Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

Naive Bayes Algorithm to Enhance Sentiment Analysis of Coursera Application Reviews on Google Play Store

Masdarul Rizqi[1*], Martanto[2], Arif Rinaldi Dikanda[3], Dede Rohman[4]

[1,2,3,4] *STMIK IKMI Cirebon*

m.rizqi1221@gmail.com[1*], martantomusijo@gmail.com[2], rinaldi21crb@gmail.com[3], dede.rohman17@hotmail.com[4]

Abstract

Coursera is an online learning platform that provides various courses and certifications. This study aims to analyze user perceptions of the Coursera application after the reviews are translated into Indonesian, identify factors that influence positive and negative sentiment, and activate the effectiveness of the Naive Bayes algorithm in classifying review sentiment. The method used is Knowledge Discovery in Databases (KDD), with stages of data collection, preprocessing, and sentiment analysis using Naive Bayes. The results of the study show that the translation of reviews does not change the essence of user perception. Analysis of key words reveals positive experiences such as "kursus", "berguna", and "terima kasih", as well as criticism related to application performance. Factors such as price, content, and user experience play an important role in positive sentiment, while technical issues are the main cause of negative sentiment. The Naive Bayes model shows high accuracy with an accuracy value of 83.62%, precision of 83.34%, recall of 87%, and F1-score of 85.2%. These results indicate that the Naive Bayes algorithm is effective in analyzing sentiment of Coursera application user reviews. Further research is recommended to explore other algorithms or expand the analysis by considering additional factors that can influence user sentiment.

Keywords: *Sentiment Analysis, Naive Bayes, Coursera, User Reviews, Google Play Store*

1. Introduction

Coursera is an online learning platform that provides courses, certifications, and lectures in a variety of subjects. Examples of lessons offered include Advanced Data Analytics with GPT Chat and SQL for Data Science. Certifications available include Google Data Analytics Professional Certificate, Google Cybersecurity Professional Certificate, and Microsoft Business Analyst Professional Certificate. In addition, Coursera also offers courses such as the Master of Data Science from the Illinois Institute of Technology. This application is also used by programs such as the Digital Talent Scholarship themed Fresh Graduate Academy, Kampus Merdeka such as Bangkit Academy, and ICE Institute, which are increasingly expanding their influence in the field of higher education. Understanding user perceptions of these services through sentiment analysis of app reviews on the Google Play Store is important to provide insights into user experience and satisfaction, which can ultimately help in improving the quality and usability of the application.

Based on the results of the research conducted, the large number of Coursera app user reviews on the Google Play Store reflects a variety of user perceptions and experiences. This poses its own challenges in analyzing sentiment data effectively and efficiently. Some of the obstacles faced in managing this review data include the very large and increasing amount of data, the variety of languages and expressions used by users, limited technical skills in performing manual text analysis, and the lack of tools that can integrate sentiment analysis results with the application development process directly. This problem is important to note because without the right solution, application developers will have difficulty identifying key issues and user needs, which can ultimately hinder product improvement and innovation. Therefore, this study focuses on the application of the Naive Bayes algorithm for analyzing sentiment data for Coursera app reviews on the Google Play Store. By using the Naïve Bayes algorithm, which is a classification method that can predict the probability of a class so that it can produce decisions based on learning data [1], it is hoped that the process of managing and analyzing review data can be carried out more effectively and efficiently, and can reveal positive and negative sentiments in user reviews. The results of this analysis will provide valuable insights for developers in improving the quality of applications, adjusting features according to user needs, and making more informed strategic decisions. Thus, this study aims to develop an optimal solution in managing user review data using Natural Language Processing techniques and provide practical recommendations for Coursera application developers. In addition, this study is expected to contribute to the academic literature on Sentiment Analysis and Natural Language Processing, as well as become a basis for further research in this field. With a better understanding of user sentiment, the Coursera application can continue to grow and meet the educational needs of the world's population.

Previous studies have shown that the Naïve Bayes Classifier (NBC) method has been used effectively in various sentiment analysis contexts, ranging from the policy of enforcing restrictions on community activities (PPKM) to opinions on applications such as TikTok.

In a study related to the PPKM policy on Twitter, NBC succeeded in classifying public opinion with a dominance of positive sentiment [2]. In addition, the application of NBC to text-based sentiment analysis on Twitter showed quite good performance with an accuracy of 82% [3]. Similar results were also found in the analysis of YouTube comments, where NBC identified neutral sentiment as the majority with an accuracy of 90.36% [4]. Other studies related to the 2020 Pilkada noted that NBC was able to classify public opinion with an accuracy of 92.2% [5], while in the analysis of TikTok application reviews, NBC showed an accuracy of 79%, but still lower than the Support Vector Machine method [6]. From these studies, it can be concluded that although Naïve Bayes is quite effective, there is still room for further development, especially in the context of larger and more complex datasets.

This study aims to analyze the sentiment data of Coursera application reviews on the Google Play Store using the Naive Bayes algorithm. This study has important significance in providing in-depth insights into user experience and satisfaction with online learning applications such as Coursera. By applying sentiment analysis, this study contributes to filling the knowledge gap in the field of Informatics related to Natural Language Processing (NLP) and user data analysis. The results of this study also have potential practical benefits, such as helping users to provide more structured feedback that can be accessed by related parties, so as to improve user satisfaction and quality of experience.

2. Literature review

2.1 Related Research Results

The first study [7] discussed "Public Sentiment Analysis Towards the Pre-Employment Card Program on Twitter Using the Support Vector Machine Method". The study used the Support Vector Machine (SVM) method with the stages of data collection from Twitter using web scraping techniques. Text preprocessing such as tokenization, stemming, and stopword removal. Classification using Support Vector Machine. Model evaluation using metrics such as accuracy, precision, recall, and F1-score. The dataset used came from Twitter posts related to the Pre-Employment Card Program. The results showed that the SVM model with a linear kernel managed to achieve an accuracy of 99.67%, precision of 98%, recall of 99%, and F1-score of 98%. Meanwhile, the Radial Basis Function kernel produced an accuracy of 98.34%, precision of 97%, recall of 98%, and F1-score of 98%. This study also shows that public sentiment on Twitter towards the Pre-Employment Card program during the pandemic tends to be neutral at 98.34%. Based on the evaluation results, the linear kernel produces an accuracy of 99.67%, while the Radial Basis Function kernel produces an accuracy of 98.34%. The benefits of this study are mainly to understand the public's response to government policies, especially in the context of the Pre-Employment Card Program.

The second study [8] discussed the "Analysis of Pro and Contra Sentiment of Indonesian Society about the COVID-19 Vaccine on Twitter Social Media". The study used the Latent Dirichlet Allocation method with the stages of data collection from Twitter using web scraping techniques. Preprocessing. Latent Dirichlet Allocation Model. The dataset used in this study was obtained from Twitter, with a certain time span and focused on Indonesian-language tweets discussing the COVID-19 vaccine. The results of the study showed that pro-vaccine sentiment was more dominant than contra sentiment, with a positive response of 30% and a negative response of 26%. Positive sentiment is usually related to belief in the benefits of the vaccine, while negative sentiment is more due to concerns about side effects and distrust of the vaccination program. The benefits of this study are to provide insight for the government and related agencies to understand public opinion regarding vaccination, so that it can be used to develop more effective communication strategies and counteract nonsense related to the COVID-19 vaccine.

The third study [9] discusses the "Naive Bayes Classification Algorithm and Support Vector Machine in Student Complaint Services". The study uses the Naive Bayes and Support Vector Machine methods with the stages of collecting complaint data from students. Data processing and pre-processing, including normalization, removal of irrelevant data, and grouping by category. Models and evaluations using Naive Bayes and Support Vector Machine. The dataset used is student complaint data obtained from the complaint system at the university. The results show that the Support Vector Machine algorithm has higher accuracy than Naive Bayes. The AUC value for SVM reaches 0.922 with an accuracy of 84.45%, while Naive Bayes has an accuracy of 69.75% and an AUC of 0.679. With these results, universities can more effectively respond to student complaints quickly and accurately, which can ultimately improve student services and satisfaction.

The fourth study [10] discusses "Public Sentiment Analysis from Twitter Regarding Covid-19 Handling Policies in Indonesia with Naive Bayes Classification". The study uses the Naive Bayes Classification method with the stages of data collection from Twitter using web scraping techniques. Text preprocessing such as Tokenization, Cleansing, Case Folding, Filtering, and Stemming. Classification using Naive Bayes Classification. The dataset used was obtained from Twitter, with keywords related to Covid-19 handling policies in Indonesia. The results of this evaluation study obtained a classification accuracy level of 87.34%, sensitivity of 93.43%, and specificity of 71.76%. The benefits that can be obtained from this study are to provide insight for the government and policy makers regarding public responses to the policies implemented, so that they can be used as evaluation material for further policies.

The fifth study [11] discusses "Sentiment Analysis and Tourism Topic Modeling in Lombok Using Naive Bayes and Latent Dirichlet Allocation Algorithms". The study uses the Naive Bayes and Latent Dirichlet Allocation methods with the stages of collecting data from Twitter using web scraping techniques. Text preprocessing such as Tokenization, Cleansing, Case Folding, Filtering, and Stemming. Classification using Naive Bayes and Latent Dirichlet Allocation. Model evaluation using metrics such as accuracy, precision, recall, and specificity. The results of the study show that the Naive Bayes algorithm managed to achieve an accuracy value of 92%, precision 100%, recall 83.84%, and specificity 100%. As for LDA, positive topics were obtained in the 8th topic with a value of 0.613, and negative topics in the 12th topic with a value of 0.528. The benefits of helping the tourism industry, especially in Lombok, to understand tourist perceptions.

3. Research methods

3.1 Resarch methods

This study uses the Knowledge Discovery in Databases (KDD) method, which is a method for gaining knowledge from existing databases [12], and has advantages in the process of identifying patterns in an organized manner from complex data sets, so that the data becomes easier to understand. The KDD process in this study was applied to extract hidden patterns from Coursera application user data on the Google Play Store. KDD allows researchers to conduct structured data analysis to find significant relationships between existing variables. The following is the sequence of research methodologies that have been applied by researchers in Figure 1.

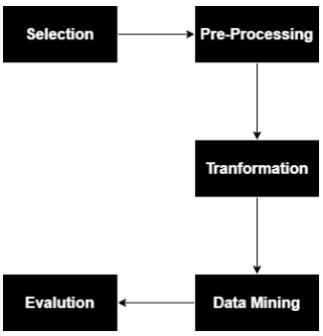

Figure 1: Knowledge Discovery in Databases

3.2 Selection

This stage begins with the collection of Coursera app review data on the Google Play Store using Python-based web scraping techniques. The data obtained includes reviews in various languages without geographical restrictions to ensure broad representation. Furthermore, reviews from various languages are translated into Indonesian using Google Sheets with the GOOGLETRANSLATE formula.

3.3 Pre-Processing

This stage, data is cleaned to ensure the quality of the analysis. The steps taken include cleansing to remove mentions, hashtags, RTs, links, numbers, punctuation, and non-alphanumeric characters. Case folding is applied to convert all text to lowercase, while slangword handling is done to normalize slang words or abbreviations. The tokenization process breaks the text into tokens, stopword removal removes irrelevant words, and sentences are rebuilt from the cleaned tokens.

3.4 Transformation

This stage converts the raw text data into a numeric format that can be used for further analysis. Numerical representation is done using the Term Frequency-Inverse Document Frequency (TF-IDF) method to give relevant weights to important words in the reviews. In addition, each review is labeled with a sentiment (positive or negative) based on a sentiment dictionary, so that the data is ready for the model analysis stage.

3.5 Data Mining

This stage involves applying machine learning algorithms to classify sentiment. The review data is divided into training data and test data. Next, the Naive Bayes algorithm is used to learn patterns from the training data and classify reviews as positive or negative sentiment on the test data.

3.6 Evaluation

This stage of evaluation is carried out to assess the performance of the model using metrics such as accuracy, precision, recall, and F1-score. The analysis uses the Confusion Matrix to calculate the number of correct predictions (true positive and true negative) and classification errors (false positive and false negative). The results of this evaluation show the level of success of the Naive Bayes algorithm in classifying review sentiment accurately.

4. Results and Discussion

4.1 Selection

Selection stage, user review data for the Coursera application on the Google Play Store is collected using web scraping techniques, where reviews are taken by sorting them based on relevance and includes all reviews with varying scores, from 1 to 5. The data is collected without geographic restrictions, to get a representation more comprehensive user sentiment. By collecting data directly from the source,

this study ensures that the dataset used is representative, relevant, and reflects the user's direct experience. The data collected consists of 27,457 reviews obtained between March 28, 2014 and September 19, 2024. Table 1. below is the result of data collection.

Table 1: Dataset Coursera

Content	Score
If I have to write a short review I would say ...	5
if Coursera have will add the functionality of...	5
I've learnt a lot from Coursera.	5
If I press the home button, receive a call, or..	3
If i change playback speed subtitles wont match.	1

After data collection, the next process is translating reviews from various languages, such as English, Japanese, Chinese, and others, into Indonesian. This translation process uses the Google Translate feature in Google Sheets to automatically transform review texts, which were originally in a foreign language, into a uniform language, namely Indonesian. This ensures consistency in sentiment analysis, even though the reviews come from various countries with different languages. Table 2 below is the result of the translation.

Table 2: Translation

Language	Before	After
China	很棒	Aplikasi terbaik untuk belajar
Korea	구맙 습 니 다	Terima Kasih
Arab	يوجد به الكثير من الأخطاء و التحديث الأخير مقيت جداااا	Ada banyak kesalahan dan pembaruan terbaru sangat menjijikan
Japanese	ありがとう！	Terima Kasih
English	You're my school	Kamu adalah sekolahku

4.2 Preprocessing

The preprocessing process is carried out to clean text data from irrelevant elements, such as numbers, blank lines, emojis, symbols, and redundant text such as RT in reviews.

4.2.1 Cleansing

Table 3. presents the results of the cleaning process, showing changes in data before and after cleaning was performed.

Table 3: Cleasing

Before	After
Jika saya harus menulis ulasan singkat, menurut saya ini adalah aplikasi yang sempurna	Jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna

The cleansing stage is carried out to clean the text by removing mentions, hashtags, retweets, links, numbers, characters other than letters and numbers, and punctuation.

4.2.2 Case Folding

Table 4. presents the results of the case folding process, showing changes in data before and after case folding was performed.

Table 4. Case Folding

Before	After
Jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna	jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna

The case folding stage is carried out to change all text to lowercase, thus maintaining consistency in the analysis.

4.2.3 Slangword Handling

Table 5. presents the results of the slangword handling process, showing changes in data before and after slang word handling was carried out.

Table 5. Slangword Handling

Before	After
jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna	jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna

The slangword handling stage is carried out to change abbreviations, slang, or words with different forms but the same meaning into one standard and uniform word form.

4.2.4 Tokenization

Table 6. presents the results of the tokenization process, showing changes in data before and after tokenization was performed.

Table 6: Tokenization

Before	After
jika saya harus menulis ulasan singkat menurut saya ini adalah aplikasi yang sempurna	['jika', 'saya', 'harus', 'menulis', 'ulasan', 'singkat', 'menurut', 'saya', 'ini', 'adalah', 'aplikasi', 'yang', 'sempurna']

The tokenization stage is carried out to break down text into smaller units, which are called tokens.

4.2.5 Stopword Removal

Table 7. presents the results of the stopword removal process, showing changes in data before and after stopword removal was performed.

Table 7: Stopword Removal

Before	After
['jika', 'saya', 'harus', 'menulis', 'ulasan', 'singkat', 'menurut', 'saya', 'ini', 'adalah', 'aplikasi', 'yang', 'sempurna']	['menulis', 'ulasan', 'singkat', 'aplikasi', 'sempurna']

The stopword removal stage is carried out to remove common words that often appear but do not provide significant meaning in text analysis, such as 'iya,' 'yaa,' 'gak,' 'nya,' and 'na'.

4.2.6 Postprocessing

Table 8. presents the results of the postprocessing process, showing changes in data before and after postprocessing was performed.

Table 8: Postprocessing

Before	After
['menulis', 'ulasan', 'singkat', 'aplikasi', 'sempurna']	menulis ulasan singkat aplikasi sempurna

The postprocessing stage in text processing aims to change the results of previous processing into a more structured format and according to analysis needs. This process involves steps such as recombining words into sentences, tidying up text by removing unnecessary characters or spaces, and rearranging data to make it easier to read or use for advanced modeling.

4.2.7 Labeling

Figure 2 presents the results of labeling with a positive word cloud, which shows frequently occurring positive words, while Figure 3 shows the labeling results for negative words.

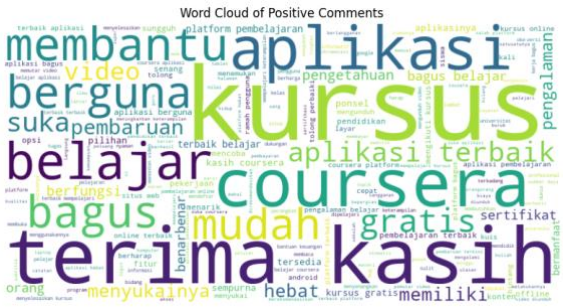

Figure 2: Positive Word Cloud

Figure 3: Negative Word Cloud

The labeling stage shows that the word cloud for positive sentiment reflects words frequently used by Coursera app users, such as 'Kursus,' 'berguna,' 'mudah,' and 'terima kasih.' In contrast, the word cloud for negative sentiment displays words that reflect complaints and dissatisfaction, such as 'salah,' 'bug,' 'berhenti,' and 'tolong perbaiki,' indicating the challenges users face in their experience.

4.3 Transformation

In the Transformation stage, the Term Frequency-Inverse Document Frequency (TF-IDF) method is used to convert review text into numeric representation.

	akses	akun	android	antarmuka	aplikasi	aplikasinya	bagus	bahasa	bantuan	belajar	...
0	0.0	0.0	0.0	0.0	1.000000	0.0	0.0	0.0	0.0	0.0	...
1	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	...
2	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	...
3	0.0	0.0	0.0	0.0	0.000000	0.0	1.0	0.0	0.0	0.0	...
4	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	...
...	...	...	...	...	...	...	...	...	...	...	...

Figure 4: Term Frequency-Inverse Document Frequency

4.4 Data Mining

At the data mining stage, the data is divided into two sets, namely the training set and the testing set with a ratio of 80:20.

4.5 Evaluation

The evaluation stage is carried out using metrics such as accuracy, precision, recall, F1-score, and Confusion Matrix to measure how well the model classifies sentiment.

```
Confusion Matrix:
[[403 103]
 [ 77 516]]

Classification Report:
              precision    recall  f1-score   support

    negative       0.84      0.80      0.82       506
    positive       0.83      0.87      0.85       593

    accuracy                           0.84      1099
   macro avg       0.84      0.83      0.83      1099
weighted avg       0.84      0.84      0.84      1099
```

Figure 5: Results of Confusion Matrix and Classification

4.5.1 Naive Bayes Evaluation Metrics

The Naive Bayes model is used to classify sentiment in Coursera app reviews. The classification results are evaluated not only based on accuracy but also on other important metrics, such as precision, recall, and F1-score, which are used to measure the performance of the model.

A. Accuracy

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{516+403}{516+403+103+77} = \frac{919}{1099} x\ 100\% = 83{,}62\%\ or\ 84\% \quad (1)$$

B. Precision

$$Precision = \frac{TP}{TP+FN} = \frac{516}{516+103} = \frac{516}{616} x\ 100\% = 83{,}34\%\ or\ 83\% \quad (2)$$

C. Recall

$$Recall = \frac{TP}{TP+FN} = \frac{516}{516+77} = \frac{516}{593} x\ 100\% = 87{,}01\%\ or\ 87\% \quad (3)$$

D. F1-Score

$$F1 = 2x\ \frac{Precision\ x\ Recall}{Precision+Recall} = \frac{0{,}83\ x\ 0{,}87}{0{,}83+\ 0{,}87} x\ 100\% = 85\% \quad (4)$$

Calculation explanation:
TP (True Positives): 516, Number of positive data correctly classified as positive.
TN (True Negatives): 403, Number of negative data correctly classified as negative.
FP (False Positives): 103, Number of negative data incorrectly classified as positive.
FN (False Negatives): 77, Number of positive data incorrectly classified as negative.

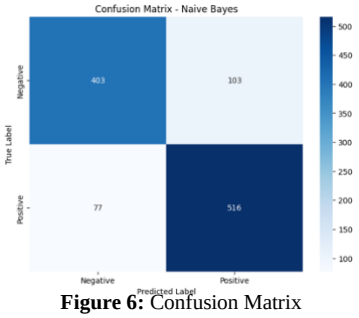

Figure 6: Confusion Matrix

5. Conclusion and recommendations

The results of the study on the analysis of sentiment of Coursera application user reviews using the Naive Bayes algorithm show that the process of translating user reviews into Indonesian does not change the essence of user perception. The analysis shows that the main words consistently reflect positive experiences and criticism of the application. Factors such as price, content, and user experience play an important role in positive sentiment, while technical issues are the main cause of negative sentiment. The Naive Bayes model has proven to be effective with high accuracy in classifying sentiment, as well as evaluation of precision, recall, and F1-score metrics that show good performance. For further research, it is recommended to explore the use of other algorithms or expand the analysis by considering additional factors that can affect user sentiment.

From the results of this study, it is recommended for further research to continue developing skills in sentiment analysis and NLP, especially related to the Naive Bayes algorithm, as well as exploring more complex machine learning algorithms such as Random Forest, XGBoost, or deep learning models such as LSTM to improve the accuracy of sentiment classification. In addition, the use of alternative translation APIs such as Microsoft Translator or Amazon Translate also needs to be considered for more consistent results. For Coursera, improving application performance and adding new features based on user reviews is essential to improving user experience. Developing broader language support will help reach a global audience and increase their convenience. Finally, campuses are advised to provide more research infrastructure support and develop curricula that include sentiment analysis, machine learning, and practical research projects to support students in facing big data-based research challenges.

References

[1] S. Styawati, A. R. Isnain, N. Hendrastuty, and L. Andraini, "Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis," *J. Inform. J. Pengemb. IT*, vol. 6, no. 1, pp. 56–60, 2021, doi: 10.30591/jpit.v6i1.3245.

[2] T. Krisdiyanto and E. M. O. Nurharyanto, "Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, pp. 32–37, 2021, doi: 10.24014/coreit.v7i1.12945.

[3] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 697–711, 2021, doi: 10.30645/j-sakti.v5i2.369.

[4] M. Hudha, E. Supriyati, and T. Listyorini, "Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamenantiterawan Dengan Metode Naïve Bayes Classifier," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.v5i1.3376.

[5] A. Muzaki and A. Witanti, "Sentimen Analisis Masyarakat Di Twitter Terhadap Pilkada 2020 Ditengah Pandemic Covid-19 Dengan Metode NaïVe Bayes Classifier," *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, 2021, doi: 10.20884/1.jutif.2021.2.2.51.

[6] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine," *TEKNOSAINS J. Sains, Teknol. dan Inform.*, vol. 10, no. 2, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.

[7] Styawati, N. Hendrastuty, and A. R. Isnain, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021, doi: 10.30591/jpit.v6i3.2870.

[8] F. F. Rachman and S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter," *Heal. Inf. Manag. J.*, vol. 8, no. 2, pp. 100–109, 2020, doi: 10.47007/inohim.v8i2.223.

[9] H. Hermanto, A. Mustopa, and A. Y. Kuntoro, "Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 5, no. 2, pp. 211–220, 2020, doi: 10.33480/jitk.v5i2.1181.

[10] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, pp. 228–238, 2021, doi: 10.32520/stmsi.v10i1.1179.

[11] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.

[12] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 2, pp. 247–253, 2020, doi: 10.32736/sisfokom.v9i2.903.