International Journal of Engineering, Science and InformationTechnology
Volume 5 No. 1 (2025) pp. 229-235
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i1.692
Research Paper, Short Communication, Review, Technical Paper

Sentiment Analysis of Customer Satisfaction Towards Shopee and Lazada E-commerce Platform Using the Random Forest Algorithm Classifier

Tursina Dewi, Asrianda, Yesy Afrillia*

Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia

**Corresponding author Email: yesy.afrillia@unimal.ac.id*

The manuscript was received on 1 June 2024, revised on 10 October 2024, and accepted on 29 December 2024, date of publication 9 January 2025

Abstract

In the digital era, e-commerce platforms like Shopee and Lazada have become the primary channels for online transactions in Indonesia, significantly shaping consumer behaviour and business strategies. This study analyses and compares consumer sentiment towards product reviews on these platforms, focusing on three prominent stores: Skintific, Originote, and Azarine. The research utilized a dataset of 4,500 comments collected from both platforms, with 3,600 comments allocated for training and 900 comments for testing. The sentiment analysis used a lexicon-based approach and machine learning techniques to ensure accuracy and reliability. The results reveal that the Skintific store achieved 88% positive sentiment on Shopee and 84.1% on Lazada. The Originote store recorded 81.4% positive sentiment on Shopee and 91.5% on Lazada, while the Azarine store achieved 87.8% on Shopee and 77.9% on Lazada. These findings highlight variations in consumer sentiment between platforms, which platform-specific features and user demographics may influence. This study provides valuable insights for businesses to tailor their marketing strategies and improve customer engagement on different e-commerce platforms.

Keywords: *E-commerce, Shopee, Lazada, Sentiment Analysis, Random Forest*.

1. Introduction

In the digital era, consumer behaviour has undergone significant changes, marked by the increasing use of e-commerce platforms as practical and efficient tools for buying and selling. As one of the countries with the most crucial number of internet users worldwide, Indonesia has experienced rapid growth in e-commerce adoption, with Shopee and Lazada emerging as dominant platforms. These platforms facilitate online transactions and offer features such as promotions, discounts, and free shipping, which attract consumers and enhance user engagement [1].
Product reviews play a crucial role in influencing consumers' purchasing decisions while also serving as a key indicator of product and seller reputation. [2]. Understanding the sentiment embedded in these reviews is essential for e-commerce platforms and sellers to improve services and enhance customer satisfaction. Sentiment analysis, leveraging Natural Language Processing (NLP) techniques and machine learning, enables businesses to identify consumer perceptions and classify them as either positive or negative sentiments [3].In e-commerce, sentiment analysis can help companies gain deeper insights into customer experiences, identify the strengths and weaknesses of their services, and formulate more effective business strategies [4]. Sentiment analysis is a branch of data mining that identifies opinions on a specific topic. It helps analyze, process, and extract textual data related to services, products, individuals, organizations, events, or issues [5]. Similar to comments on social media platforms like Twitter, sentiment analysis can be performed on tweets to gauge public responses to implementing "Permendikbud Number 30 of 2021." This involves identifying and categorizing the polarity of a text to determine whether a particular document holds a positive or negative value based on predefined categories [6]. This study uses the Random Forest algorithm to analyze and compare consumer sentiment towards the same stores operating on Shopee and Lazada. This research focuses on stores selling skincare products, such as Skintific, Originote, and Azarine, to further explore consumer

perceptions on both platforms. Through this analysis, the study is expected to provide valuable insights that can assist sellers and platform developers enhance service quality. Additionally, this research seeks to significantly contribute to developing sentiment analysis studies in the e-commerce field. Furthermore, the findings are expected to be beneficial not only for sellers but also for customers who wish to have a better shopping experience. By comparing reviews on both platforms, customers can make more informed decisions about the stores and products they choose.

2. Literature Review

In recent years, sentiment analysis using natural language processing techniques has gained significant attention, especially in the context of e-commerce platforms. A study by Nanda Aurelia Salsabila (2024) titled Sentiment Analysis of Tokopedia App Reviews Using Naïve Bayes Classification analyzes user sentiment towards the Tokopedia app to improve user experience using the Naïve Bayes method on R Studio software. The review data was collected from Google Play via scraping, yielding 2,000 reviews sorted by relevance, which were then filtered by removing neutral reviews (3-star ratings), leaving 1,819 data points. The reviews were classified into negative (1 and 2 stars) with 1,481 reviews and positive (4 and 5 stars) with 338 reviews. The analysis process involved exploration, data preparation, and sentiment classification, resulting in an accuracy of 82.97% in identifying positive and negative sentiments [7].

In recent years, sentiment analysis using natural language processing techniques has gained significant attention, especially in the context of e-commerce platforms. A study by Nanda Aurelia Salsabila (2024) titled Sentiment Analysis of Tokopedia App Reviews Using Naïve Bayes Classification analyzes user sentiment towards the Tokopedia app to improve user experience using the Naïve Bayes method on R Studio software. The review data was collected from Google Play via scraping, yielding 2,000 reviews sorted by relevance, which were then filtered by removing neutral reviews (3-star ratings), leaving 1,819 data points. The reviews were classified into negative (1 and 2 stars) with 1,481 reviews and positive (4 and 5 stars) with 338 reviews. The analysis process involved exploration, data preparation, and sentiment classification, resulting in an accuracy of 82.97% in identifying positive and negative sentiments [8].

Next, a study conducted by Talha Ahmed Khan (2024) titled Sentiment Analysis Using Support Vector Machine and Random Forest provides deeper insights into sentiment analysis. It sheds light on the effectiveness of machine-learning approaches in this field. Based on the results obtained, two machine learning algorithms, Random Forest and SVM, were evaluated based on their accuracy in classification tasks. The Random Forest algorithm achieved an accuracy of 0.78564, while SVM performed slightly better with an accuracy of 0.80394. Both algorithms demonstrated their strength in attaining accuracy in the given classification tasks. These results suggest that SVM, with its slightly higher accuracy, maybe a more suitable choice if accuracy is the primary priority. However, the characteristics of the problem at hand should also be considered when selecting the most appropriate algorithm [9].

A recent study by Gunawan (2020) titled Sentiment Analysis System on Product Reviews Using Naive Bayes Method developed a sentiment analysis system for online product reviews in Indonesia using the Naive Bayes method. The system involves five stages: crawling, preprocessing, word weighting using the TF-IDF method, model creation, and sentiment classification. The data is classified into five classes: very negative, negative, neutral, positive, and very positive. The evaluation used confusion matrix testing with accuracy, recall, and precision parameters. The results of testing with three classes (negative, neutral, positive) using 90% training data and 10% testing data showed an accuracy of 77.78%, recall of 93.33%, and precision of 77.78%. In contrast, the five-class testing yielded an accuracy of 59.33%, recall of 58.33%, and precision of 59.33%. Although the results were not entirely accurate, the sentiment class predictions were relevant compared to the classes marked by the supervisor [10].

A recent study by Ramadhan et al. (2022) titled Sentiment Analysis of Product Reviews in Marketplaces Using Naïve Bayes Algorithm analyzes the sentiment of product reviews in marketplaces using the Naïve Bayes algorithm. The study results show this algorithm effectively classifies product review sentiments satisfactorily [11].

A study conducted by Hadiguna Setiawan (2024) titled Comparison of Machine Learning and Deep Learning Algorithms for Sentiment Analysis of Feedback on Lecturer Evaluation evaluates lecturer performance using text analysis with a combination of machine learning, deep learning, and word embedding techniques. The dataset comprised 663 positive, 552 negative, and 465 neutral data points. The results showed that the Bi-LSTM method outperformed others, achieving a training accuracy of 95.91% and a testing accuracy of 72.25%, compared to other methods such as LSTM (70.81%), Random Forest (69.37%), KNN (59.82%), and SVM (62.88%). This method proved effective in identifying areas for improvement and providing insights into lecturer performance, making it the best choice for sentiment analysis in this context [12].

A recent study by Assami Muzaki (2024) titled Sentiment Analysis on Product Reviews in E-Commerce Using the Naive Bayes Method develops a sentiment analysis system for product reviews on e-commerce platforms using the Naive Bayes method to classify sentiments as positive or negative. The system is implemented as an interactive website capable of performing real-time sentiment analysis and accepting CSV files to analyze multiple reviews simultaneously. Using an exploratory approach and system development, product review data is processed through preprocessing stages to prepare the text data. The results show that the Naive Bayes-based system was successfully implemented, creating an interactive website that can efficiently assist in sentiment analysis of product reviews [13].

A recent study by Novresia Wijaya (2024) titled Sentiment Analysis of Instagram App Reviews on Google Play Store aims to develop an accurate sentiment analysis method for Instagram user reviews using a combination of Naïve Bayes and a lexicon-based approach. This approach is designed to address the mismatch between star ratings and the content of reviews. The main challenge faced is accurately analyzing sentiment, given this potential inconsistency. The study successfully determined user reviews' positive and negative feelings with the applied method. The results showed that this method achieved an accuracy of 92%, with precision of 84%, recall of 91%, and an F1-score of 87%, demonstrating the effectiveness of this approach in understanding Instagram user experiences [14].

Overall, the studies indicate that various sentiment analysis algorithms, such as Naive Bayes, SVM, Random Forest, and deep learning, effectively analyze product reviews on e-commerce platforms. Each method has its strengths and limitations, which must be considered based on the objectives and complexity of the data being analyzed. Therefore, further research is needed to explore algorithm combinations that can provide optimal results in classifying sentiment in product reviews on e-commerce platforms and enhance the overall accuracy and efficiency of sentiment analysis.

3. Research Methods

Data collection is a crucial step in preparing the dataset for sentiment analysis. This study collected data through direct observation by manually recording product reviews and ratings from the e-commerce platforms Shopee and Lazada. This method allows researchers to deeply understand customer sentiment, reflected in their opinions within reviews and ratings. The total dataset used in this study consists of 4,528 reviews, with 3,622 reviews (80%) as training data and 906 reviews (20%) as testing data. This division ensures that the model can learn effectively from the training data and be tested on previously unseen data to measure the model's generalization. The reviews include text comments and numerical ratings used to build and test the sentiment analysis model. The preprocessing steps undertaken are as follows:

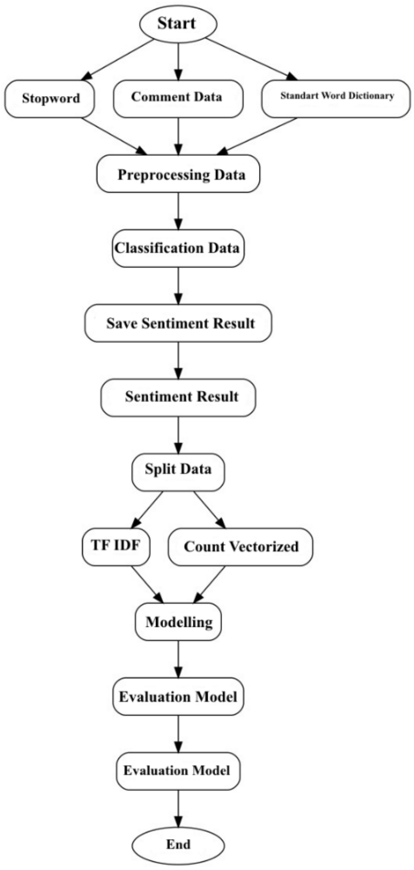

Fig 1. Model Training Flowchart

The sentiment analysis process is carried out through several key stages, from data preprocessing to model evaluation. In the preprocessing stage, the text data is cleaned of irrelevant elements using preprocessing techniques. Preprocessing is a crucial stage in sentiment analysis because raw comment texts typically contain elements that are irrelevant for modeling, such as punctuation, numbers, and special characters. Therefore, there are several steps in preprocessing, which are:

a. Case Folding: Convert all characters in the text to lowercase to eliminate the distinction between uppercase and lowercase letters.
b. b. Tokenizing: Tokenizing is breaking down the text into individual words (tokens).
c. Normalization: The normalization process is carried out to convert non-standard words to their standard forms using a pre-prepared dictionary of standard words.
d. d. Stemming: The stemming process is used to return words to their root or base form (e.g., "membeli" becomes "beli") to reduce word variation that has the same meaning.

Next, stopword removal is performed to eliminate common words that do not carry significant meaning, text normalization is done using a standard word dictionary, and tokenization is used to break the text into word units. This process is concluded with stemming, which reduces words to their root form to standardize the writing pattern.

After the preprocessing stage, feature extraction uses techniques like Count Vectorizer or TF-IDF to convert the text data into a numerical representation that machine learning algorithms can process. The sentiment analysis model is built using the Random Forest algorithm, known for its ability to handle complex datasets. Hyperparameter tuning is performed using GridSearchCV to optimize the model's performance. After training the model, the evaluation uses accuracy, precision, recall, and F1-score metrics to ensure the model can accurately classify positive, negative, or neutral sentiment in the review data.

As the final step, the trained model and the vectorization are saved using the Joblib module, allowing it to be reused for future predictions without retraining. Overall, this process is designed systematically to ensure that the resulting sentiment analysis model can provide accurate and relevant results and offer insights into customer perceptions of products on the two e-commerce platforms studied. The data results from the preprocessing process can be seen in the following image:

	rating	author_username	comment	Text Case Folding	Text Tokenizing	Text Normalization	Text Stemming	Text Filtering
0	5	annissa.saputri	Cocok banget buat di bawa traveling kemasannya...	cocok banget buat di bawa traveling kemasannya...	[cocok, banget, buat, di, bawa, traveling, kem...	[cocok, banget, buat, di, bawa, traveling, kem...	[cocok, banget, buat, di, bawa, traveling, kem...	[cocok, banget, bawa, traveling, kemas, jangka...
1	5	m*****1	Barang sampai sesuai pesanan. Nyampenya juga c...	barang sampai sesuai pesanan nyampenya juga ce...	[barang, sampai, sesuai, pesanan, nyampenya, j...	[barang, sampai, sesuai, pesanan, nyampenya, j...	[barang, sampai, sesuai, pesan, nyampenya, jug...	[barang, sesuai, pesan, nyampenya, cepett, mog...
2	5	v*****a	Wangi:tidak berbau\nPengalaman Penggunaan:bagu...	wangitidak berbau pengalaman penggunaanbaguss ...	[wangitidak, berbau, pengalaman, penggunaanbag...	[wangitidak, berbau, pengalaman, penggunaanbag...	[wangitidak, bau, alam, penggunaanbaguss, bang...	[wangitidak, bau, alam, penggunaanbaguss, bang...
3	5	n*****a	Tekstur:cream\nCocok Untuk:kulit normal\nPerfo...	teksturcream cocok untukkulit normal performag...	[teksturcream, cocok, untukkulit, normal, perf...	[teksturcream, cocok, untukkulit, normal, perf...	[teksturcream, cocok, untukkulit, normal, perf...	[teksturcream, cocok, untukkulit, normal, perf...
4	4	diniemaniz	Pengiriman cepat,barang original cuma biasanya...	pengiriman cepatbarang original cuma biasanya ...	[pengiriman, cepatbarang, original, cuma, bias...	[pengiriman, cepatbarang, original, cuma, bias...	[kirim, cepatbarang, original, cuma, biasa, pa...	[kirim, cepatbarang, original, pakai, kardus, ...

Fig 2. Preprocessing Data

Figure 2 above shows the transformation of the comment data from the initial input in the "comment" column to the final result in the "Text Filtering" column, which displays the cleaned words. This process starts with case folding to convert the text to lowercase, followed by tokenization to split the sentence into individual words. Next, normalization is performed to convert non-standard words into their standard form, followed by stemming to reduce words to their root form. The final step is filtering by removing stopwords, resulting in a more concise and relevant set of words for sentiment analysis.

4. Results and Discussion

This study analyzes user reviews from two e-commerce platforms, Shopee and Lazada, to compare customer sentiment for the same store on different platforms. The data comprises 4,528 reviews, divided into 80% training data (3,622 reviews) and 20% testing data (906 reviews). Sentiment analysis is conducted using two feature representation methods, namely Count Vectorizer and TF-IDF, with the Random Forest classification model.

After the training and testing process, the results show that both methods, Count Vectorizer and TF-IDF, have almost the same accuracy, around 77%, with Count Vectorizer slightly outperforming TF-IDF. Based on the confusion matrix analysis, Count Vectorizer recorded higher True Negatives (457) and True Positives (417) compared to TF-IDF (375 and 321). This indicates that Count Vectorizer performs better at recognizing negative and positive sentiment classes than TF-IDF.

```
Confusion matrix:
 [[457 127]
 [131 417]]
Classification report:
              precision    recall  f1-score   support

    Negative       0.78      0.78      0.78       584
    Positive       0.77      0.76      0.76       548

    accuracy                           0.77      1132
   macro avg       0.77      0.77      0.77      1132
weighted avg       0.77      0.77      0.77      1132

Inset Random Forest Classifier on Count Vectors: 0.7720848056537103
```

Fig 3. Count Vectorizer

```
Confusion matrix:
 [[375 100]
 [110 321]]
Classification report:
              precision    recall  f1-score   support

    Negative       0.77      0.79      0.78       475
    Positive       0.76      0.74      0.75       431

    accuracy                           0.77       906
   macro avg       0.77      0.77      0.77       906
weighted avg       0.77      0.77      0.77       906

Inset Random Forest Classifier on TF-IDF Vectors: 0.7682119205298014
```

Fig 4. TF-Idf Vector

In comparison to other evaluation metrics:

a. The precision for the negative class is almost the same between both methods (0.78 vs. 0.77). However, Count Vectorizer has a slight advantage in recall for the negative class (0.78 vs. 0.77) and performs better in precision and recall for the positive class (0.77 vs. 0.76 and 0.76 vs. 0.74).

b. The F1-score also shows comparable performance, but the Count Vectorizer yields better results, especially for the negative class (0.78 vs. 0.77).

The table below presents a detailed comparison between Count Vectorizer and TF-IDF:

Table 1. Comparison of Count Vectorizer and TF-IDF Models

Matriks	Count – Vectorizer	TF - IDF
Accuracy	77.21 %	76.82%
True Negatives	457	375
True Positives	417	321
Precision (Negatif)	0.78	0.77
Precision (Positif)	0.77	0.76
Recall (Negatif)	0.78	0.77
Recall (Positif)	0.76	0.74
F1-Score (Negatif)	0.78	0.77
F1-Score (Positif)	0.75	0.75

4.1. Sentiment Analysis Results

Here is the result of the model testing on the stores (Skintific, Azarine, and The Originote) on two platforms (Lazada and Shopee):

4.1.1. Model Testing on Skintific Store

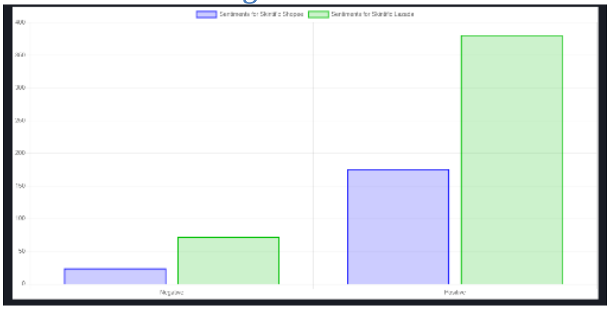

Fig 5. Comparison of Skintific Sentiment on Lazada and Shopee

Table 2. Comparison of Skintific Sentiment on Lazada and Shopee

No	Platform	Number of Data	Sentiment positif (%)	Sentiment Negatif (%)
1	Lazada	453	381 (84.1 %)	72 (15.9 %)
2	Shopee	200	176 (88 %)	24 (12 %)

Based on the data in Table 1, it can be seen that on the Lazada platform, Skintific has a proportion of 84.1% positive and 15.9% negative reviews. Meanwhile, on the Shopee platform, Skintific recorded a higher percentage of positive reviews at 88%, with only 12% negative reviews. This difference indicates that while Skintific has a high satisfaction rate on both platforms, customers on Shopee tend to be more satisfied than those on Lazada. This may reflect differences in the quality of user experience or services each platform provides. For example, factors such as shipping processes, product quality, or even customer service interactions could play a role in the varying satisfaction levels. This data also suggests that Shopee performs better in terms of positive reviews, even though it has a smaller dataset compared to Lazada, which may indicate a tendency for customers on Shopee to give higher ratings.

4.1.2. Model Testing on Azarine Store

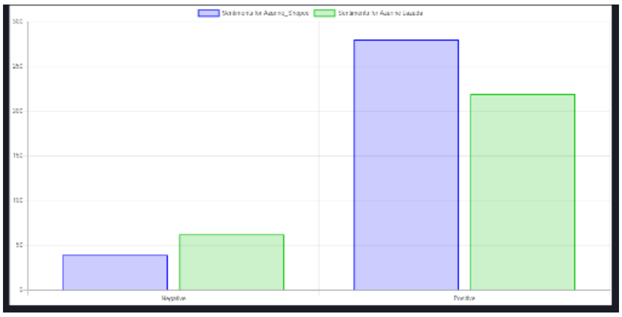

Fig 6. Comparison of Azarine Sentiment on Lazada and Shopee

Table 3. Comparison of Azarine Sentiment on Lazada and Shopee

No	Platform	Number of Data	Sentiment positif (%)	Sentiment Negatif (%)
1	Lazada	281	291 (77.9 %)	62 (22.1 %)
2	Shopee	319	176 (87.8 %)	24 (12.2 %)

Based on the data in Table 1 regarding the comparison of Azarine's sentiment on the Lazada and Shopee platforms, it can be concluded that there is a significant difference in customer satisfaction levels between the two platforms. On the Lazada platform, of the 281 review data, 77.9% expressed positive sentiment, and 22.1% expressed negative sentiment. Meanwhile, on the Shopee platform, of the 319 review data, 87.8% expressed positive sentiment, and 12.2% expressed negative sentiment. This indicates that Azarine customers are likelier to leave positive reviews on Shopee than on Lazada, where dissatisfaction is slightly higher. This difference could be attributed to various factors, including differences in service quality between the platforms, user experience, or even factors related to price or shipping, which may significantly impact customer satisfaction on each platform.

4.1.3. Model Testing on Originote Store

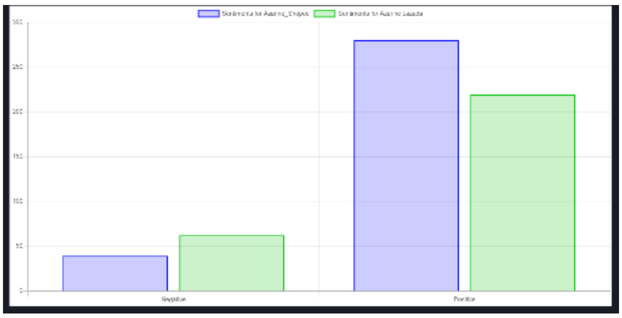

Fig 7. Comparison of Originote Sentiment on Lazada and Shopee

Table 4. Comparison of Originote Sentiment on Lazada and Shopee

No	Platform	Number of Data	Sentiment positif (%)	Sentiment Negatif (%)
1	Lazada	281	291 (77.9 %)	62 (22.1 %)
2	Shopee	319	176 (87.8 %)	24 (12.2 %)

Based on the data in Table 3 regarding the comparison of Originote's sentiment on Lazada and Shopee platforms, it can be seen that customers on both platforms have different responses to the Originote product. On the Lazada platform, of the 440 review data, 91.5% expressed positive sentiment, and only 8.5% expressed negative sentiment. Meanwhile, on the Shopee platform, of the 336 review data, 82.4% expressed positive sentiment, and 17.6% expressed negative sentiment. These results indicate that customers on Lazada tend to be more satisfied with the Originate product than customers on Shopee, as the percentage of positive sentiment on Lazada is higher. On the other hand, while most reviews on Shopee are also positive, there is a higher proportion of negative reviews, indicating a greater level of dissatisfaction on this platform than on Lazada. This could be due to various factors such as shopping experience, customer service, or perceptions of product quality on each platform.

4.2. System Implementation

WordCloud visualization is performed after sentiment classification to display the words that frequently appear in positive and negative comments. Using the WordCloud library, all filtered text is combined to generate an image showing word frequency. This process involves removing common words using a predefined stopwords list, so the visualization focuses more on significant terms related to positive or negative sentiments. The image below shows the WordCloud visualization results for both positive and negative sentiments:

Fig 8. Positive Sentiment WordCloud

Fig 9. Negative Sentiment WordCloud

This study used two methods, and it can be seen that both the Count Vectorizer and TF-IDF methods show almost the same accuracy of around 77%, with the Count Vectorizer model being slightly superior. In the confusion matrix analysis, Count Vectorizer had higher True Negatives (457) and True Positives (417) compared to TF-IDF (375 and 321), indicating its better ability to recognize negative and positive classes. The precision for the negative class was nearly identical between both methods. Still, Count Vectorizer had a slight advantage in recall (0.78 vs. 0.77) for the negative class and was slightly better in both precision and recall for the positive class. The F1-score also showed comparable performance, with the Count Vectorizer achieving better results, especially for the negative class.

```
Confusion matrix:
 [[457 127]
 [131 417]]
Classification report:
              precision    recall  f1-score   support

    Negative       0.78      0.78      0.78       584
    Positive       0.77      0.76      0.76       548

    accuracy                           0.77      1132
   macro avg       0.77      0.77      0.77      1132
weighted avg       0.77      0.77      0.77      1132

Inset Random Forest Classifier on Count Vectors: 0.7720848056537103
```

Fig 9. Model Evaluation with Count Vectorizer

```
Confusion matrix:
 [[375 100]
 [110 321]]
Classification report:
              precision    recall  f1-score   support

    Negative       0.77      0.79      0.78       475
    Positive       0.76      0.74      0.75       431

    accuracy                           0.77       906
   macro avg       0.77      0.77      0.77       906
weighted avg       0.77      0.77      0.77       906

Inset Random Forest Classifier on TF-IDF Vectors: 0.7682119205298014
```

Fig 10. Model Evaluation with TF-IDF

5. Conclusion

The sentiment analysis of customer comments using Random Forest was conducted on 4500 data points from two platforms, Shopee and Lazada, across three stores: Skintific, Originote, and Azarine. A total of 3600 comment data was used for training gathered from both platforms, while 900 comment data were used for testing and taken from both platforms. Using Random Forest with 4500 comments from Shopee and Lazada, the study's results show the following sentiment outcomes: Skintific Official Store achieved a positive sentiment of 88% on Shopee and 84.1% on Lazada. The Originote. indo store had a positive sentiment of 82.4% on Shopee and 91.5% on Lazada. Azarine Cosmetic Official Shop had a positive sentiment of 87.8% on Shopee and 77.9% on Lazada.

This research found that customers frequently discuss and comment on all three products—Skintific, Originote, and Azarine. Each product shows varying positive sentiment across platforms. Skintific received a higher positive sentiment on Shopee than on Lazada, while Originote had a higher positive sentiment on Lazada compared to Shopee. Azarine, on the other hand, received a higher positive sentiment on Shopee than on Lazada.

It can be concluded that the Skintific store, with 200 comments on Shopee and 453 comments on Lazada, shows a better positive sentiment for Shopee. This proves that the Skintific store is more successful in sales on the Shopee platform. For Lazada, improvements should be made to increase sales, such as offering promotions or discounts on best-selling products and enhancing customer service to attract more shoppers. For the Originote product, with 336 comments on Shopee and 440 comments on Lazada, the results are balanced, indicating that the product is performing well in sales on both platforms. Similarly, for the Azarine product, with 319 comments on Shopee and 281 comments on Lazada, the results are also balanced, confirming that the Azarine product is performing well in sales on both Shopee and Lazada platforms.

References

[1] Shofianti, "Dampak E-Commerce Terhadap Perubahan Kebiasaan Belanja Masyarakat Di Indonesia (Studi Kasus: Pembeli Online Di Platform E-Commerce Shopee)," -, pp. -, 2021.

[2] R. Trianto, "Klasifikasi Rating Otomatis pada Dokumen Teks Ulasan Produk Elektronik Menggunakan Metode N-gram dan Naïve Bayes," *Jurnal Informatika Universitas Pamulang,* pp. 295-301, 2020.

[3] A. F. Hidayatullah, "Sentiment Analysis on Twitter using Neural Network: Indonesian Presidential Election 2019 Dataset," *IOP Conference Series: Materials Science and Engineering,* pp. 1-6, 2021.

[4] Kusuma, "Penerapan Machine Learning dalam Analisis Sentimen di E-Commerce," *Jurnal Ilmiah Teknologi Informasi,* pp. 99-108, 2023.

[5] N. Aula, "ANALISIS SENTIMEN REVIEW CUSTOMER TERHADAP PERUSAHAAN EKSPEDISI JNE, J&T EXPRESS DAN POS INDONESIA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)," *Journal of Informatics and Computer Science,* pp. 81-86, 2023.

[6] Y. Afrillia, "Jurnal Informatika: Jurnal pengembangan IT Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine," *G-Tech : Jurnal Teknologi Terapan,* vol. 6, no. 2, pp. 387-384, 2021.

[7] M. Devika, "Analisis Sentimen Pada Ulasan Aplikasi Tokopedia Menggunakan Klasifikasi Naïve Bayes," *PRISMA,* vol. 7, no. -, pp. 44-51, 2024.

[8] V. T. Tarigan, "Perbandingan Algoritma Maching Learning dalam Analisis Sentimen Mobil Listrik di Indonesia pada Media Sosial Twitter/X," *Jurnal Informatika Polinema,* vol. 10, no. 4, 2024.

[9] T. A. Khan, "Sentiment Analysis using Support Vector Machine and Random Forest," *Journal of Informatics and Web Engineering,* vol. 3, no. 1, pp. 67-75, 2024.

[10] B. Gunawan, "Analisis Sentimen Ulasan pada Aplikasi E-Commerce dengan Menggunakan Algoritma Naïve Bayes," *Jurnal Edukasi dan Penelitian Informatika,* vol. 4, no. 2, pp. 113-118, 2020.

[11] Ramadhan, "Analisis Sentimen Ulasan Produk di Marketplace Menggunakan Algoritma Naïve Bayes," *Jurnal Teknologi dan Sistem Informasi,* vol. 7, no. 2, pp. 45-55, 2022.

[12] H. Setiawan, "Analisis Perbandingan Algoritma Machine Learning Dan Deep Learning Untuk Sentimen Analisis Teks Umpan Balik Tentang Evaluasi Pengajaran Dosen," *JSAI: Journal Scientific and Applied Informatics* , vol. 7, no. 2, pp. 379-385, 2024.

[13] A. Muzaki, "ANALISIS SENTIMEN PADA ULASAN PRODUK DI E-COMMERCE DENGAN METODE NAIVE BAYES," *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI),* vol. 5, no. 4, pp. 758-765, 2024.

[14] N. Wijaya, "Analisis Sentimen Ulasan Aplikasi Instagram di Google Play Store: Pendekatan Multinomial Naive Bayes dan Berbasis Leksikon," *Building of Informatics, Technology and Science (BITS)* , vol. 6, no. 2, pp. 921 - 929, 2024.

[15] R. Apriani, "Analisis Sentimen dengan Naïve Bayes terhadap Komentar Aplikasi Tokopedia," *Jurnal Rekayasa Teknologi Nusa Putra,* vol. 6, no. 1, pp. 54-62, 2019.