Sentiment Analysis of YouTube Comments on the Internet Data Quota Rollover Policy Using Support Vector Machine and Naïve Bayes Algorithms

Khairunnisa, Muhammad Iqbal

Abstract

The growing use of the internet has made internet data quota an important need for supporting communication, education, work, business, and entertainment. The issue of unused quota expiring at the end of its validity period has generated various public responses, especially following the introduction of an internet data quota rollover policy as a form of protection for user rights. This study aims to analyze public sentiment toward the internet data quota rollover policy and compare the performance of Support Vector Machine (SVM) and Naïve Bayes algorithms in sentiment classification. The research data consisted of 3,281 public comments collected through scraping from a Liputan6 YouTube source discussing the internet data quota rollover policy. The data underwent preprocessing, including case folding, cleaning, tokenizing, normalization, stopword removal, and stemming. The text was then converted into numerical representations using Term Frequency–Inverse Document Frequency (TF-IDF) and classified into three categories: positive, negative, and neutral. Model evaluation used accuracy, precision, recall, F1-score, and a confusion matrix. The results show that neutral sentiment dominated at 64.54%, followed by negative sentiment at 25.51% and positive sentiment at 9.95%. SVM achieved an accuracy of 91%, while Naïve Bayes achieved 74%. SVM also performed better in most sentiment classes, with F1-scores of 0.94 for neutral, 0.91 for negative, and 0.74 for positive sentiment. Meanwhile, Naïve Bayes produced F1-scores of 0.83, 0.52, and 0.09, respectively. Thus, SVM performed better than Naïve Bayes in classifying public comments about the internet data quota rollover policy.

Keywords: *Sentiment Analysis, Quota Rollover, Support Vector Machine, Naïve Bayes, TF-IDF, YouTube.*

Khairunnisa[1]
Master of Information Thecnology Program, Pembangunan Panca Budi University, Medan, Indonesia
e-mail: nisak030720@gmail.com

Muhammad Iqbal[2]
e-mail: muhammadiqbal@dosen.pancabudi.ac.id
3rd International Conference Epicentrum of Economic Global Framework (ICEEGLOF)
Theme: Cross-Sector and Cross-Disciplinary Strategies for Achieving the SDGS
https://proceeding.pancabudi.ac.id/index.php/ICEEGLOF

Introduction

The development of information and communication technology has made the internet an essential part of people's lives. The internet is used for communication, education, work, business, public services, entertainment, and access to information. Therefore, the quality and availability of telecommunications services require government attention. The government plays a role in establishing policies, regulation, supervision, and control of telecommunications operations while considering the interests of the public as users. Law No. 36 of 1999 on Telecommunications stipulates that telecommunications are guided by the government through policy, regulation, supervision, and control [1]. User protection is also strengthened through Law No. 8 of 1999 on Consumer Protection, which regulates consumers' rights to accurate, clear, and honest information and fair treatment [2]. More recently, the government has also focused on protecting customers' remaining internet quota as part of consumer protection in telecommunications services [3].

One issue related to user protection is internet data quota rollover. Rollover is a mechanism that allows unused quota from a previous period to remain available or be carried over to the next period, subject to the telecommunications provider's terms. This mechanism is beneficial because quota purchased by customers does not immediately lose its value when the usage period ends. However, its implementation may vary according to the operator, package type, validity period, quota type, accumulation limits, and repurchase requirements. For example, Telkomsel provides the Quota Data Rollover feature, which allows remaining main quota to be used in the next period if certain requirements are met [4]. These differences in terms may lead to varied public perceptions. Policy developments in 2026 have further increased attention to this issue, as the government has emphasized the importance of protecting customers' remaining quota through mechanisms such as rollover, validity extensions, benefit transfers, compensation, or refunds [5].

The research problem arises because the internet data quota rollover policy has generated diverse public opinions, including support, criticism, complaints, and neutral responses. Digital media, particularly YouTube, is a place where people express their experiences and views on a policy. Comments on videos discussing internet data quota rollover may contain information about public acceptance or rejection. However, the large number of comments makes manual sentiment identification time-consuming and inefficient. In addition, YouTube comments commonly use informal language, abbreviations, nonstandard words, spelling errors, and various expressions that complicate classification. Therefore, computational methods are needed to identify public sentiment trends systematically and objectively.

The solution used in this study is sentiment analysis based on text mining and machine learning. The research data were obtained by scraping comments from a Liputan6 YouTube video discussing the internet data quota rollover policy. The data then underwent preprocessing, including case folding, cleansing, tokenizing, word normalization, stopword removal, and stemming. The comments were classified into three categories: positive, negative, and neutral. Term Frequency–Inverse Document Frequency (TF-IDF) was used to convert text into numerical representations based on the importance of words in documents [6]–[8]. The TF-IDF results were then used as input for classification using SVM and Naïve Bayes algorithms [9].

Support Vector Machine (SVM) and Naïve Bayes were selected because both are widely used in text classification and sentiment analysis. SVM identifies the best hyperplane for separating data from different classes and has good generalization capability [6], [10]. Naïve Bayes is a probabilistic classification method that determines a class based on the probability of data characteristics. Research by Pang, Lee, and Vaithyanathan has shown that SVM and Naïve Bayes can be used for sentiment classification [11]–[13]. Using both algorithms in this study enables performance comparison based on accuracy, precision, recall, and F1-score, making it possible to determine which algorithm is more suitable for the characteristics of public comments about the internet data quota rollover policy.

This study aims to analyze sentiment toward the internet data quota rollover policy using SVM and Naïve Bayes, classify comments as positive, negative, or neutral, and compare the performance of the two algorithms. Academically, this study is expected to contribute to the application of text mining, sentiment analysis, and machine learning to Indonesian-language public opinion data. Practically, the results may provide an overview of public perceptions and support the government and telecommunications providers in evaluating policies for protecting remaining internet quota.

Previous studies have shown that Support Vector Machine (SVM) and Naïve Bayes are widely used for sentiment analysis across various topics. Khaira, Aryani, and Hardian compared the two algorithms to analyze sentiment toward the Ministry of Education, Culture, Research, and Technology's internet quota policy during COVID-19. This study is closely related because both examine sentiment toward internet quota policies [14].

Setyawan, Azizah, and Pradani used SVM and Naïve Bayes to analyze public responses to the Indonesian government based on Twitter data [15]. Setiawan and Suryono applied both algorithms to sentiment analysis of Indonesia's new capital (IKN) using SMOTE, with SVM accuracy of 84% and Naïve Bayes accuracy of 77%. These results indicate that SVM performed better [16].

Asro, Azizah, and Sudaryono analyzed 1,800 YouTube comments about the 2024 election using Naïve Bayes and SVM. SVM achieved 76.33% accuracy, while Naïve Bayes achieved 72.19% [17]. Evanthe et al. compared SVM, Naïve Bayes, and Random Forest on 12,372 Telegram reviews, with accuracy of 88% for SVM and 85% for Naïve Bayes [18]. Setiadi applied both algorithms to sentiment about ShopeeFood services, while Fahrezy and Utomo analyzed sentiment about electric vehicles using SVM and Naïve Bayes, achieving accuracies of 82.95% and 73%, respectively [19], [20].

This study is expected to provide an objective, data-based overview of public sentiment trends regarding the internet data quota rollover policy and identify the best-performing algorithm. The findings may serve as an academic reference for future sentiment analysis research and provide the government and telecommunications providers with information about users' complaints, support, and perceptions. The recommendations suggest that telecommunications policy evaluations should consider not only technical and regulatory aspects but also public opinion expressed in digital spaces. Thus, sentiment analysis may support the development of more transparent, fair, and user-oriented policies on rollover, validity periods, and protection of remaining quota.

Research Methodology

This descriptive study uses a quantitative approach to apply Support Vector Machine (SVM) and Naïve Bayes algorithms to analyze sentiment toward the internet data quota rollover policy. The research data consist of public comments collected by scraping a Liputan6 YouTube video discussing the policy. The comments were classified into three sentiment categories: positive, negative, and neutral.

The research framework describes the systematic stages of the study, beginning with problem identification, data collection, data cleaning, sentiment labeling, text preprocessing, TF-IDF weighting, data splitting, application of SVM and Naïve Bayes, evaluation, and analysis of sentiment results. Data processing was carried out using Google Colab with Python and libraries such as Pandas, NumPy, NLTK, and Scikit-learn.

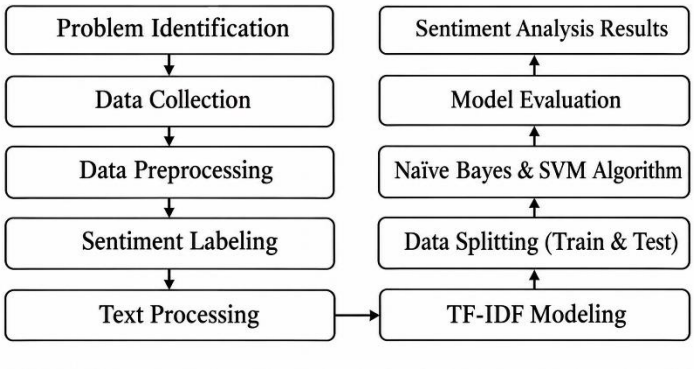


Figure 1. Research Stages

1. Data Collection
   Data were collected from comments on a Liputan6 YouTube video related to the internet data quota rollover policy. Scraping was used to obtain public comments. The data were then stored in a structured format, such as Excel or CSV. The main attribute was the comment text, while additional attributes could be used as supporting information.
2. Data Cleaning and Labeling
   The scraped data were examined to remove empty and duplicate comments, unnecessary characters, and comments unrelated to the research topic. The comments were then labeled by sentiment as positive, negative, or neutral. Positive sentiment indicates support or a favorable response, negative sentiment indicates criticism, complaints, or rejection, while neutral sentiment indicates neither a positive nor negative tendency.
3. Data Preprocessing
   The labeled data underwent preprocessing, including case folding, cleansing, tokenizing, word normalization, stopword removal, and stemming. These stages were intended to remove unnecessary elements from the text and standardize word forms so that the data could be processed more easily during classification.
4. Weighting and Data Splitting
   The processed text was converted into numerical form using Term Frequency–Inverse Document Frequency (TF-IDF). The resulting weights were used as features for classification. The dataset was then split into 80% training data and 20% testing data using train_test_split. The training data were used to build the model, while the testing data were used to assess the model's ability to classify data that had not been used during training.
5. Algorithm Application
   The SVM and Naïve Bayes algorithms were applied to the TF-IDF-weighted data. SVM determines the decision boundary between sentiment classes based on data characteristics, while Naïve Bayes determines the sentiment class based on the probabilities of word characteristics in comments. For text data, a linear-kernel SVM and Multinomial Naïve Bayes can be used because they are suitable for TF-IDF-based data.
6. Evaluation and Analysis
   The classification results were evaluated using accuracy, precision, recall, F1-score, and a confusion matrix. The evaluation measured each algorithm's ability to classify comments as positive, negative, or neutral. The sentiment distribution was then analyzed to identify public opinion trends regarding the internet data quota rollover policy. The evaluation values of both algorithms were presented as additional information to identify which model produced the best classification results for the research dataset. The final results were used to describe public perceptions and provide input to relevant parties when evaluating policies to protect remaining internet quota.

Results and Discussion

This study used public comments discussing the internet data quota rollover policy. The data were obtained from comments on the Liputan6 YouTube channel video titled “Era Kuota

Internet 'Hangus' Berakhir?" collected on August 12, 2026. The collection process yielded 3,281 comments as raw research data. All raw data were then prepared through preprocessing to clean and transform the text into a form suitable for sentiment analysis.

Table 1. Comment Data

No.	Username	Comment
1	@slametsutikno9239	Forget about it. Telkomsel is ignoring a Constitutional Court ruling. Remaining data quotas are still being forfeited (reduced to zero). Telkomsel is ignoring a Constitutional Court ruling.
2	@aguspras8888	TRADE MONOPOLY—it seems the Constitutional Court (MK) has to step in; it really is difficult to find an HONEST company in this COUNTRY.
3	@AwanKimika	No wonder I feel like I just bought it, yet suddenly it's already unusable, even though it hasn't even been a month yet.
4	@wayanutama5174	It's a lie; my Telkomsel data quota expired too.
5	@Netizen_nyinyir	A nightmare for Telkomsel that incredibly stingy provider.
6	@fizzx3081	Did no one notice, or what? That guest looks just like Abah Anies, lol! 😂
7	@susilosilo3640	I still had plenty of data left, BUT it's been forfeited... they said it's been officially confirmed.
8	@_BonjeR.	If you get 10k from each person, and multiply that by 3 million people, that's a really good profit.
9	@ramekeunskuyy1957	Telkomsel's network is absolutely terrible right now.
10	@Hery77syah12Syah	Even so, they still haven't heeded the Constitutional Court's ruling; our quota remains lost and forfeited.
11	@mrpsg3707	When did it start?
12	@masboy8552	Did my data quota expire because I bought it last month?
13	@clipper.magang_yt	The lawsuit might be granted, but the cost will surely be astronomical there are pros and cons to consider, folks.
14	@ahmadriyadi4204	Where's the proof? My data quota is still missing, even though I had more than 10GB left. I was happy when I heard the news, but in reality, it's just gone.
15	@kelincibulan4866	It's a hoax. It's still burnt.
16	@Bang_Edho-p7n	This still has an expiration date....
17	@SugandaSuganda-l8f	No way my remaining data quota is still getting swiped by the provider.
…	…	
…	…	
3281	@buzzz411	Consumer rights 😅

Source: https://www.youtube.com/watch?v=WP4qR2nCA8I&t=7

The collected data were prepared through several preprocessing stages: case folding, cleaning, tokenizing, normalization, stopword removal, and stemming. These stages transformed unstructured text comments into more structured data that could be used for sentiment analysis. After preprocessing, the data were used for sentiment labeling and classification with Support Vector Machine (SVM) and Naïve Bayes.

Table 2. Preprocessing Results

Stage	Result

Original Text	It should be like this. I usually top up 150 GB of data for a month. When the validity period ends, there's still 60 GB of data left. In the end, it just goes to waste. I really hope this news is true.
Case Folding	It should be like this. I often purchase a 150 GB data plan for the month. When the validity period ends, I still have 60 GB of data left. In the end, it just goes to waste. I really hope this news is true.
Cleaning	It should be like this; I often top up 150 GB of data for a month, but when the validity period ends, there's still 60 GB left, which just goes to waste. I really hope this news is true.
Tokenizing	It should be like this: I often top up 150 GB of data for a month, but when the validity period ends, there's still 60 GB left which just goes to waste. I really hope this news is true.
Normalization	It should be like this: I often top up 150 GB of data for a month, but when the validity period ends, there's still 60 GB left which just goes to waste. I really hope this news is true.
Stopword Removal	I usually top up with a 150 GB data package, but by the time the validity period ends, there's still 60 GB left which just goes to waste. I really hope this news is true.
Stemming	I have to top up 150 GB of data frequently; by the time the validity period ends, there's still 60 GB left, which just goes to wasteit's such a shame. I really hope this news is true.

After all preprocessing stages were completed, the data were analyzed by sentiment distribution, grouping comments into three categories: neutral, negative, and positive sentiment.

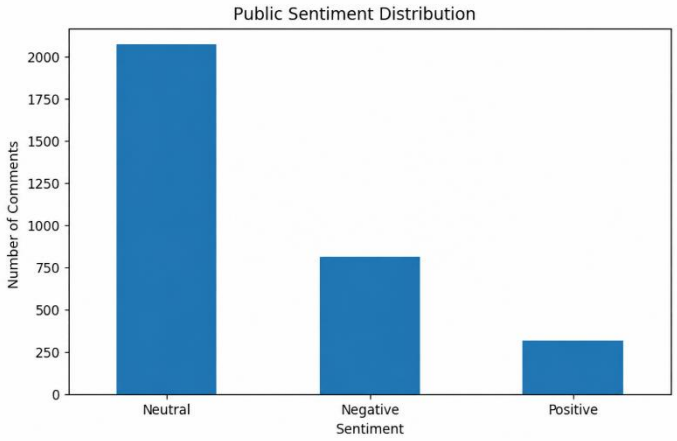


Figure 2. Sentiment Distribution

Figure 2 shows that neutral sentiment had the highest percentage at 64.54%, followed by negative sentiment at 25.51% and positive sentiment at 9.95%. The dominance of neutral sentiment indicates that most comments tended to provide responses, questions, or statements about the internet data quota rollover policy without strongly supporting or opposing it.

Negative sentiment, at 25.51%, indicates that many comments contained complaints, criticism, or public dissatisfaction with the policy. Meanwhile, positive sentiment, at 9.95%, indicates that some people supported the policy and considered internet data quota rollover beneficial to users. This distribution formed the basis for classification using SVM and Naïve Bayes.

Table 3. Classification Results

Metric	**Support Vector Machine**			**Naïve Bayes**		
	Neutral	Negative	Positive	Neutral	Negative	Positive
Accuracy		0.91			0.74	
Precision	0.90	0.95	0.93	0.71	0.95	1.00
Recall	0.98	0.88	0.61	0.99	0.36	0.05
F1-Score	0.94	0.91	0.74	0.83	0.52	0.09

Figure 3 presents the performance of Support Vector Machine (SVM) and Naïve Bayes in classifying sentiment toward the internet data quota rollover policy. The test results show that SVM achieved an accuracy of 0.91, while Naïve Bayes achieved 0.74. This indicates that SVM performed better than Naïve Bayes on this research data.

For SVM, precision for neutral, negative, and positive sentiment was 0.90, 0.95, and 0.93, respectively. Recall was 0.98, 0.88, and 0.61, with F1-scores of 0.94, 0.91, and 0.74. These results indicate that SVM recognized neutral and negative sentiment reasonably well, while recognition of positive sentiment was lower.

Naïve Bayes achieved precision values of 0.71, 0.95, and 1.00, recall values of 0.99, 0.36, and 0.05, and F1-scores of 0.83, 0.52, and 0.09 for neutral, negative, and positive sentiment, respectively. These values indicate that Naïve Bayes had difficulty recognizing positive sentiment. Overall, SVM outperformed Naïve Bayes in terms of accuracy and F1-score.

In addition to evaluation results for accuracy, precision, recall, and F1-score, this study presents confusion matrices for SVM and Naïve Bayes. A confusion matrix provides a more detailed view of each algorithm's ability to classify data into the three sentiment categories: neutral, negative, and positive. It shows the number of correctly classified data points and the misclassifications in each category. The confusion matrices were then used to compare the algorithms' ability to recognize each sentiment category.

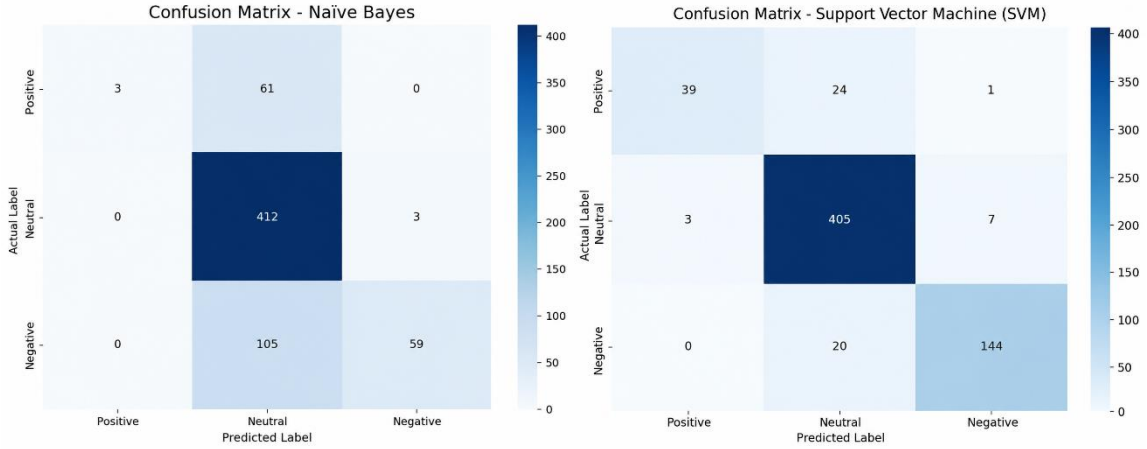


Figure 3. Confusion Matrix Results

Based on the Naïve Bayes confusion matrix, 474 of the 643 test data points were classified correctly: 3 positive, 412 neutral, and 59 negative. Naïve Bayes recognized neutral sentiment very well, as shown by the 412 neutral data points correctly predicted. However, the model had difficulty recognizing positive and negative sentiment. Of the 64 positive data points, only 3 were predicted as positive, while the remaining 61 were predicted as neutral. For negative sentiment, 59 data points were correctly predicted, while 105 were instead predicted as neutral.

The SVM confusion matrix showed better results. Of the 643 test data points, 588 were classified correctly, consisting of 39 positive, 405 neutral, and 144 negative data points. Misclassifications still occurred: 24 positive data points were predicted as neutral, 1 positive data point as negative, 3 neutral data points as positive, 7 neutral data points as negative, and 20 negative data points as neutral. Compared with Naïve Bayes, SVM was better able to recognize positive and negative sentiment and made fewer classification errors. These results show that SVM performed better than Naïve Bayes on this dataset.

This study used a word cloud visualization to provide an overview of the words most frequently appearing in public comments about the internet data quota rollover policy. A word cloud is a text visualization that displays words according to their frequency, with more frequent words shown in larger sizes. This visualization helps identify the words most commonly used by the public to express their opinions.



Figure 4. Word Cloud Visualization

Based on Figure X, some words, such as “seperti” and “bagus,” appear larger than others. Larger word sizes indicate that those words occur more frequently in public comments. The words “kuota” and “hangus” indicate that much of the discussion concerns remaining quota that cannot be used after an internet package expires. In addition, the words “masa” and “paket” indicate that validity periods and internet package terms are frequently discussed. Overall, the word cloud provides an overview of the words most frequently used by the public when expressing opinions about the internet data quota rollover policy.

Conclusion

Based on the sentiment analysis of public comments about the internet data quota rollover policy using Support Vector Machine (SVM) and Naïve Bayes, it can be concluded that sentiment analysis based on text mining and machine learning can identify public opinion trends concerning the protection of remaining internet quota. The research data were obtained by scraping comments from a Liputan6 YouTube source discussing the policy. The data underwent preprocessing, including case folding, cleaning, tokenizing, normalization, stopword removal, and stemming, before being weighted using Term Frequency–Inverse Document Frequency (TF-IDF).

The analysis shows that neutral sentiment dominated the data at 64.54%, followed by negative sentiment at 25.51% and positive sentiment at 9.95%. The dominance of neutral sentiment indicates that most comments consisted of responses, questions, or statements that did not clearly support or oppose the policy. Meanwhile, the higher percentage of negative sentiment compared with positive sentiment indicates considerable public criticism and complaints about remaining quota and internet package validity periods.

Based on the test results, SVM achieved 91% accuracy, while Naïve Bayes achieved 74%. SVM also performed better in terms of precision, recall, and F1-score for most classes. For SVM, the F1-scores for neutral, negative, and positive sentiment were 0.94, 0.91, and 0.74, respectively. Naïve Bayes produced F1-scores of 0.83, 0.52, and 0.09, respectively. These results indicate that SVM handled the characteristics of the comment data in this study better than Naïve Bayes.

Therefore, SVM is the more suitable algorithm for sentiment classification on this dataset. However, the small number of positive sentiment examples and the imbalance in class distribution should be considered in future research to improve the model's ability to recognize each sentiment category.

References

[1] P. R. Indonesia, “UU No. 36 Tahun 1999 tentang Telekomunikasi,” no. 1, 1989.
[2] Undang-Undang No. 8, “Peraturan UU No. 8 Tahun 1999 tentang Perlindungan Konsumen,” no. 8, 1999.
[3] S. Penghangusan and K. Internet, “Keabsahan Klausula Baku dalam Menentukan Wanprestasi pada Perjanjian Telekomunikasi,” pp. 2836–2843, 2026.

[4] J. M. Sumilat, W. H. F. Rorimpandey, and S. Siruru, "Analisis Dampak Pembelajaran Daring Terhadap Hasil Belajar Siswa," *J. Basicedu*, vol. 6, no. 5, pp. 8537–8544, 2022, doi: 10.31004/basicedu.v6i5.3739.

[5] I. Wasahua and T. Elawati, "Perlindungan Hukum Konsumen terhadap Praktik Penghangusan Kuota Internet oleh Penyelenggara Jasa Telekomunikasi di Indonesia," pp. 3185–3200, 2026.

[6] P. H. C. Samanmali and R. A. H. M. Rupasingha, "Sentiment analysis on google play store app users' reviews based on deep learning approach," *Multimed. Tools Appl.*, 2024, doi: 10.1007/s11042-024-19185-w.

[7] A. Gerliandeva, Y. H. Chrisnanto, and H. Ashaury, "Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naïve Bayes dan Ekstraksi Fitur TF-IDF serta N-grams Optimization of Sentiment Classification on Online Comments using Multinomial Naïve Bayes and TF-IDF Feature Extraction and N-grams," no. X, pp. 259–272, 2024, doi: 10.56873/jpkm.v9i2.5585.

[8] P. Sheridan *et al.*, "A Fisher ' s Exact Test Justification of the TF – IDF Term-Weighting Scheme," *Am. Stat.*, vol. 0, no. 0, pp. 1–24, 2026, doi: 10.1080/00031305.2025.2539241.

[9] D. Nurmadewi, Z. F. Jailani, and H. Rafi, "Comparison of Support Vector Machine and Naïve Bayes for E-Commerce Review Sentiment Classification Perbandingan Support Vector Machine dan Naïve Bayes untuk Klasifikasi Sentimen Ulasan E-Commerce," vol. 6, no. April, pp. 923–933, 2026.

[10] F. A. Indriyani, A. Fauzi, and S. Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine Tiktok application sentiment analysis using naïve bayes algorithm and support vector machine," vol. 10, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.

[11] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.

[12] A. Kausar *et al.*, "IMPLEMENTASI ALGORITMA NAÏVE BAYES CLASSIFIER UNTUK," vol. 10, no. 2, 2023.

[13] A. Novriandy, "Implementasi Algoritma Naive Bayes dan Algoritma C4 . 5 dalam Klasifikasi Kelayakan Bantuan UMKM," vol. 4, no. 1, pp. 208–217, 2023, doi: 10.30865/klik.v4i1.1099.

[14] U. Khaira, R. Aryani, and R. W. Hardian, "Perbandingan Algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam Analisis Sentimen Kebijakan Kemdikbudristek Tentang Kuota Internet Selama Pandemi Covid-19," vol. 18, no. 2, pp. 183–191, 2023.

[15] H. Setyawan, L. Ma, A. Y. Pradani, U. M. Yogyakarta, and J. Brawijaya, "Sentiment Analysis of Public Responses on Indonesia Government Using Naïve Bayes and Support Vector Machine," vol. 4, no. 1, pp. 1–7, 2023.

[16] A. Setiawan and R. R. Suryono, "Edumatic : Jurnal Pendidikan Informatika Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes," vol. 8, no. 1, pp. 183–192, 2024, doi: 10.29408/edumatic.v8i1.25667.

[17] A. N. S. Asro, N. Azizah, and S. Sudaryono, "Public Sentiment Analysis on the 2024 Presidential Election Using Naive Bayes Classifier (NBC) and Support Vector Machine (SVM) On Social Media Data," 2024.

[18] N. A. Evanthe, C. J. Angel, V. Sutanto, and E. D. Madyatmadja, "Text-to-Sentiment Analysis: Performance Evaluation of SVM, Naive Bayes, and Random Forest for Telegram User Review Classification in Indonesia," in *2025 2nd Beyond Technology Summit on Informatics International Conference (BTS-I2C)*, 2025, pp. 836–841.

[19] K. Setiadi, "KOMPUTA : Jurnal Ilmiah Komputer dan Informatika SHOPEEFOOD PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) KOMPUTA : Jurnal Ilmiah Komputer dan Informatika," vol. 12, no. 1, pp. 29–38, 2023.

[20] D. C. Fahrezy and F. S. Utomo, “Analisis Sentimen Opini Publik Tentang Kendaraan Listrik di Indonesia,” vol. 1, no. 2, pp. 127–139, 2024.