. 139–150, Nov. 2021, doi: 10.30812/matrik.v21i1.1092.

[8] A. M. Rahat, A. Kahir, and A. K. M. Masum, "Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset," in *Proceedings of the SMART–2019 8th International Conference on System Modeling & Advancement in Research Trends*, 2019, pp. 266–270.

[9] J. Song, K. T. Kim, B. Lee, S. Kim, and H. Y. Youn, "A novel classification approach based on Naïve Bayes for Twitter sentiment analysis," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 11, no. 6, pp. 2996–3011, 2017.

[10] D. Zimbra, A. Abbasi, D. Zeng, and H. Chen, "The state-of-the-art in twitter sentiment analysis: A review and benchmark evaluation," *ACM Transactions on Management Information Systems*, vol. 9, no. 2. Association for Computing Machinery, Apr. 01, 2018. doi: 10.1145/3185045.

[11] A. Deviyanto and M. D. R. Wahyudi, "Application of Sentiment Analysis on Twitter Users using the K-Nearest Neighbor Method," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 3, no. 1, pp. 1–13, 2018.

[12] M. Vadivukarassi, N. Puviarasan, and P. Aruna, "Sentimental Analysis of Tweets Using Naive Bayes Algorithm," *World Appl Sci J*, vol. 35, no. 1, pp. 54–59, 2017, doi: 10.5829/idosi.wasj.2017.54.59.

[13] Bahrawi, "SENTIMENT ANALYSIS USING RANDOM FOREST ALGORITHM-ONLINE SOCIAL MEDIA BASED," *JOURNAL OF INFORMATION TECHNOLOGY AND ITS UTILIZATION*, vol. 2, no. 2, 2019.

[14] M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler : Twitter," in *Fourth International Conference on Informatics and Computing (ICIC)*, 2019, pp. 1–5.

[15] J. A. Septian, T. M. Fahrudin, and A. Nugroho, "Sentiment Analysis of Twitter Users Towards Indonesian Football Polemics Using TF-IDF Weighting and K-Nearest Neighbor," *JOURNAL OF INTELLIGENT SYSTEMS AND COMPUTATION*, vol. 1, no. 1, pp. 443–49, 2019, [Online]. Available: https://t.co/9WloaWpfD5

[16] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electronics (Switzerland)*, vol. 9, no. 3, Mar. 2020, doi: 10.3390/electronics9030483.

[17] M. R. Huq, A. Ali, and A. Rahman, "Sentiment Analysis on Twitter Data using KNN and SVM," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, 2017, [Online]. Available: www.ijacsa.thesai.org

[18] Ankit and N. Saleena, "An Ensemble Classification System for Twitter Sentiment Analysis," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 937–946. doi: 10.1016/j.procs.2018.05.109.

[19] A. Alsaeedi and M. Z. Khan, "A Study on Sentiment Analysis Techniques of Twitter Data," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019, [Online]. Available: www.ijacsa.thesai.org

[20] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin, and L. Marlinda, "Comparison of SVM & Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter," in *The 6th International Conference on Cyber and IT Service Management (CITSM 2018)*, 2018. [Online]. Available: www.twitter.com

[21] R. T. Aldisa and P. Maulana, "Sentiment Analysis of Public Opinion on COVID-19 Booster Vaccination with Comparison of Naive Bayes, Decision Tree and SVM Methods," *Technology and Science (BITS)*, vol. 4, no. 1, pp. 106–109, 2022, doi: 10.47065/bits.v4i1.1581.

[22] E. Undamayanti *et al.*, "Sentiment Analysis Using the Naive Bayes Method Based on Particle Swarm Optimization Towards the Implementation of the Merdeka Learning Program at Merdeka Campus," 2022.

[23] F. F. Mailoa and L. Lazuardi, "Sentiment Analysis of Twitter Data Using Text Mining Method About Obesity Problem in Indonesia," *Journal of Information Systems for Public Health*, vol. 6, no. 1, pp. 44–51, 2019.

[24] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrainthos, "Sentiment Analysis of Online Learning on Twitter during the COVID-19 Pandemic Using the Naïve Bayes Method," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 1, pp. 157–163, Jan. 2021, doi: 10.30865/mib.v5i1.2604.

[25] A. Harun and D. P. Ananda, "Analysis of Public Opinion Sentiment About Covid-19 Vaccination in Indonesia Using Naïve Bayes and Decission Tree," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 58–63, 2021.

[26] E. Ditendra, Suryani, S. Romelah, M. H. A. Tanjung, and M. Sarah, "Comparison of Classification Algorithms for Sentiment Analysis of Islam Nusantara in Indonesia," vol. 2, pp. 71–77, 2022.

[27] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Sentiment Analysis of Ruang Guru Application on Twitter using Classification Algorithm," *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.

[28] A. M. Pravina, I. Cholissodin, and P. P. Adikara, "Sentiment Analysis of Airline Opinions on Twitter Documents Using Support Vector Machine (SVM) Algorithm," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 3, pp. 2789–2797, 2019, [Online]. Available: http://j-ptiik.ub.ac.id

[29] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, "Comparison of the Naïve Bayes Method and Support Vector Machine on Twitter Sentiment Analysis," *SMATIKA Jurnal: STIKI Informatika Jurnal*, vol. 10, no. 2, pp. 71–76, 2020.