JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP TIKTOK SHOP MENGGUNAKAN METODE NAÏVE BAYES DAN RANDOM FOREST

Didin Saadduddin El Said[*1], Erza Sofian[2]
[1,2]Sistem Informasi, Universitas Ary Ginanjar
didin.s.el.said@students.esqbs.ac.id

ABTRACT

This study aims to determine the sentiment of the Indonesian people towards TikTok Shop using sentiment analysis. This study uses a quantitative approach by collecting 2340 sentiment data taken from Twitter social media. For sentiment analysis, two machine learning algorithms are applied, namely Naïve Bayes and Random. The first stage is data collection through crawling from Twitter. Then the selection of relevant attributes for analysis is carried out. And the raw data is converted into structured data that is ready to be analyzed. Furthermore, data labeling and classification model development are carried out. Finally, an evaluation of model performance is carried out using and k-fold cross validation to ensure the accuracy of the analysis results. The results show that the Naïve Bayes algorithm produces an accuracy of 66.41%, while the Random Forest algorithm produces an accuracy of 57.81%. Thus, Naïve Bayes is superior with an accuracy difference of 8.60% compared to Random Forest. Sentiment analysis with Naïve Bayes produces 375 positive sentiments and 346 negative sentiments. These results indicate that the public has diverse views on TikTok Shop. This study is expected to provide valuable insights for TikTok Shop managers to improve the quality of their services based on user feedback. In addition, the results of this study can also be a reference for further research in the field of sentiment analysis and e-commerce.

Keywords: *Tiktok Shop, Social Media, Sentiment Analysis, Digital Commerce, Marketing Strategy.*

INTRODUCTION

TikTok Shop is an increasingly popular online shopping platform in Indonesia. With the increasing use of TikTok Shop, it is important to understand how the public responds to this service. This study aims to determine the sentiment of the Indonesian public towards TikTok Shop using sentiment analysis.

A number of new innovations have been introduced to increase sales on various Indonesian e-commerce platforms. One of them is the application of the affiliate link principle, which allows general users to find the products they are looking for through links provided by affiliates. The use of this affiliate link concept has proven effective in boosting sales on various e-commerce platforms in Indonesia. Its development also includes various affiliate programs such as the Tokopedia Affiliate Program, Tiktok Affiliate Program, Lazada Affiliate Program, Shopee Affiliate Program.

Indonesia is the first market where TikTok opened TikTok Shop, and was once the market with the highest Gross Merchandise Value (GMV). Since TikTok Shop was introduced in Indonesia in February 2021, significant achievements have been recorded. In 2022, TikTok Shop GMV in Indonesia reached 25 billion US dollars, contributing 57% of the total GMV of the Southeast Asian market.

Based on SimilarWeb data, Shopee is the e-commerce with the most site visits in Indonesia in the first quarter of 2023. During the January-March period this

year, the Shopee site achieved an average of 157.9 million visits per month, far surpassing its competitors. In the same period, the Tokopedia site achieved an average of 117 million visits, the Lazada site 83.2 million visits, the BliBli site 25.4 million visits, and the Bukalapak site 18.1 million visits per month (Katadata, 2023).

TikTok Shop service was officially closed on Wednesday, October 4, 2023 .This decision was taken by TikTok Indonesia in response to the implementation of new regulations by the Ministry of Trade that prohibit social commerce.

The official closure of TikTok Shop has drawn various reactions from various groups. Reported by detikjatim.com, the reason behind the ban on TikTok Shop services is a licensing issue. The Minister of Cooperatives and Small and Medium Enterprises, Teten Masduki, clarified that TikTok Shop does not have a permit to trade in e-commerce, it only has a permit from the Foreign Trading Company Representative Office (KP3A).

There are three other reasons that led to the ban on TikTok Shop operations in Indonesia, as reported by tribune Jateng (5/10). First, TikTok Shop considered to be carrying out predatory pricing by offering prices below production costs, resulting in market dominance and losses for competitors. Second, buying and selling activities on TikTok Shop are considered a company strategy to collect data on products that are popular among customers through the TikTok Shop algorithm. Third, the large number of imported goods on TikTok Shop that are sold at prices below local products is also another reason.

From the literature analysis, there are differences in research results regarding the effectiveness of Naïve Bayes and Random Forest in sentiment analysis of the TikTok Shop application. Two studies, namely those conducted by (Siswanto et al., 2022) and (Friska Aditia Indriyani et al., 2023) present a positive perspective on Naïve Bayes. Siswanto evaluated the results of sentiment analysis on TikTok using Naïve Bayes and Lexicon-Based, while Fauzi conducted opinion analysis on climate change issues using the Random Forest and Naïve Bayes methods.

Furthermore, research on sentiment towards the elimination of honorary workers (Miftahusalam et al., n.d.-a) provides support for the Random Forest algorithm with an accuracy level of 96.6%. On the other hand, research by (Friska Aditia Indriyani et al., 2023) applied Naïve Bayes and found an accuracy level of 79% in describing public sentiment towards TikTok.

With this comparison, it can be concluded that the research results provide diverse perspectives. As a follow-up study, this study will consider the differences in results to contribute to further understanding regarding the choice of sentiment analysis algorithms on the Tiktok Shop application. This study responds to the differences in results by choosing to directly compare the effectiveness of Naïve Bayes and Random Forest. This step was taken to identify the most effective algorithm in the context of sentiment analysis on Tiktok Shop.

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

METODELOGY RESEARCH

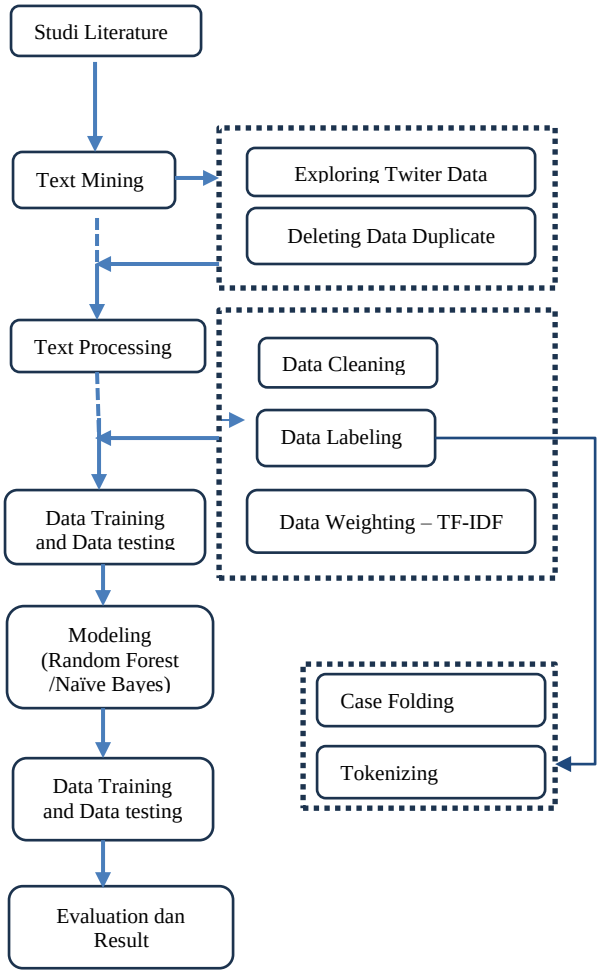


Figure 1. Researh Frame Work

The following is an explanation of the research steps in the research framework:

1. Literature Study

a. Input: Collecting previous research journals.
b. Process: Studying the findings in previous research journals.
c. Output: Results: Obtaining approaches to solving problems when conducting sentiment analysis.

2. Text Mining

a. Input: The data to be used is in the form of tweets from the Twitter application.
b. Process

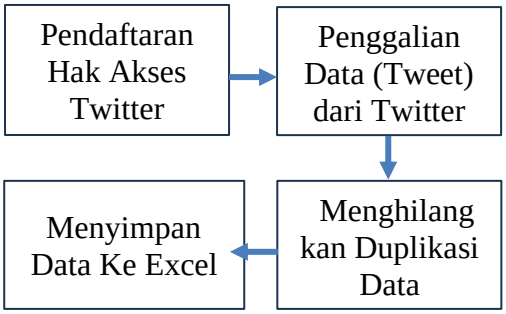


Figure 2. Text Mining Process

Text data collection was carried out using the Rapidminer Studio application. The data collection process includes the following steps:

1) Access registration is carried out through a Rapidminer account to allow data retrieval
from Twitter.
2) Data extraction (tweets) is carried out by searching for the keyword "Tiktok Shop".
3) Before saving the data mining results, duplicate data is removed to ensure the cleanliness of the data results.
4) The data results that have gone through the clearing and extraction process are stored in excel file format. Text preprocessing which consists of:

1) Cleaning tweets data so that it only contains pure text and also applying data deletion that still escapes during data mining
2) Manually selecting tweets for tweets that have no meaning because they will become noise when applied to the classification.

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

3) Labeling each tweet to indicate the sentiment category that the tweet has.
4) Changing all capital letters in tweets to lowercase.
5) Tokenizing words from tweets.
6) Giving word weighting to tweets using the TF-IDF method.

3. Text Processing

a. Input: Tweets data that has been successfully obtained from the text mining process

b. Process:

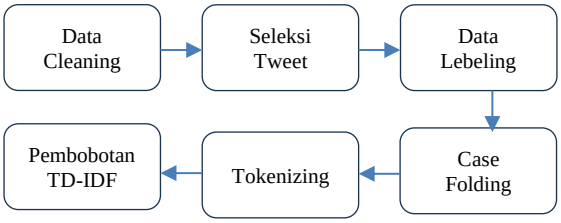


Figure 3. Data Labeling Process

Figure 2 is a picture of the process flow when doing text preprocessing, the flow of which consists of:

1) Cleaning tweets data so that it only contains pure text and also applying data deletion
that still escapes during data mining
2) Manually selecting tweets for tweets that have no meaning because they will become
noise when applied to the classification.
3) Labeling each tweet to indicate the sentiment category of the tweet.
4) Changing all capital letters in tweets to lowercase.
5) Creating word tokenization from tweets.
6) Giving word weighting to tweets using the TF-IDF method.

c. Output: Getting data that is ready to be used in the sentiment analysis process.

4. Traininga and Testing Model

a. Input: The data that has been processed in the previous stage will be divided into two to be used as training data and as testing data using K-fold Cross validation.

b. Process:

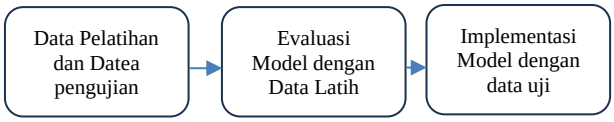


Figure 4. Process of Training Algorithm Model

The flowchart in Figure 4. illustrates the process of training an algorithm model, which involves the following steps:

1) Compiling training data and testing data by determining the K value in cross validation
2) The model is formed using training data.
3) The model that has been formed is tested using testing data.

b. Output: Test results from each algorithm.

5. Result Evaluation

a. Input: Results of testing each Algorithm
b. Process: Selecting the best test results for each algorithm to then be used as a comparison between the two algorithms used.
c. Output: Best algorithm

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

RESULTS AND DISCUSSION

1. Crawling Data

Crawling data is the first step in collecting data to be processed to the text mining stage. After doing the crawling process, then the data is stored in the form of an Excel format, the results of the output of the crawling process will be stored in Excel format.

Table. 1 Table of Crawling Data

Source_ Name	Created_At	Full_Text	In_Reply	Lang
Crawling.csv	Thu May 16 02:15:24 2024	Kualitas barang TikTokShop cukup sesuai dengan harga	wildf_dandelion	in
Crawling.csv	Thu May 16 01:53:33 2024	Pelayanan TikTokShop cukup baik	alifmegat	in
Crawling.csv	Thu May 16 01:52:17 2024	Customer service TikTokShop cukup responsif	rowsih	in

In this researcher, the crawling process uses the keyword "Tiktok Shop". Crawling data was conducted on 18-20 May 2024 totaling 4621 tweets. After getting data and saving in excel format, the next step is to do text preprocessing this aims to reduce or eliminate noise (disturbance) in order to facilitate the next step.
The results of the data crawling process produce 13 attribute columns, but to facilitate the data analysis process, therefore in this study only uses text attributes.

2. Pre-processing Data

At this stage, a pre-processing process is carried out on data that has been collected through crawling techniques. Pre-processing is an important step in data analysis, especially in natural language processing (NLP). This step aims to clean and prepare data to be ready for further analysis. In this study, pre-processing includes several main stages: the removal of duplicate data, normalization of text, removal of punctuation and numbers, as well as removal of words that are not meaningful (stop words). Pre-processing is an essential step to get better insight than complex and varied text data.

2.1 Cleansing

In this cleaning process serves to clean or remove unnecessary attributes and has no significant meaning on tweet data, such as hashtag, mention, retweets, whitespace, links, and other symbols, so that it can be processed to do the next process. The following is a description of the steps of the Cleaning stage:

1. Input Crawling results with Excel to RapidMiner format.
2. Elimination of the word RT@. In this process the word RT and Mention or Tag will automatically disappear.

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

Table 3. Result of Cleansing Link

Initial Text	Final Text
masih ada nih di tiktokshop https://t.co/5eM7dAfmH8	masih ada nih di tiktokshop

3. Hashtag word deletion. In this process, data that has hashtags will be automatically lost.

Table 4. Result of Hastag Cleansing.

Initial Text	Final Text
#tiktokshopindonesia Harga yang sangat kompetitif. Sangat merekomendasikan TikTokShop	Harga yang sangat kompetitif. Sangat merekomendasikan TikTok Shop

4. Deleting symbols. Next, every tweet data that has symbols will be cleaned, and will be automatically deleted.

Text Awal	Text Akhir
Customer Service TikTokShop cukup responsif meski kadang lambat TikTokShop	customer service tiktokshop cukup responsif meski kadang lambat tiktokshop

5. Stopwords Removal

Text Awal	Text Akhir
(*&^$ Pelayanan TikTokShop biasa saja. Tidak terlalu bagus tapi juga tidak terlalu buruk	Pelayanan TikTokShop biasa saja. Tidak terlalu bagus tapi juga tidak terlalu buruk

3. Tokenizing

The tokenization process is carried out to convert raw text into smaller units known as tokens. Tokenization helps in handling text data in a more structured and organized manner, enabling more efficient and accurate analysis.

Table 6. Hasil Tokenize

Intial Text	Final Text
Kualitas barang di TikTok Shop, rata-rata sesuai dengan harganya TikTokShop..	Kualitas barang di TikTok Shop rata rata sesuai dengan harganya TikTokShop

4. Transform Case

In text data processing, one of the important steps that needs to be done is letter transformation or transformation is the process of changing all characters in the text to lowercase or uppercase, in this process all letters in the data set will be changed to lowercase.

Tabel 7. Result of Transform Case

The stopwords removal stage is the process of removing words that are

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

frequently used but do not affect the sentiment of a sentence. In this study, stopwords removal was carried out using a dictionary which is a corpus of stopwords in Indonesian

Tabel 8. Result of Stopwords Removal

Initial Text	Final Text
Aku jajan di tiktokshop slalu puas sih	jajan tiktokshop slalu puas sih
Beli di tiktokshop lebih puas	beli tiktokshop puas
kak di tiktok shop ada kak bagus juga	kak tiktok shop kak bagus

6. Filter Token(by Length)

The Filter Tokens (by Length) Stage is the process of removing words that are too short or too long,

with the provision that the word must consist of a minimum of 4 letters and a maximum of 25 letters.

Table 9. Result of Filter Tokens (by Lenght)

Initial Text	Final Text
kalo minta uang ke nyokap selalu di marahin katanya jajan mulu padahal beliau cekot tiktok shop live SETIAP HARI	kalo uang nyokap marahin jajan mulu beliau cekot tiktok shop live

7. Steam

The Steam stage is a process used to process text by cutting words into their basic form, or root words.

Table 10. Result of Steam

Text Awal	Text Akhir
"Event diskon besar-besaran menarik banyak pembeli"	"Event diskon besar-besar tarik banyak beli"

3. Modeling

The next stage is machine learning modeling using the Naïve Bayes and Random Forest algorithms, with the Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction method. This study divides the data (tweets) that have gone through the text

pre-processing process into two parts: training data and testing data with a ratio of 5:5. After that, feature extraction is carried out from the tweets so that machine learning modeling can be carried out using the training data to predict sentiment labels on tweets in the testing data (Alizah et al., 2020). Of the total 2340 data, 1170 are used as training data and the other 1170 as testing data. Manual data labeling is used to train machine learning algorithms. According to research (Siswanto et al., 2022), this manual labeling functions to train machine learning to recognize sentence patterns that will be labeled automatically. This manual labeling process involves observing text patterns in tweets to determine whether the tweet contains support or hate speech. While testing data will be processed without labels.

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

Table 11. Label Data

Tweet	**Sentimen**
Kualitas barang TikTokShop ratarata sesuai dengan harganya.	Positif
Customer service TikTokShop cukup responsif	Positif
Harga produk di TikTokShop sangat kompetitif Nilai tambah untuk TikTokShop	Positif
Pengiriman sangat terlambat barang tidak sampai tepat waktu tiktokshop	Negatif
pernah nanyain stok papan skate di live tiktok shop yg jual gaenak bgt jawabnya	Negatif

3.1 Naïve Bayes

This study divides the dataset into two parts, namely training data and test data, with a ratio of 50% for training data and 50% for test data (Alizah et al., 2020).

The following is the Naïve Bayes model process using rapidminer software with the operators provided.

3.2 Random Forest

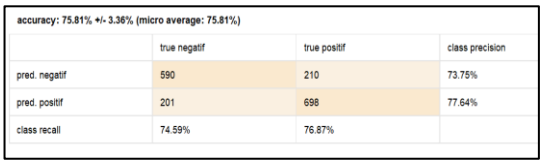
Similar to the use of the Naïve Bayes method, this study divides the data set into two parts, namely test data and training data with a test and train ratio of 50% training data and 50% test data.

3.3 K-Fold Cross validation

After obtaining the training and testing data models on the resulting model, the two algorithms, namely Naïve Bayes and Random Forest, then compare the algorithms, where the algorithm with the highest accuracy will be used for model testing. The results of the evaluation of these two methods are in the form of Confusion matrix values containing accuracy, precision and Recall values taken from the test data. In determining the confusion matrix value, this study uses k-fold cross validation with a value of k = 10 so that the resulting value is maximized (Haganta Depari et al., n.d.)

3.3.1 Naïve Bayes

The cross-validation process with the Naïve Bayes algorithm includes the training and testing stages (Apply Model and Naïve Bayes performance). After that, modeling is carried out using the Naïve Bayes Algorithm to obtain accuracy,precision, and Recall values based on the performance vector (performance-NB).


accuracy: 75.81% +/- 3.36% (micro average: 75.81%)

	true negatif	true positif	class precision
pred. negatif	590	210	73.75%
pred. positif	201	698	77.64%
class recall	74.59%	76.87%	

Figure 5. Confusion Matrix**Naïve Bayes**

The following equation is the calculation of Naïve Bayes accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$= \frac{698 + 590}{698 + 590 + 201 + 210} = \frac{1288}{1699}$$

$$= 75.81\%$$

The following equation is the calculation of the precision values of the positive and negative classes of the method *Naïve Bayes* :

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541-576X

$$Presisi\ Positif = \frac{TP}{TP + FP}$$
$$= \frac{698}{698 + 201} = \frac{698}{899}$$
$$= 77.64\%$$

$$Presisi\ Negatif = \frac{TN}{TN + FN}$$
$$= \frac{590}{590 + 210} = \frac{590}{800}$$
$$= 73.75\%$$

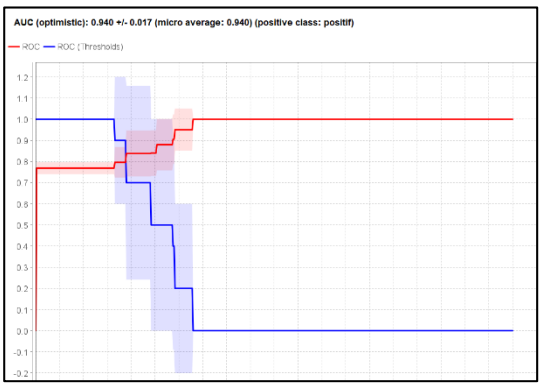
The following is the calculation equation for positive and negative Recall values using the method *Naïve Bayes* :$Recall\ Positif = \frac{TP}{TP+FN} = \frac{698}{698+210} = \frac{698}{908} =$ 76.87%

$$Recall\ Negatif = \frac{TN}{TN + FP} = \frac{590}{590 + 201}$$
$$= \frac{590}{791} = 74.59\%$$

The accuracy value obtained is 75.81% with a margin of error of ± 3.36%. The precision for positive predictions is 77.64% and the precision for negative predictions is 73.75%. The Recall value for positive data is 76.87% and the Recall for negative data is 74.59%.

From the accuracy results that have been described, here is a ROC-AUC graph to evaluate the algorithm's capabilities.

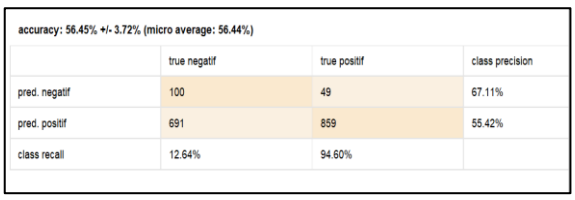
Figure 6. Graphic ROC-AOC of *Naïve Bayes*



The ROC-AUC graph above shows that the Naïve Bayes model has excellent performance with an AUC value of 0.940. This means that the model is able to distinguish between positive and negative sentiments with high accuracy. The ROC curve close to the upper left corner indicates a high True Positive Rate (TPR) and a low False Positive Rate (FPR), indicating good classification performance. An AUC value close to 1.0 reflects the model's ability to classify sentiments with a low level of uncertainty.

3.3.2 Random Forest

The cross validation process with the Naïve Bayes algorithm includes the training and testing stages (Apply Model and Naïve Bayes performance). After that, modeling is carried out using the Naïve Bayes Algorithm to obtain accuracy, precision, and Recall values based on the performance vector (performance-NB).


accuracy: 56.45% +/- 3.72% (micro average: 56.44%)

	true negatif	true positif	class precision
pred. negatif	100	49	67.11%
pred. positif	691	859	55.42%
class recall	12.64%	94.60%	

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

The following equation is the calculation of Random Forest accuracy:

Figure 7 . Result of Confuion Matrix of Random Forest

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$= \frac{859 + 100}{859 + 100 + 691 + 49}$$
$$= \frac{959}{1699} = 56.45\%$$

The following equation is the calculation of the precision values of the positive and negative classes of the method *Random Forest*:

$$Presisi\ Positif = \frac{TP}{TP + FP}$$
$$= \frac{859}{859 + 691} = \frac{859}{1550}$$
$$= 55.42\%$$

$$Presisi\ Negatif = \frac{TN}{TN + FN}$$
$$= \frac{100}{100 + 49} = \frac{100}{149}$$
$$= 67.11\%$$

The following is the calculation equation for positive and negative Recall values using the Random Forest method:

$$Recall\ Positif = \frac{TP}{TP + FN} = \frac{859}{859 + 49}$$
$$= \frac{859}{908} = 94.60\%$$

$$Recall\ Negatif = \frac{TN}{TN + FP}$$
$$= \frac{100}{100 + 691} = \frac{100}{791}$$
$$= 12.64\%$$

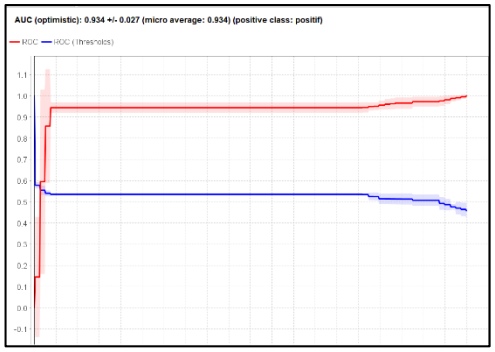
The following is the equation for calculating the positive and negative F1-score values of the Random Forest method:

$$F1 - Score\ Positif$$
$$= 2 \times \frac{Presisi\ Positif\ \times Recall\ Positif}{Presisi\ Positif + Recall\ Negatif}$$
$$= 2 \times \frac{0.5542 \times 0.9460}{0.5542 + 0.9460} = 2 \times \frac{0.5245}{1.5002}$$
$$= 0.6988 \approx 69.88\%$$

$$F1 - Score\ Negatif$$
$$= 2 \times \frac{Presisi\ Negatif\ \times Recall\ Negatif}{Presisi\ Negatif + Recall\ Negatif}$$
$$= 2 \times \frac{0.6711\ \times 0.1264}{0.66711 + 0.1264} = 2 \times \frac{0.0849}{0.7975}$$
$$= 0.2128 \approx 21.28\%$$

The accuracy value obtained is 56.45% with a margin of error of ± 3.72%. The precision for positive predictions is 55.42% and the precision for negative predictions is 67.11%. The Recall value for positive data is 94.60% and the Recall for negative data is 12.64%. The F1-Score value for positive predictions is 69.88% and for negative predictions is 21.28%. From this stage, it is known that the accuracy value of Naïve Bayes is superior

to the accuracy of Random Forest with an accuracy comparison of 75.81% for Naïve Bayes and 56.45% for Random Forest.



JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

Figure 8. Result of Confusion Matrix of Random Forest

The ROC-AUC graph above shows the performance of the Random Forest model with an AUC value of 0.934. This indicates that the Random Forest model has excellent ability to distinguish between positive and negative sentiments. The ROC curve close to the upper left corner indicates a high True Positive Rate (TPR) and a low False Positive Rate (FPR), indicating excellent classification performance. An AUC value close to 1.0 reflects that the model has strong predictive ability and a low error rate in classifying sentiment.

Sentiment Analysis

In the previous stage, an accuracy comparison was carried out and the accuracy values of the two algorithms were obtained. The test results showed that the accuracy value of Naïve Bayes was superior to Random Forest. Therefore, at this stage, labeling will be carried out on the test data that we have also obtained in the previous stage using Naïve Bayes. From the data labeling process carried out using the Naïve Bayes algorithm, the results of the testing data were obtained in the form of 620 entries labeled with positive sentiment and 550 other entries labeled with negative sentiment. Then added training data with 599 positive sentiments and 571 negative sentiments, so the total of the whole is 1219 positive sentiments and 1121 negative sentiments.

Table 12. Label Data

No	Sentimen	Prediction	Text
1.	?	Negatif	mahal tiktok shop
2.	?	Positif	coba cocok skin tone
3.	?	Positif	adaaa kmrn beli tiktokshop murah

CONCLUSION

This study has implemented a text classification model from Twitter related to the Tiktok Shop application, which is then used to label text classes using the Naïve Bayes and Random Forest methods. In addition, word cloud visualizations are created and the accuracy, precision, and Recall values are evaluated. From the results of the classification analysis carried out, the following conclusions can be drawn in this study:

1. After going through the text preprocessing process, 2621 tweets of data were obtained, which were then separated into test data and training data with a ratio of 5:5, then modeling and accuracy testing were carried out for both methods, namely Naïve Bayes and Random Forest with the highest accuracy produced by Naïve Bayes. The Naïve Bayes method produces 620 positive sentiments and 550 negative sentiments.
2. The Naïve Bayes method produces an accuracy value of 75.81% with a positive prediction precision of 77.64%, a negative prediction of 73.75%, a positive data recall of 76.81%, and a negative data recall of 74.59%. The Random Forest method produces an accuracy value of 56.45% with a positive prediction precision of 55.42%, a negative prediction of 67.11%, a positive data recall of 94.60%, and a negative data recall of 12.64%. 3. From the assessment of accuracy, precision, and recall between the two classification methods, it is known that the Naïve Bayes method shows a higher level of accuracy, reaching 66.41%, exceeding the performance of the Random Forest method. Therefore, based on the results

of the evaluation, it was decided to use the Naïve Bayes method to process the data further.

References

Nachouki, Mirna, et al. "Student Course Grade Prediction Using the Random Forest Algorithm: Analysis of Predictors' Importance." Trends in Neuroscience and Education (2023): 100214.

Al Amrani, Yassine, Mohamed Lazaar, and Kamal Eddine El Kadiri. "*Random Forest* and support vector machine, based hybrid approach to sentimen analysis." *Procedia Computer Science* 127 (2018): 511-520.

Miftahusalam, Akhmad Miftahusalam. "Perbandingan Metode *Random Forest* dan *Naïve Bayes* pada Analisis Sentimen Review Aplikasi BCA Mobile." *SIPTEK: Seminar Nasional Inovasi dan Pengembangan Teknologi Pendidikan*. Vol. 1. No. 1. 2023.

Rahmadani, Putri Suci, et al. "Tiktok Social Media Sentimen Analysis Using the Nave Bayes Classifier Algorithm." *Sinkron: jurnal dan penelitian teknik informatika* 7.3 (2022): 995-999.

Zulqornain, Junda Alfiah, Indriati Indriati, and Putra Pandu Adikara. "Analisis sentimen tanggapan masyarakat aplikasi tiktok menggunakan metode *Naïve Bayes* dan categorial propotional difference (cpd)." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 5.7 (2021): 2886-2890.

Cahyani, Rafiqa, Indri Sudanawati Rozas, and Nita Yalina. "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019." *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi (Journal of Computer Science and Information Technology)* 12.1 (2020): 79-86.

Wulandari, Delima Ayu, Rd Rohmat Saedudin, and Rachmadita Andreswari. "Analisis Sentimen Media Sosial Twitter Terhadap Reaksi Masyarakat Pada Ruu Cipta Kerja Menggunakan Metode Klasifikasi Algoritma *Naïve Bayes*." *eProceedings of Engineering* 8.5 (2021).

Siswanto, Siswanto, et al. "The Sentimen Analysis Using *Naïve Bayes* with Lexicon-Based Feature on TikTok Application." *Jurnal Varian* 6.1 (2022): 89-96.

Van der Heide, E. M. M., et al. "Comparing regression, *Naïve Bayes*, and *Random Forest* methods in the prediction of individual survival to second lactation in Holstein cattle." *Journal of dairy science* 102.10 (2019): 9409-9421.

Fitri, Veny Amilia, Rachmadita Andreswari, and Muhammad Azani Hasibuan. "Sentimen analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using *Naïve Bayes*, decision tree, and *Random Forest* algorithm." *Procedia Computer Science* 161 (2019): 765-772.

Fitri, F., Gunawan, M. A., & Fauzi, P. P. A. (2023). "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma *Naïve Bayes* , *Random Forest* Dan Support Vector Machine." Jurnal Transformatika.

Gunawan, D., Dwiza, R., Ardiansyah, D., Akba, F., & Alfariz, S. (2020). "Komparasi Algoritma Support Vector Machine Dan *Naïve Bayes* Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023." Jurnal Teknologi Informasi Dan Ilmu Komputer.

JURSIMA
Jurnal Sistem Informasi & Manajemen
ISSN : 2338-1523, E-ISSN: 2541- 576X

Rizki, M. F., Auliasari, K., & Primaswara Prasetya, R. (2021). "Analisis Sentimen Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine." JATI (Jurnal Mahasiswa Teknik Informatika).

Nugraha, K. A. (2021). "Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan *Naïve Bayes*." Jurnal Teknik Informatika Dan Sistem Informasi.

Alizah, M.D., Nugroho, A., Radiyah, U. and Gata, W., 2020. Sentimen Analisis Terkait Lockdown pada Sosial Media Twitter. Indonesian Journal on Software Engineering (IJSE), 6(2), pp.223-229.