Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

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

Naïve Bayes Optimization by Implementing Genetic Algorithm in Sentiment Analysis of BCA Mobile Reviews

Muhammad Enricco Rizqy[1]*, Ahmad Faqih[2], Gifthera Dwilestari[3]

[1,2]*Teknik Informatika, STMIK IKMI Cirebon*
[3]*Sistem Informasi, STMIK IKMI Cirebon*
*menricooo@gmail.com**

Abstract

The development of the digital era has encouraged the adoption of mobile banking applications that facilitate banking transactions, including the BCA Mobile application which is simple but still adheres to a slightly outdated, user-friendly appearance but to provide the best service, it is necessary to evaluate the various problems that arise through review analysis. This study aims to conduct sentiment analysis of BCA Mobile application reviews taken from the Google Play Store, with data totaling 1,200 reviews scraping results using Google Collaboratory python programming language, to categorize negative and positive reviews used manual labeling for more accurate results, the Naïve Bayes approach is used in classifying positive and negative category reviews due to the ability of this algorithm to handle text data. However, the weakness of Naïve Bayes which is sensitive to irrelevant features can cause a decrease in accuracy. This research implements Genetic algorithm to improve the performance of Naïve Bayes. The results showed that the application of Genetic algorithm successfully increased the accuracy, precision of Naïve Bayes classification 95%, precision 92% to accuracy 98%, precision 99%, which proved the effectiveness of Genetic algorithm in optimizing the model and improving the quality of sentiment analysis.

Keywords: *Genetic Algorithm; Naive Bayes; Sentiment Analysis; Bank Application; Optimization*

1. Introduction

In today's digital era, mobile-based banking applications have become a major necessity for customers who want convenience and efficiency in transactions. One of the leading mobile banking apps in Indonesia is BCA Mobile, which has been named the number one favorite m-banking app by 2024 according to GoodStats.id. The app was developed by PT Bank Central Asia Tbk (BCA) and offers a variety of features, such as fund transfers, bill payments, and credit purchases, that make it easier for customers to do their banking activities[1]. To stay competitive and meet user expectations, it is important for BCA to understand user perceptions of its applications. Sentiment analysis of user reviews is an effective solution to identify the strengths and weaknesses of the app, allowing developers to improve service quality in the future[2]. Sentiment analysis of user reviews of applications such as BCA Mobile is becoming increasingly important amid the increasing adoption of digital technology[3]. While technology continues to evolve, the challenge of understanding user perception and experience remains significant. Many previous studies have focused more on technical aspects, while user experience is often overlooked. Issues such as data security, privacy, and app reliability are major concerns that affect user satisfaction and loyalty. Negative reviews that are not handled properly can damage an app's reputation. Therefore, sentiment analysis offers a solution to identify problems and improve app quality, while providing valuable insights to increase competitiveness in a competitive market. According to Hendra & Fitriyani in a study entitled “Sentiment Analysis of Halodoc Reviews Using Naïve Bayes Classifier”.Halodoc is a digital application service that allows users to access various health services anytime and anywhere. Reviews from users on the Google Play Store are a source of evaluation to optimize application performance. The research categorizes user reviews into positive sentiment, with the results of the sentiment analysis accuracy rate of 81.68% using the naïve bayes method[4]. Meanwhile, according to Rahman & Utami reported the results of their research entitled “Sentiment analysis of applications on Google Playstore using Naïve Bayes Algorithm and Genetic Algorithm.” Previous research evaluated the accuracy of sentiment analysis on four applications, namely Shopee, Gojek, Ruangguru, and Tokopedia, using two naïve bayes methods optimized with genetics. The results showed that the average accuracy using only naïve bayes reached 90%, but after applying the genetic algorithm wrapper feature, the average accuracy increased to 95%. This proves that genetic algorithm can significantly improve the performance of naïve bayes in sentiment analysis[5]. Furthermore, there is another study entitled “Naïve Bayes Optimization Using Genetic Algorithms on the Classification of Cyberbullying Comments on Social Media X”.This study uses two models to test the optimization of the naïve bayes algorithm with the genetic algorithm wrapper feature. The purpose of the research is to evaluate the effectiveness of optimization in addressing the issue of cyberbullying which often occurs in society and is a controversial topic on the X social media platform. The data used consists of 1,176 scrapped comments on the social media. After applying the genetic algorithm feature, the classification accuracy of comments related to cyberbullying increased to 77%[6]. The first objective of the research discussion was to evaluate user perceptions of the BCA Mobile application. This was done using the naive bayes method and google-colaboratory genetic features, as well as the python programming language[7]. The purpose of this research is to find the general perception of users, categorize frequently commented elements, and understand the elements that influence negative sentiment. Naive Bayes will be used to classify BCA Mobile user reviews into two labels: positive and negative. By conducting a comprehensive sentiment analysis, the research is expected to fill the existing knowledge gap in the literature related to user experiences with mobile banking applications. Significantly, this research will contribute to the field of informatics by providing an effective method for automatically evaluating user reviews. Additionally, the results of this research will have practical distribution benefits for software developers and BCA management in improving application quality, identifying areas for improvement, and designing better development strategies to meet user expectations. The genetic algorithm selection feature is very helpful in optimizing the hyperparameters of Naïve Bayes, reducing classification errors, increasing accuracy, and helping Naïve Bayes work more effectively. To achieve the research objectives, the method to be applied is sentiment analysis of BCA Mobile application user reviews using a genetic algorithm, an algorithm that can be used for optimization based on the principles of natural selection and natural genetics, which have been successfully applied in machine learning and optimization. The Naïve Bayes algorithm will use Google Colaboratory and the Python programming language. The latest user review data will be collected from BCA Mobile through platforms such as the Google Play Store. After data collection, the next step is data pre-processing, which includes data cleaning, tokenization, case folding, stopwords, and stemming to prepare the data before analysis. To categorize the sentiment of reviews into positive or negative categories, the naive bayes algorithm will be used. Patterns and trends in user perception and experience can be identified through this data analysis technique, which allows for the identification of sentiment patterns and trends. This computational method is not only efficient but also provides accurate results in understanding user sentiment, which can be used as a basis for decision-making in the development strategy of the BCA mobile application.

2. Research Methodology

This research evaluates the influence of independent variables on dependent variables. To study the events discussed, a quantitative approach was used. Quantitative data were obtained from reviews of the BCA Mobile application available on the Google Play Store[8]. The collection of review data was carried out using Google Play Scraper to extract data from the Google Play Store. The data obtained is processed using the Text Mining method, with word weighting calculated using Term Frequency-Inverse Document Frequency. (TF-IDF)[9]. After the manual labeling process, the reviews are categorized into positive or negative sentiment. To improve accuracy, genetic algorithms are used to optimize the hyperparameters of the Naive Bayes model, which serves as the main text classification method.

2.1. Data Analysis Techniques

This research uses the Knowledge Discovery in Databases (KDD) approach, a systematic data analysis method to extract valuable information. The process begins with understanding the research objectives and its domain, followed by data processing to handle missing values or anomalies. The next stage is Exploratory Data Analysis (EDA), which involves descriptive analysis and visualization to deeply understand the characteristics of the data[10].The feature engineering process can be carried out to create new features that enhance the model's performance. Model selection becomes a crucial step, followed by training using the training data. Model evaluation is conducted to measure performance using metrics such as accuracy and precision. The results of the analysis are interpreted to gain a deeper understanding of the data and findings. The knowledge obtained from this analysis can be used for business decision-making or scientific development. This KDD approach provides a strong foundation for exploring and understanding the meaning within the dataset. Figure 1 shows the stages in KDD.

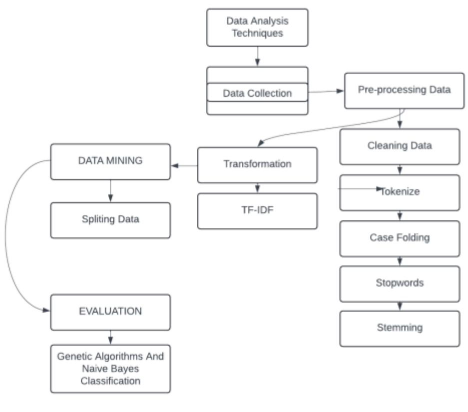

Fig. 1: Research methodology.

From the example, it can be seen in Fig 1 showing the stages of knowledge discovery in databases, data selection, preprocessing, transformation, data mining, evaluation.

2.2. Data Collection

The initial process in this research is data collection. Data is taken from reviews of the BCA Mobile application found on the Google Play Store using scraping techniques. This process is carried out using the Python programming language through the Google Colaboratory platform, utilizing the google-play-library package. The data obtained is then stored in .csv file format for further analysis.

2.3. Preprocessing

The preprocessing stage aims to prepare the data for analysis. This process includes several steps, namely:
1. Cleansing data: Removing irrelevant elements to ensure data quality.
2. Tokenization: Breaking down text into small units (tokens) such as words or phrases.
3. Case folding: Converting all letters to lowercase for uniformity.
4. Stopwords removal: Removing common words that do not have analytical value.
5. Stemming: Returning a word to its base form or root word.

2.4. Transformation

At the data transformation stage, the data is changed from its original form to a more suitable format so that it can be processed efficiently for analysis or modeling. One of the methods used is TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF works by assigning weights to each word based on two aspects. Term Frequency (TF) Measures how often a word appears in a specific document, Inverse Document Frequency (IDF) Assesses the uniqueness of a word by comparing its occurrence in a specific document with the entire collection of documents. A word that rarely appears across all documents will have a higher IDF value. This TF-IDF technology generates a numerical representation of the text, making it easier for further analysis or application in data modeling.

2.5. Data Mining

At the data mining stage, the collected data is divided into two parts, namely training data and test data. Training data is used to train the model to learn patterns in the data, while test data is used to measure how well the model can predict new data that it has not encountered before. This division is important to ensure that the model not only performs well on the training data but also provides reliable results when applied to real data.

2.6. Evaluation

The evaluation stage in this research uses a Confusion Matrix to assess the performance of the Naïve Bayes algorithm optimized with a Genetic Algorithm. The evaluation was conducted by measuring the accuracy and precision values, which are the main indicators in assessing the model's performance. Confusion Matrix is used to analyze the effectiveness of the model in predicting classes or labels. This matrix provides a comparative overview between true predictions and false predictions, both for positive and negative classes. By involving four main elements, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), this evaluation ensures that the model resulting from Genetic Algorithm optimization can provide more accurate and precise predictions compared to the Naïve Bayes model without optimization.

3. Results And Discussion

3.1. Data Collection

This research utilizes secondary data, which is data obtained from media or other intermediaries. The data was collected using Python through Google Colaboratory with the help of libraries, utilizing the google-play-scraper library to search for and collect reviews and user opinions on the BCA Mobile app in the Google Play Store[11]. Latest reviews NEWEST. The data obtained from scraping this research consists of 1,200 reviews in Indonesian. Fig 2 is the result of scrapping using the google-play-scrapper library

```
df = pd.DataFrame(np.array(result),columns=['review'])

df = df.join(pd.DataFrame(df.pop('review').tolist()))
df.to_csv("bca_dataset.csv", index = False)
df.head()
```

reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	appVersion
8d0a389-c5-4b45-9543-b951354	Muhamad Dinik	https://play-lh.googleusercontent.com/a/ACg8oc...	Buruk	1	0	4.5.3	2024-12-10 16:20:19	Mohon maaf atas ketidaknyamanan Bapak/Ibu, moh...	2024-12-10 16:59:33	4.5.3
c608ae9-e1-453d-b434-c8f05ab8	heri sandy	https://play-lh.googleusercontent.com/a-/ALV-U...	Kenapa FC trus yaaa ini BCA skrg	2	0	4.5.3	2024-12-10 16:13:19	Mohon maaf apabila Bapak/Ibu mengalami untuk k...	2024-12-10 16:59:30	4.5.3
101c0a4-12-45d1-af20-741d8ae	Aryan Pribadi	https://play-lh.googleusercontent.com/a-/ALV-U...	Tolong di bantu cara aktifasi Notifikasi uang ...	3	0	4.5.3	2024-12-10 15:31:51	Mohon maaf atas ketidaknyamanan Bapak/Ibu, Nam...	2024-12-10 15:43:22	4.5.3
4db64e3-c0f-475f-af73-83b99c8	Imanuel Yulio	https://play-lh.googleusercontent.com/a-/ALV-U...	Tolong kembalikan font tulisan yang lama	3	0	4.5.3	2024-12-10 15:28:40	Mohon maaf atas ketidaknyamanannya, ulasan Bap...	2024-12-10 15:43:23	4.5.3

Fig. 2: Scraping Results

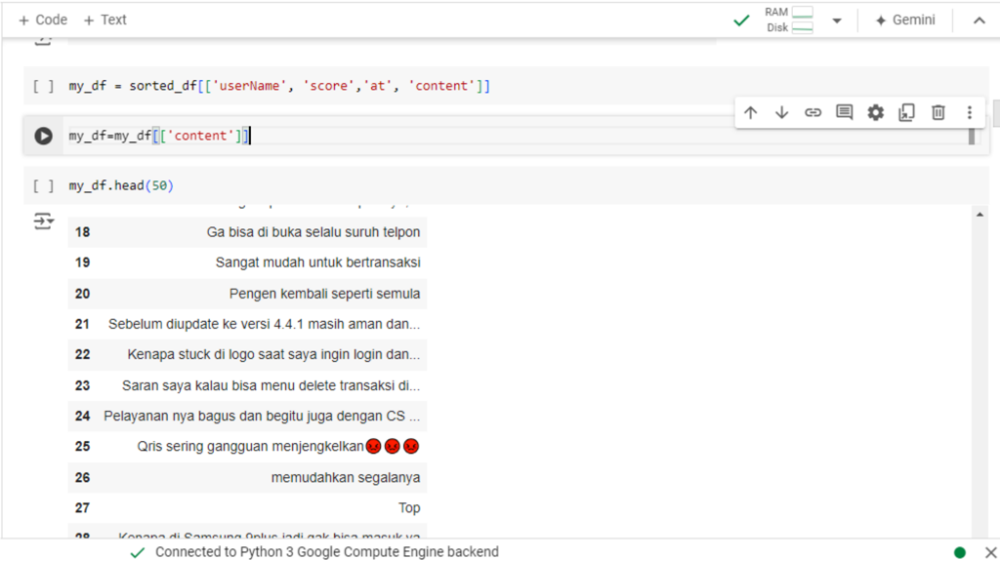

Fig. 3: Attribut Content

Fig 2 and 3 show the results of scraping from Google Colab. After scraping, the data will be filtered because the obtained data has many attribute columns, so only the content attribute will be retained. We will create a variable my_df to store the attributes username, score, at, and content, and then the my_df variable will filter and only keep the content attribute.

3.2. Labeling

In the next step, specifically labeling, data labeling is done manually. This process involves reading each review one by one to label each word that contains emotion as negative feelings and, conversely, each word that conveys praise as positive sentiment[12] the sentiment distribution generated manually labels 716 positive reviews, 484 negative reviews, from a total dataset of 1,200 review rows. Table 1 shows the results of the manual labeling.

Table 1: Manual Labeling

No	Content	Sentiment
1	Bagus sekali karena banyak membantu kita dalam beberapa urusan...	Positive
2	Ribet mau daftar aja banyak syarat	Negative

3.3. Case Folding

Case folding is a technique to convert all letters in a text to lowercase, such as converting "A-Z" to "a-z". This process also removes unnecessary symbols to make the text more uniform and consistent. This aims to facilitate the matching, comparison, and analysis of text data efficiently[13]. Table 2 results from Case Folding.

Table 2: Case Folding Results

No	Content	Case Folding
1	Kalau ganti hp semua transaksi hilang termasuk aktivasi 🤦‍♂️	kalau ganti hp semua transaksi hilang termasuk aktivasi
2	Luar biasa, apalagi menjadi priority... 👍	luar biasa apalagi menjadi priority

3.4. Stopwords Removal

The process of removing common words that are usually used in large quantities and considered meaningless is called stopword removal. Common stopwords include conjunctions such as "is," "to," "in," "from," and "and." The purpose of stopwords is to simplify the text and focus on more creative and meaningful words. Table 3 shows the results of Stopword Removal.

Table 3: Stopwords Removal Results

No	Case Folding	Stopwords Removal
1	bca tidak koperatif pengembalian nasabah salah transfer	bca koperatif pengembalian nasabah salah transfer
2	suka eror akhir akhir ini jd persulit pembayaran	suka eror jd persulit pembayaran

3.5. Tokenize

At this stage, tokenization is the process of breaking down sentences into smaller parts and removing special symbols, resulting in a unique dataset[14]. Table 4 shows the results of the Tokenize.

Table 4: Tokenize Results

No	Stopwords Removal	Tokenize
1	eror ya bca mobile nya	eror,ya,bca,mobile,nya
2	membantu mengatasi transaksi	membantu,mengatasi,transaksi

3.6. Stemming

Stemming is the process of removing affixes from words, such as prefixes and suffixes, to return the word to its base form. This technique helps identify words with the same root. The stemming process is carried out through the mapping and parsing of words using algorithms from the Python library Sastrawi. Table 5 Stemming.

Table 5: Tokenize Results

No	Tokenize	Stemming
1	pembaruan,transaksi,tranfer	baru transaksi tranfer
2	adakan,menu,live,chat,kalo,kendala,tlpn,buang,pulsa	adakan menu live chat kalo kendala tlpn buang pulsa

3.7. Transformasi

This procedure is a critical stage in the subsequent process of the Transformation Stage.

Fig. 4: TF-IDF

Fig 4 shows the results of the TF-IDF process. Data manipulation is necessary for TF-IDF weighting to prepare the data for the analysis that will be conducted. By knowing the ratio to be divided, TF is used to determine how often a word appears in a text. IDF compares the frequency of a word across the entire collection of documents to determine how unique or informative that word is based on the total number of words in the document. Higher IDF values can be found for terms that are used more sporadically across the entire collection. Each word in the document is weighted with TF-IDF, which is the result of multiplying TF with IDF. Where TF IDF Vectorizer and tfidfTransform are used for TF-IDF. The author will use Tfidftransformer to systematically calculate terms using CountVectorizer, calculate the Inverse Document Frequency (IDF) value, and only then will we determine the TF IDF score. The author uses this in the example, completing all three processes simultaneously. The number of words, IDF values, and TF-IDF scores are calculated using the same dataset for the reverse.

3.8. Data Mining

The goal of the current data mining categorization process is to uncover hidden relationships or patterns in the data, which results in the processes of the model that will be run[15]. In this study, the data used consists of 960 training data and 240 testing data, with the testing data used to train the model and test its performance. This division aims to ensure that the model can understand data patterns well during training and can be accurately evaluated on new data that it has never seen before. Fig 5 Results and data split process.

.

```
[ ]
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(data_clean['content'], data_clean['sentiment'],
                                                    test_size = 0.20,
                                                    random_state = 0)
```

```
print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)
```

```
(960,)
(960,)
(240,)
(240,)
```

Fig. 5: Split Data

3.9. Evaluation

The analysis of the combination of Naïve Bayes and Genetic Algorithm (GA) was conducted to evaluate the model's effectiveness after optimization. The Genetic Algorithm was used to select the best features and optimize parameters, such as population size (50) and the number of generations. (50). The results show an improvement in the performance of the Naïve Bayes model, where the accuracy after optimization reached 98% with a precision of 98%. This proves that the Genetic Algorithm significantly enhances the model's ability to detect and classify BCA Mobile user reviews.

```
Confusion_matrix:
 [[ 96   3]
 [  1 140]]
==================================================

Classification Report:
              precision    recall  f1-score   support

    negative       0.99      0.97      0.98        99
    positive       0.98      0.99      0.99       141

    accuracy                           0.98       240
   macro avg       0.98      0.98      0.98       240
weighted avg       0.98      0.98      0.98       240

NB+GA: 0.9833333333333333
```

Fig. 6: After Genetic Algorthm

```
MultinomialNB Accuracy: 0.9458333333333333
MultinomialNB Precision: 0.9215686274509803
MultinomialNB Recall: 0.9494949494949495
MultinomialNB f1_score: 0.9353233830845771
confusion_matrix:
 [[ 94   5]
 [  8 133]]
====================================================

              precision    recall  f1-score   support

    negative       0.92      0.95      0.94        99
    positive       0.96      0.94      0.95       141

    accuracy                           0.95       240
   macro avg       0.94      0.95      0.94       240
weighted avg       0.95      0.95      0.95       240
```

Fig. 7: Before Genetic Algorthm

In the standard Naïve Bayes model, the evaluation shows an accuracy of 95% with a precision of 92%. Out of 240 test data, the model successfully classified 94 data correctly as the negative class (True Negative/TN), but there were 8 positive class data incorrectly classified as negative (False Negative/FN). Conversely, 5 negative class data were incorrectly classified as positive (False Positive/FP), and 133 data were correctly classified as the positive class (True Positive/TP). These results indicate a significant improvement after the application of the Genetic Algorithm.

3.10. Implementation and evaluation Results

The final results of this study show that the application of genetic algorithms as a wrapper successfully optimized the parameters of the naive bayes method, significantly improving the model's performance compared to the use of naive bayes without optimization. Genetic algorithms allow the feature selection process to be conducted iteratively to choose the most relevant features, which contributes to the increase in accuracy and classification capability of the model.In this study, evaluation was conducted using accuracy, precision, and confusion matrix metrics. For the unoptimized naive bayes model, the resulting accuracy was 95%, with a precision of 92%. Table 6 Confusion Matrix.

Table 6: Before After Naïve Bayes And Algorithm Genetic

Model	True Negatif	False Positive	False Negative	True Positive
Naïve Bayes	94	5	8	133
Naïve Bayes + Genetic Algorithm	96	1	3	140

Table 6 is the result of the confusion matrix, from which NB (naïve bayes) & NB + GA (naïve bayes optimized by genetic algorithm feature selection) are derived. The genetic algorithm is used to select the best features while also optimizing hyperparameters, such as a population size of 50 and a generation count of 50. The evaluation results show an increase in accuracy to 98%, with precision rising to 99%. The reduction in FP and FN values indicates that the model with genetic algorithm optimization is capable of effectively minimizing classification errors, resulting in a more accurate and reliable model.

4. Conclusion

After a series of experiments, the optimization of the genetic algorithm on the Naïve Bayes method proved capable of increasing accuracy and reducing classification errors in BCA Mobile sentiment analysis. This combination resulted in an accuracy of 98%, with 3 False Positives (FP) and 1 False Negative (FN), compared to the standard Naïve Bayes accuracy without a wrapper, which only reached 95%. Genetic algorithms use the best fitness and best solution parameters to optimize the model, with evaluation through a population size of 50 generations. This process significantly improves the model's performance in classifying positive and negative sentiments, and also helps Naïve Bayes handle sentiment variations more effectively. In addition to accuracy, metrics such as precision also show significant improvement, making the model more reliable and accurate in sentiment analysis.

References

[1] V. Tundjungsari and S. Wijaya, "Sentimen Analisis Pada Aplikasi E-Branch BCA Menggunakan Metode Naive Bayes," *IKRA-ITH Teknologi Jurnal Sains dan Teknologi*, vol. 8, no. 1, pp. 78–87, Mar. 2024.

[2] D. Wahyu Bhatara and R. Randy Suryono, "Analisis Sentimen Aplikasi BCA Mobile Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine," *https://jurnal.stkippgritulungagung.ac.id/index.php/jipi/article/view/5536*, vol. 9, no. 4, pp. 1907–1917, 2024, doi: 10.29100/jipi.v9i4.5536.

[3] M. Rizky Pratama, Y. R. Ramadhan, and M. A. Komara, "Analisis Sentimen BRImo dan BCA Mobile Menggunakan Support Vector Machine dan Lexicon Based," vol. 12, 2023.

[4] A. Hendra and F. Fitriyani, "Analisis Sentimen Review Halodoc Menggunakan Nai ̈ve Bayes Classifier," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 6, no. 2, pp. 78–89, May 2021, doi: 10.14421/jiska.2021.6.2.78-89.

[5] A. Rahman, E. Utami, and S. Sudarmawan, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 1, pp. 60–71, Jul. 2021, doi: 10.31603/komtika.v5i1.5188.

[6] S. F. Tahir and C. A. Sugianto, "Optimasi Naive Bayes Menggunakan Algoritma Genetika Pada Klasifikasi Komentar Cyberbullying Pada Media Sosial X," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, pp. 3350–3356, Aug. 2024, doi: 10.23960/jitet.v12i3.4834.

[7] W. Astuti, R. Kurniawan, and Y. A. Wijaya, "Jurnal Informatika dan Rekayasa Perangkat Lunak Analisis Data Sentimen Ulasan Aplikasi Dana di Google Play Store Menggunakan Algoritma Naïve Bayes," vol. 6, no. 1, pp. 158–163, 2024.

[8] W. Wahyudi, R. Kurniawan, and A. Y. Wijaya, "Analisis Sentimen Pengguna Terhadap Aplikasi Blu Bca Di Playstore Mengunakan Algoritma Naïve Bayes," vol. 8, no. 3, pp. 2511–2517, May 2024.

[9] J. Teknika, M. K. Rifa, M. H. Totohendarto, and M. R. Muttaqin, "Analisis Sentimen Penguna E-Wallet Dana Dan Gopay Pada Twitter Menggunakan Metode Support Vector Machine (SVM)," *IJCCS*, vol. 17, no. 2, pp. 323–332, Aug. 2023.

[10] A. Suryana, A. I. Purnamasari, and I. Ali, "Mengoptimalkan Kepuasan Pengguna: Analisis Sentimen Review Aplikasi Grab Di Indonesia," Jl.Perjuangan No.10 B Majasem, Kec.Kesambi, Kota Cirebon Jawa Barat 45135, Indonesia, Jun. 2024. Accessed: May 29, 2024. [Online]. Available: https://ejournal.itn.ac.id/index.php/jati/article/view/9688/5524

[11] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier," *Jurnal Minfo Polgan*, vol. 11, no. 2, pp. 30–35, Aug. 2022, doi: 10.33395/jmp.v11i2.11640.

[12] M. K. Khoirul Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 478–483, Mar. 2023, doi: 10.36040/jati.v7i1.6373.

[13] N. Agustina, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 1, pp. 47–54, Apr. 2022, doi: 10.57152/malcom.v2i1.195.

[14] K. Anwar, "Analisa sentimen Pengguna Instagram Di Indonesia Pada Review Smartphone Menggunakan Naive Bayes," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 4, pp. 148–155, Feb. 2022, doi: 10.30865/klik.v2i4.315.

[15] A. F. Watratan, A. P. B, and D. Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," *Journal of Applied Computer Science and Technology*, vol. 1, no. 1, pp. 7–14, Jul. 2020, doi: 10.52158/jacost.v1i1.9.