Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

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

Sentiment Analysis to Classify Tiktok Shop Users on Twitter with Naïve Bayes Classifier Algorithm

Ayu Lestari[1*], Ade Irma Purnamasari[2], Agus Bahtiar[3], Edi Tohidi[4]

[1,2,3,4]*STMIK IKMI Cirebon*
ayulestari61219@gmail.com[1], irma2974@yahoo.com[2], agusbahtiar038@gmail.com[3], editohidi00@gmail.com[4]*

Abstract

Advances in information technology have facilitated the use of social media as an e-commerce platform, with TikTok Shop enabling in-person transactions. This research addresses the gap in understanding user perceptions of TikTok Shop through sentiment analysis on Twitter. Sentiment classification is performed using the Naïve Bayes Classifier algorithm. The dataset consists of 1,907 Indonesian tweets, collected from January 2023 to July 2024, and processed using RapidMiner in the Knowledge Discovery in Database (KDD) framework. The preprocessing stages include data cleaning, normalization, tokenization, stopword removal, and stemming. To overcome data imbalance, Synthetic Minority Oversampling Technique (SMOTE) was applied. The model achieved 93.98% accuracy, with balanced precision and recall for positive, neutral, and negative sentiments. The sentiment distribution among TikTok Shop users on Twitter was 35.5% positive, 35.5% negative, and 29.0% neutral. This research provides insights into consumer behavior on social media and emphasizes the importance of sentiment analysis to increase user engagement and understand market perception. This research is expected to provide information to platform developers and businesses looking to improve TikTok.

Keywords: *Tiktok Shop; sentiment analysis; Naïve Bayes Classifier; Twitter;Synthetic Minority Oversampling Technique (SMOTE)*

1. Introduction

Rapid advances in information technology have triggered profound transformations in various sectors, including digital and social commerce. Social media platforms such as TikTok have evolved into e-commerce facilitators through its TikTok Shop feature, which enables direct transactional interactions within the app. User sentiment towards TikTok Shop, expressed on platforms such as Twitter, provides important data to evaluate the acceptance of such services [1].

However, understanding this sentiment presents challenges, especially in classifying user opinions into positive, negative or neutral categories. Factors such as product quality, delivery efficiency, and customer service influence user opinions [2]. TikTok Shop faces an urgent need to understand these sentiment trends to improve service quality and reputation in the market.

This research aims to analyze TikTok Shop's user sentiment on Twitter using the Naïve Bayes Classifier algorithm. With this approach, the research is expected to provide strategic insights for TikTok Shop managers to improve user experience and service quality, and contribute to the development of more effective e-commerce strategies [3],[4], [5], [6]

2. Literature Review

2.1. Data Mining

Data mining is the process of discovering patterns and knowledge from large amounts of data, such as tweets on Twitter that mention TikTok Shop. In sentiment analysis, data mining techniques are used to extract information from large datasets, for example by using appropriate classification techniques to classify the sentiment of TikTok Shop users. Research by [7], showed that sentiment analysis of social media texts can provide valuable insights to understand user perceptions of platforms such as TikTok Shop on Twitter.

2.2. Klasifikasi

Classification is the process of mapping data to predefined categories.In sentiment analysis of TikTok Shop on Twitter, elaboration techniques are used to classify user tweets as positive, negative, or neutral.The research of [8], [9], [10] investigated sentiment analysis methods applied in Twitter, identifying that elaboration procedures such as Naive Bayes & SVM put good output on sentiment elaboration for social media data. Another study by [11]discussed the application of K-Means splitting procedure for data mining, which can be customized to identify patterns in a dataset of tweets related to TikTok Shop. Classification is the process of mapping data to predefined categories. In sentiment analysis of TikTok Shop on Twitter, the elaboration technique is used to classify user tweets as positive, negative, or neutral categories.

2.3. Algoritma Naïve Bayes Classifier

Naive Bayes algorithm is a classification algorithm that is often used in sentiment analysis because of its ability to handle large and diverse text data. Research by [12] discussed the application of the Naive Bayes algorithm to analyze sentiment from reviews on Twitter and found that this algorithm is very effective in classifying sentiment even though text data contains a lot of noise. [13] also used Naive Bayes to analyze the sentiment of product reviews on Tokopedia, which gave effective results in sentiment analysis of e-commerce products, relevant to the application of sentiment analysis on TikTok Shop on Twitter.

3. Research Methods

3.1. Research Methods

This research uses a quantitative method with a data mining approach, using the Knowledge Discovery In Database (KDD) process for research entitled Naïve Bayes Classifier Algorithm to improve the Tiktok Shop user sentiment classification model on Twitter. the steps of the research method can be shown in the figure 1 which includes data understanding, data cleaning, feature extraction, sentiment model building, model evaluation, result analysis, and technical implementation using the Naïve Bayes Classifier algorithm [14].

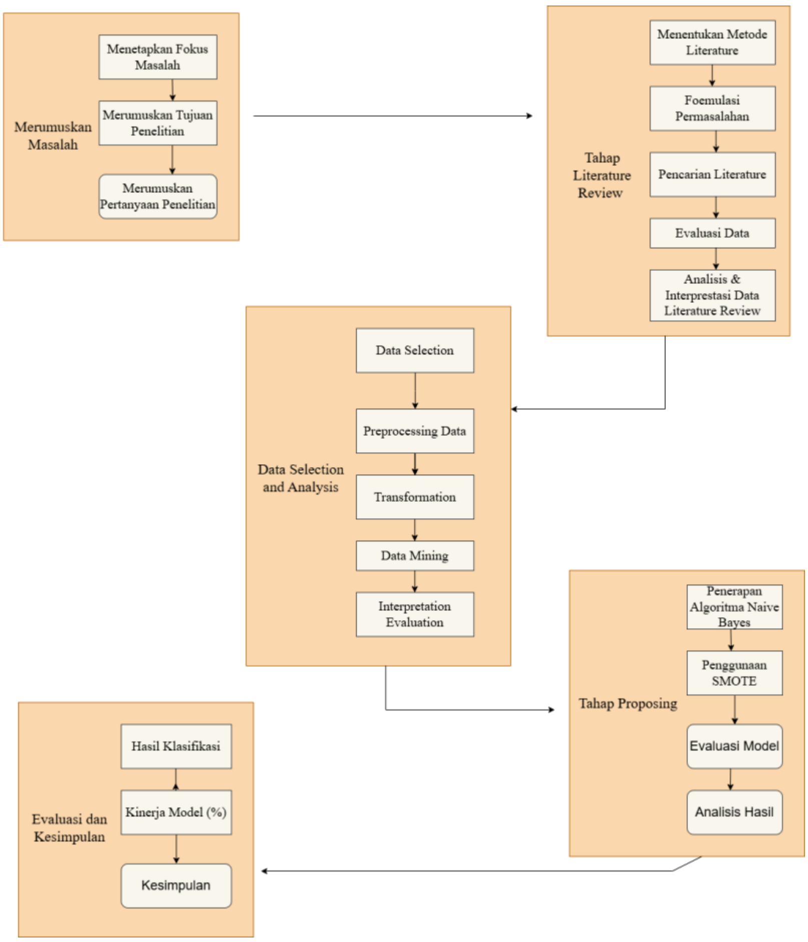

Fig.1: Description of Research Method Activities

Table 1: Description of Research Method Activities

Tahapan	Aktivitas	Deskripsi Aktivitas
Formulating the Problem	Establishing Problem Focus	Determining specific aspects or topics of Tiktok Shop to analyze, such as user shopping experiences.
	Formulating Research Objectives	Outlining the main objectives of this research, such as understanding netizen sentiment towards Tiktok Shop.
	Formulating Research Questions	Formulating specific research questions.
Literature Review	Determining Literature Method	Selecting methods to review relevant literature.
	Problem Formulation	Identifying and formulating problems based on literature review.

	Literature Search	Searching for relevant articles and journals using keywords such as Analysis, Naïve Bayes Classifier, and E-commerce.
	Data Evaluation	Evaluating and assessing the quality and relevance of the data from the literature found.
	Analysis and Interpretation of Literature Review Data	Analyzing and interpreting data from literature to build a strong research foundation.
Data Collection and Analysis	Data Collection	Collecting tweets related to Tiktok Shop on Twitter using Tweet-Harvest and including Twitter authentication from the Twitter account.
	Data Selection	Choosing and ensuring relevant data from the dataset by selecting only certain criteria such as Tiktok Shop.
	Preprocessing	Cleaning data from noise such as symbols, URLs, and irrelevant characters, as well as handling duplicates and missing data to enhance raw data quality.
	Transformation	Converting raw data into relevant features through tokenization, stemming, and text normalization for sentiment classification.
	Data Mining	Using the Naïve Bayes Classifier algorithm to build a sentiment classification model, where data is trained and tested to classify Twitter users' sentiments (positive, negative, or neutral) regarding Tiktok Shop.
	Interpretation Evaluation	Evaluating classification results with metrics such as accuracy, precision, and recall to assess model performance and determine if the model can be further optimized.
Proposing Stage	Application of Naïve Bayes Algorithm	Applying the Naïve Bayes Algorithm to ensure the model has been well-trained using training data.
	Use of SMOTE	Utilizing SMOTE technique to address class imbalance issues in the dataset so the model can learn from more balanced data.
	Model Evaluation	Using test data to measure model performance with evaluation metrics such as accuracy, precision, and recall.
	Results Analysis	Analyzing classification results to understand user sentiment distribution and discussing the implications of the findings.
Evaluation and Conclusion	Classification Results	Presenting the final results of the sentiment classification performed, including the number and percentage of each sentiment category
	Model Performance	Summarizing model performance based on the calculated evaluation metrics and providing model interpretation.
	Conclusion	Concluding research findings, providing recommendations for future research, and discussing existing limitations.

4. Results and Discussion

4.1. Results

At the Data Selection stage, tweet data about Tiktok Shop on Twitter social media is collected using crawling techniques with the tool used is Tweet Harvest to retrieve data automatically and systematically. Data is retrieved using the keyword “Tiktok Shop” in Indonesian to ensure relevance, the data retrieval time span is January 1, 2023 to July 9, 2024. The raw data obtained is 1907 and includes attributes such as text content (full_text), time (created_at), and other attributes. The results of this raw data collection are stored in CSV (Comma Separated Values) format to facilitate the further analysis process.

1	conversation	created_a	favorite_c	full_text	id_str	image_url	in_reply_t	lang	location	quote_co	reply_cou	retweet_c	tweet_url	user_id_s	username
2	1.81027E+18	Mon Jul 08	0	@chochob	1.81E+18		chochobal	in	i'm unitsta	0	1	0	https://x.	1.62E+18	juiccyhyuck
3	1.81027E+18	Mon Jul 08	0	@juiccyhy	1.81E+18		juiccyhyuc	in	selective.	0	1	0	https://x.	1.64E+18	leehaechan_jpg
4	1.81027E+18	Mon Jul 08	0	kalian ada	1.81E+18	https://pbs.twimg.c		in	i'm unitsta	0	5	0	https://x.	1.62E+18	juiccyhyuck
5	1.81026E+18	Mon Jul 08	1	WTS Avos	1.81E+18	https://pbs.twimg.c		in		0	2	0	https://x.	1.4E+18	asdfghjklzxcYZ
6	1.81025E+18	Mon Jul 08	0	@1996oit	1.81E+18		pinkshade	in		0	1	0	https://x.	1.8E+18	nicotphine
7	1.81023E+18	Mon Jul 08	0	@bluezet	1.81E+18		bluezetea	in	rumah lan	0	0	0	https://x.	1.42E+18	ustadzmalfoy
8	1.81019E+18	Mon Jul 08	0	@staysell	1.81E+18	https://pb	stayselling	in		0	0	0	https://x.	1.25E+09	LindaaKanny
9	1.81023E+18	Mon Jul 08	0	Hiii temer	1.81E+18			in	The Non-j	0	3	0	https://x.	1.52E+18	aoristoteles
10	1.81022E+18	Mon Jul 08	0	guys bikin	1.81E+18			in	MENTION	0	0	0	https://x.	1.34E+18	kwonhorangii

Fig.2: Data Selection

4.1.1. Preprocessing Data

Preprocessing is the beginning of processing raw data so that the data can be used effectively in analytical or modeling efforts. In the context of natural language processing or data analysis, preprocessing aims to transform data that is unstructured, unorganized or contains a lot of noise and the data will be in a more organized and refined format, thus improving the operational efficacy of the algorithm.

Fig. 3: Operators Preprocessing Data

a) Replace : The first operator in the preprocessing process is Replace, which serves to replace certain values or text in the dataset with new values or text. This process is very important to correct errors in the data or change the format of the data to make it more consistent. In this operator, the parameters in the attribute filter type section use single, and for the attribute adjust to the data, for example “full_text”. then the replace what section is filled with RT.
b) Replace URL 1 : This operator is specifically used to replace certain URLs that may be present in the data set. This is necessary to remove irrelevant URLs or to equalize the format of existing URLs, so that the data becomes more consistent and for its parameters in the replace what section include https.*? (space). By performing this replacement, the resulting dataset becomes cleaner and more structured, making it easier to analyze at a later stage.
c) Replace URL 2 : This operator serves to perform any additional URL replacement that may be required. This is an advanced step from the previous process and aims to clean up the data further. In some cases, there may be more than one type of URL that needs to be fixed or customized. After this step, the data set will have cleaner and standardized URLs, thus improving the quality of the data to be used in the analysis. In the parameters section, replace what is filled with https.* without spaces.
d) Replace Hashtag : Removes empty or irrelevant values from the data to clean the dataset, with its parameter in the "replace what" section containing #.*? (Space).
e) Mention : Tags or mentions a specific entity in the data, often used in the context of social media or communication, with its parameter in the "replace what" section containing @.*?.
f) Simbol : Converts data into a symbolic format, often used for visual representation or programming, with its parameter in the "replace what" section containing [|~?.,;":'#*@%4{}-].
g) Trim : Removes unnecessary spaces or characters at the beginning and end of the data to ensure cleanliness and consistency in format.

Row No.	full_text
420	Aku di tiktokshop
421	Foxbox Jam Tangan Pria Anti air Kreatif Besi Tahan karat Shopee Tiktokshop Tokopedia Lazada
422	abis buat mie lidi ala ala tiktokshop
423	Tiktokshop kagak ada freeingkir dah
424	Masalah ni dekat tiktokshop pun sama jeee beria kita ingat dia buat RM1 sale rupa nya yang RM1 tu kotak kosong je
425	Gais yang mau beli kerupuk via tiktokshop boleh dm aku aja yh Aku promosiin tiktokshop punya pacar koko gua 1 partner kok
426	mau kurban beli di tiktokshop
427	aku jugaaa udah punya 4 malah lebih murce di shopee daripada tiktokshop
428	mohon bantuan nya untuk bisa followup masalah saya min Saya berbelanja menggunakan tiktokshop
429	open kredivo spaylatter bca paylatter yup paylatter jenius paylatter gopaylatter tokopediacard ovo u card credinex all cc/kartu kredit traveloka vn/non v...
430	HACK BUAT YG MAU GANTI HP BATERAI SECARA MURAH &amp AMANï¸□ beli batterai nya di shopee/tiktokshop gue pake yg rakkipanda: 128K pa...
431	Pati karo pekalongan iki podo panturane jateng siji viral kampung penadah siji viral isine bos2 garment menguasai pasar tiktokshop
432	Adanya tiktokshop punya pacar abangku kalo mau dm aja yh semisal pengen beli
433	di tiktokshop coba
434	emg sopi jahat bet sm gua njir :(halo tiktokshop

Fig. 4: Data Preprocessing Results

h) Labeling : The process of labeling tweet data, used manually to determine the sentiment category of each tweet that has been collected. Manual labeling is done because the raw data obtained through the crawling process does not have ready-made sentiment label information. the labels used in this study are divided into three categories, namely positive, negative, and neutral. The labeling process is done by reading and understanding each tweet manually, using a spreadsheet application such as Microsoft Excel where each tweet is listed in one column, and one additional column is used to store the sentiment label.

Table.2: Labeling

Full_text	Sentiment
@chochobalsun gatauu knp kalo di tiktokshop tuh ovp.	Neutral
Nanti kalo belanja di tiktokshop gue mending cod aja deh soalnya ini barang gue yg udah dibayar gatau bakalan nyampe atau enggak soalnya gak ada kabar terus. Statusnya diantar terus dari hari apa	Negatif
tiktokshop kenapa murah murah bangett dah nih baju bajunya jadi kalap gua	Positif

Tiktokshop adalah seburuk-buruknya marketplace yang prnah aku coba. Harganya emang jauh di bawah tapi sistemnya jelek dan ga 'jujur'	Negatif
kalimat ini masuk ke sentimen apa ? Platform tiktok jauhhhhh lebih murah jual baju dan tudung. Merujuk yg brand2 tak famous. Kalau dh famous mmg mahal even kt tiktok pun. Dari RM70 boleh turun ke RM35 campur postage. Tambah kalau dpt baucer2. Patutlah indo dulu stop tiktokshop	Positif
guys bikin tiktokshop gimana deh	Neural

4.1.2 Transformation

Transformation is the process of changing data from a particular format or representation to another form that is more suitable for analysis or modeling. This process aims to improve the quality of the data, make it more consistent, or prepare the data to fit the needs of the algorithm to be used.

Fig.5: Operators Transformation

a) Nominal to Text : The nominal to text operator is used to convert nominal type attributes into categories or labels for processing with Natural Language Processing (NLP) techniques.
b) Proses Documents From Data : Processing text by removing stopwords, performing stemming, or other transformations so that the data is ready for further analysis. This operator is used in text processing for sentiment classification. in the documents from data operator there are several other operators such as, Tokenize, Generate N-Gram, Transform Cases, Filter Tokens, Filter Stopwords, Stem.
c) SMOTE (Synthetic Minority Oversampling Technique) Upsampling : Works to address class imbalance in a data set by adding data. Upsampling adds data by using the original samples from the minority class until all classes have the same size. Three uses of SMOTE show that the oversampling process is applied to three pieces of data to improve the class distribution. This can improve the performance of machine learning models used for classification, such as accuracy, precision, and recall.
d) Split Data : This operator is used to split the dataset into two separate subsets: one for training and one for testing. This division is important in machine learning modeling because it trains the model on training data and tests it on testing data to evaluate the model's performance.

4.1.3. Data Mining

Data mining is to explore the value of large, complex, and possibly irregularly structured data. Meanwhile, Naïve Bayes is one of the classification methods in machine learning that relies on Bayes' theorem.

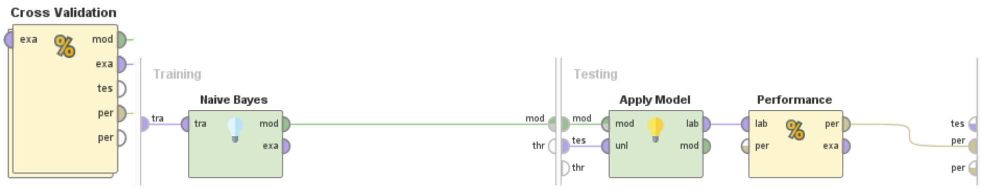

Fig.6: Operator Data Mining

4.1.4 Evaluation

a) Analysis results before applying SMOTE

accuracy: 84.28% +/- 3.14% (micro average: 84.27%)

	true netral	true positif	true negatif	class precision
pred. netral	1107	29	6	96.94%
pred. positif	103	17	2	13.93%
pred. negatif	68	2	1	1.41%
class recall	86.62%	35.42%	11.11%	

Fig. 7: Analysis results before applying SMOTE

In the confusion matrix table above, you can see the results of evaluating the sentiment classification model for three classes: neutral, positive and negative, with a data ratio of 70:30, the accuracy result is 84.28%. The model has the best performance in the neutral class, with precision values reaching 96.94% and recall of 86.62%, indicating that most of the neutral class predictions are correct. In contrast, the model performance for the positive and negative classes is much lower, with precision of 13.93% and 1.41%, respectively, and recall of 35.42% and 11.11%. Data imbalance is the main factor that causes the model to be more accurate in the majority class (neutral) but has difficulty recognizing the minority class (positive and negative).

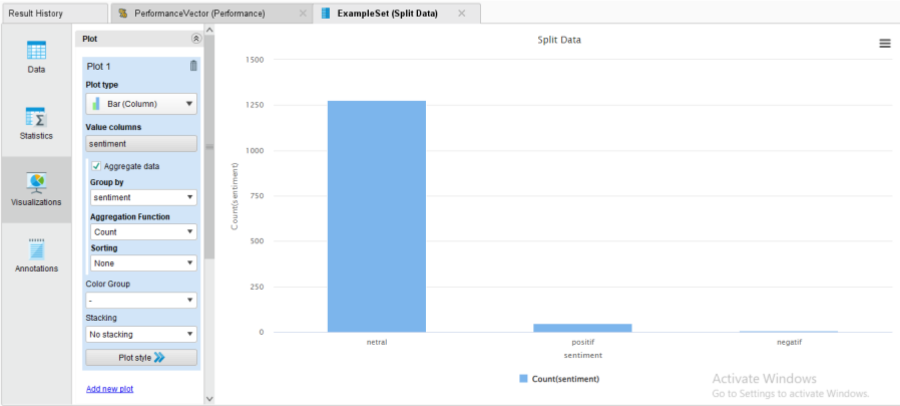

Fig. 8: Plot Results before using SMOTE

In the picture above, before using SMOTE, you can see that there is an imbalance in the amount of data between sentiments, where the neutral category dominates with a much larger amount of data than the positive and negative categories. This imbalance has the potential to reduce the model's performance in detecting positive and negative sentiment. Below is a table of data ratio comparison results before implementing SMOTE.

Table. 3: Results of comparison of ratio data before implementing SMOTE

Data Ratio	Accuracy
80:20	83.29%
70:30	84.28%
60:40	83.22%

From the results of the analysis of the ratio of training data and testing data (80:20, 70:20, and 60:40) in the table above before using the SMOTE (Synthetic Minority Over-sampling Technique) technique, the ratio that produces the greatest accuracy is a ratio of 70:30 with accuracy results of 84.28%.

a) Analysis results after applying SMOTE

The results of training data and testing data of 70:30 produce an accuracy of 93.98% ± 1.39%. Based on the confusion matrix, the precision and recall results for each category are as follows:

accuracy: 93.98% +/- 0.84% (micro average: 93.97%)

	true netral	true positif	true negatif	class precision
pred. netral	1047	0	0	100.00%
pred. positif	154	1278	0	89.25%
pred. negatif	77	0	1278	94.32%
class recall	81.92%	100.00%	100.00%	

Fig. 9: Analysis results after applying SMOTE

The sentiment percentage calculation is carried out using the formula:
Total Data = 1047 (true neutral) + 1278 (true positive) + 1278 (true negative) = 3603. Calculation results for each sentiment:

a. Neutral:

$$\text{Percentage Netral} = \frac{True\ Neutral}{Total\ Data} \times 100\%$$
$$= \frac{1047}{3603} \times 100\% \approx 29.0\%$$

b. Positif:

$$\text{Percentage Positif} = \frac{True\ Positif}{Total\ Data} \times 100\%$$
$$= \frac{1278}{3603} \times 100\% \approx 35.5\%$$

c. Negatif:

$$\text{Percentage Negatif} = \frac{True\ Negatif}{Total\ Data} \times 100\%$$

$$= \frac{1278}{3603} \times 100\% \approx 35.5\%$$

Table. 4: Ratio data comparison results after implementing SMOTE

Rasio Data	Akurasi	Netral	Positif	Negatif
80:20	93.90%	29.0	35.5	35.5
70:30	93.98%	29.0	35.5	35.5
60:40	93.79%	28.9	35.6	35.6

From the results of the analysis of the ratio of training data and testing data (80:20, 70:20, and 60:40) which has been calculated above using the SMOTE (Synthetic Minority Over-sampling Technique) technique, the ratio that produces the greatest accuracy is a ratio of 70: 30 with accuracy results of 93.98 ± 1.39%.

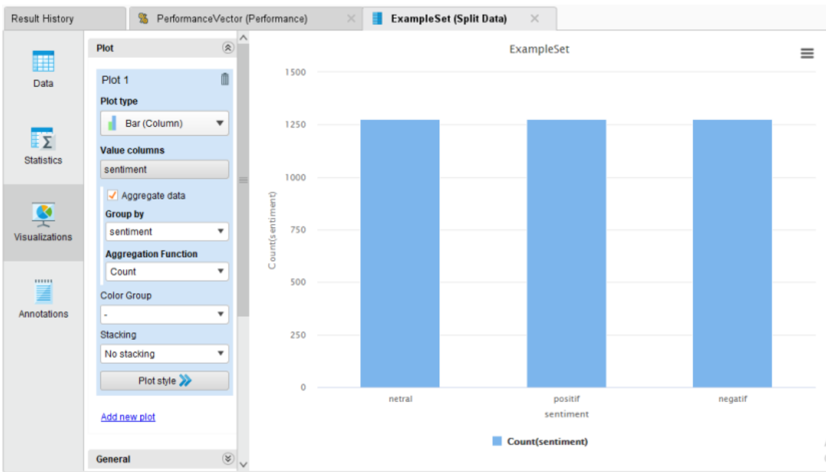

Fig. 10: Plot Results using SMOTE

In the picture above, the data distribution between sentiment categories (neutral, positive and negative) has been balanced after applying the SMOTE (Synthetic Minority Over-sampling Technique) technique. Each category has a similar amount of data, around 1,250 data per category which eliminates the class imbalance that previously existed. This aims to improve the model's performance in recognizing patterns in categories (positive and negative), so that predictions become more accurate.

Table. 5: Accuracy Results Before and after applying SMOTE

Conditions	Accuracy
Before Using SMOTE	84.28 %
After Using SMOTE	93.98%

Before applying the Synthetic Minority Sampling Technique (SMOTE), the model achieved an accuracy of 84.28%, indicating that it classified approximately 84.28% of the examples in the dataset correctly. However, this level of accuracy suggests potential problems with class imbalance, where minority classes are underrepresented, possibly leading to biased predictions. After applying SMOTE, which generates synthetic samples for minority classes to balance the dataset, the model accuracy increased significantly to 93.98%.

4.2. Discussion

```
PerformanceVector:
accuracy: 93.98% +/- 0.84% (micro average: 93.97%)
ConfusionMatrix:
True:   netral  positif negatif
netral: 1047    0       0
positif:        154     1278    0
negatif:        77      0       1278
```

Fig. 11: Result PerformanceVector

Description:
The figure above displays the evaluation results of a machine learning model, specifically in the sentiment classification task. These evaluation results are generally used to measure how well the model can predict the class of data (in this case, sentiment: positive, negative, or neutral). The accuracy value indicates the overall percentage of correct predictions. In this case, the model had an accuracy of 93.98%, meaning it managed to correctly predict the sentiment almost 94% of the time.

5. Conclusions and Suggestions

5.1. Conclusions

Based on the research conducted, it can be concluded that the Naïve Bayes Classifier algorithm has been successfully proven to overcome data imbalance by applying SMOTE, and can effectively classify the sentiment of Twitter users towards TikTok Shop. The analysis results show a fairly balanced sentiment distribution between positive and negative, as well as the model's ability to provide high accuracy, precision, and recall. This research shows that the application of the Naïve Bayes Classifier algorithm combined with the SMOTE oversampling technique successfully improves the accuracy of the classification of Twitter user sentiment towards TikTok Shop by 93.98% with the ratio used is 70:30.

5.2. Sugestions

For future research, it is recommended to improve data quality by expanding the variety of keywords and hashtags, as well as extending the data collection period. This will provide a more comprehensive and accurate picture of user sentiment. In addition, it is necessary to further explore the optimal parameters in the SMOTE method to improve classification performance, especially in minority classes. The use of other algorithms such as SVM, Random Forest, or LSTM can also be considered to compare their performance. Finally, real-time sentiment monitoring can provide more up-to-date insights into the development of users' perceptions of TikTok Shop, thus enabling faster responses to changing trends.

References

[1] N. Al Maeni, N. Suarna, and W. Prihartono, "Analisis Sentimen Pengguna Shopee Berdasarkan Data Tweet Dari Twitter Menggunakan Metode Naive Bayes Classifier," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1834–1840, 2024, doi: 10.36040/jati.v8i2.8298.

[2] E. Andrian and A. Rahman Isnain, "Analisis Sentimen Masyarakat Terhadap Tiktok Shop di Twitter Menggunakan Metode Naive Bayes Classifier," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 2, pp. 788–796, 2024, doi: 10.30865/mib.v8i2.7530.

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

[4] S. Saepudin, S. Widiastuti, and C. Irawan, "Analisis Sentimen Ulasan Platform Media Sosial Menggunakan Algoritma Naïve Bayes Classifier," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 2, pp. 236–243, 2023.

[5] A. M. H. Pardede, "PERANCANGAN SISTEM PAKAR DIAGNOSA PENYAKIT TANAMAN KELAPA SAWIT DENGAN METODE BAYES STUDY KASUS PT.UKINDO BLANKAHAN ESTATE," *Jurnal KAPUTAMA*, 2016, doi: 10.21248/contrib.entomol.13.7-8.894-901.

[6] A. F. Sari, M. Safii, D. Suhendro, and I. S. Damanik, "Implementation Data Mining of Employement Contract Exten-sion at Indosat Using Naïve Bayes," *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 1, no. 1, 2021, doi: 10.59934/jaiea.v1i1.52.

[7] N. A. Alabdulkarim, M. A. Haq, and J. Gyani, "Exploring Sentiment Analysis on Social Media Texts," *Engineering, Technology and Applied Science Research*, vol. 14, no. 3, pp. 14442–14450, 2024, doi: 10.48084/etasr.7238.

[8] M. F. Fakhrezi, Adian Fatchur Rochim, and Dinar Mutiara Kusomo Nugraheni, "Comparison of Sentiment Analysis Methods Based on Accuracy Value Case Study: Twitter Mentions of Academic Article," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 1, pp. 161–167, 2023, doi: 10.29207/resti.v7i1.4767.

[9] P. Putra, A. M. H. Pardede, and S. Syahputra, "Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga," *JTIK (Jurnal Teknik Informatika Kaputama)*, vol. 6, no. 1, pp. 297–305, 2022.

[10] S. Suhada and A. M. H. Pardede, "PENGKLASTERAN DOKUMEN DENGAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR CLUSTERING," *Jurnal Sistem Informasi Kaputama*, vol. 3, no. 2, pp. 1–6, 2019.

[11] D. Aryani, B. Irawan, and A. Bahtiar, "Implementasi Data Mining Pada Data Penjualan Pakaian Menggunakan Algoritma K-Means Dengan Optimize Parameter Grid," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1673–1680, 2024, doi: 10.36040/jati.v8i2.9147.

[12] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.

[13] H. Hajaroh, T. Suprapti, and R. Narasati, "Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Ulasan Produk Makanan Dan Minuman Di Tokopedia," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 111–118, 2024, doi: 10.36040/jati.v8i1.8237.

[14] A. M. H. Pardede, M. Sihombing, and others, "Comparison Of K-Nearest Neighbor And CNN Classification Methods In Diabetic Data Sets," *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 3, no. 1, pp. 128–134, 2023.