Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

Sentiment Analysis of TikTok Shop Prohibition using a Random Forest and Decision Tree

Yudhistira Imam Praja, Kemas Muslim L*

School of Computing Informatics, Telkom University, Bandung, Indonesia
Email: [1]imampraja@student.telkomuniversity.ac.id, [2,*]kemasmuslim@telkomuniversity.ac.id
Correspondence Author Email: kemasmuslim@telkomuniversity.ac.id
Submitted: **04/06/2024**; Accepted: **29/06/2024**; Published**: 30/06/2024**

Abstract−This research explores the impact of the closure of TikTok Shop by the Indonesian government on various aspects of the economy, the e-commerce industry, consumer behavior, and social media dynamics. As an e-commerce platform within the TikTok social media application, TikTok Shop has become a significant business information system that collects, provides, and stores information related to electronic buying and selling activities. Understanding the public's reaction to the closure of TikTok Shop is essential because it can influence consumer confidence, market stability, and future regulatory decisions. Public sentiment provides valuable insights into the potential economic and social consequences, guiding policymakers and businesses in making informed decisions. The closure of this platform has elicited both positive and negative reactions from the public, which are widely expressed through social media, especially Twitter. To analyze public sentiment regarding this issue, two relevant machine learning methods were used: Random Forest and Decision Trees. Random Forest is known for its efficiency in data mining and its ability to handle data imbalance in large datasets. Decision Trees offer similar accuracy and can be applied in both serial and parallel modes, depending on the available data capacity and memory. The results of this study are expected to provide in-depth insights into the implications of the closure of the TikTok Shop and the effectiveness of using machine learning algorithms in social sentiment analysis. This research yielded effective results with a 75.24% accuracy, 80.18% precision, 67.06% recall, and 73.04% F1 score.

Keywords: Random Forest; Decision Tree; TikTok Shop; E-commerce; Sentiment Analysis

1. INTRODUCTION

E-commerce is an abbreviation for electronic commerce, referring to the electronic business segment related to buying and selling transactions of goods and services carried out through an internet network [1]. TikTok Shop is an e-commerce platform that has operated within the context of the social media application TikTok, which has become one of the most popular social media platforms in the world, especially in Indonesia. TikTok Shop can be classified as a business information system because its use for transactions via electronic media enables the collection, provision, and storage of information, especially those related to buying and selling activities in business aspects [2]. However, the closure of the TikTok Shop by the Indonesian government is a topic currently being hotly discussed in the world of e-commerce in Indonesia.

The closure of TikTok Shop in Indonesia has had broad impacts on the economy, e-commerce industry, consumer behavior, and social media dynamics. The government's decision to close the TikTok Shop has elicited both pro and con reactions among the public. One of the social media platforms used to exchange opinions regarding the closure of the TikTok Shop in Indonesia is Twitter. The use of machine learning techniques in data analysis is increasingly becoming a major focus in various fields. In this context, two highly relevant methods are Random Forest and Decision Trees, given their ability to address sentiment issues in society.

Random Forest is an ensemble method that uses decision trees as its base classifiers, which are built and combined. Key aspects of the Random Forest method include bootstrap sampling to create multiple decision trees, where each tree makes predictions using randomly selected predictors. When making predictions, Random Forest combines the results of each decision tree using majority voting for classification and averaging for regression [3]. Random Forest is a highly efficient machine learning algorithm for data mining, especially with large datasets. Its advantage lies in its ability to handle data imbalance among various classes, particularly in extensive datasets. For instance, Rafly Ghazali Ramli conducted research using the Random Forest algorithm with Word2Vec feature expansion, achieving an accuracy rate of approximately 99.49% [4]. Random Forest is a technique in Ensemble Learning that involves eight steps in forming tree structures. When used, Decision Trees are created by taking data samples randomly. To classify data, a Random Forest uses a voting method based on the results produced by various Decisions Trees. Research by Christian Cahyaningtyas demonstrated that the Random Forest algorithm achieved the highest accuracy rate of 99.91% [5],[8].

Decision Trees have a similar level of accuracy as Random Forests. The implementation of Decision Trees can vary between serial and parallel modes, depending on the amount of data, the memory capacity available on the computer, and the algorithm's ability to measure data scalability [6]. Decision trees are data classification methods that produce a model in the form of a tree structure consisting of root, internal, and terminal nodes. Root and internal nodes represent variables or features, while terminal nodes depict class labels. In the classification process, the data to be categorized follows a path from root and internal nodes to terminal nodes. The class label for the query data is determined based on the labels existing in the internal nodes. In conventional Decision Trees, the data used have definite feature values [7]. In Decision Trees, each internal node represents a test on an attribute, where the results of the test are represented by branches emanating from the node, and class labels are stored by each node. In Decision

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

Trees, the top node is called the root node. The root node can be determined by finding the attribute that gives the highest gain or has the lowest entropy value. This usually involves calculating the entropy value using the formula in equation one. Furthermore, the calculation of the gain value is performed using the formula given in equation two [8].

Sentiment analysis is a branch of text research aimed at assessing the views or levels of subjectivity held by society on a topic, event, or issue. This is a classification task that categorizes text into two categories, namely positive or negative orientation[3]. Sentiment analysis involves data that tends to be subjective, such as opinions or opinions that cannot be measured concretely. Moreover, these data comes from individuals, and each individual has preferences and diverse ways of expressing their opinions [9].

Based on this analysis, the research compared the results of accuracy using the Random Forest and Decision tree algorithms. Using sentiment analysis, this research calculates the accuracy of both algorithms and classify various forms of sentiment related to the TikTok Shop closure event. Using a dataset from Twitter, the data used is in the Indonesian language and consists of 6000 data points, which are then divided into two classes, namely positive sentiment and negative sentiment.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The research was conducted based on the flowchart image that has been created, which can be seen in Figure 1.

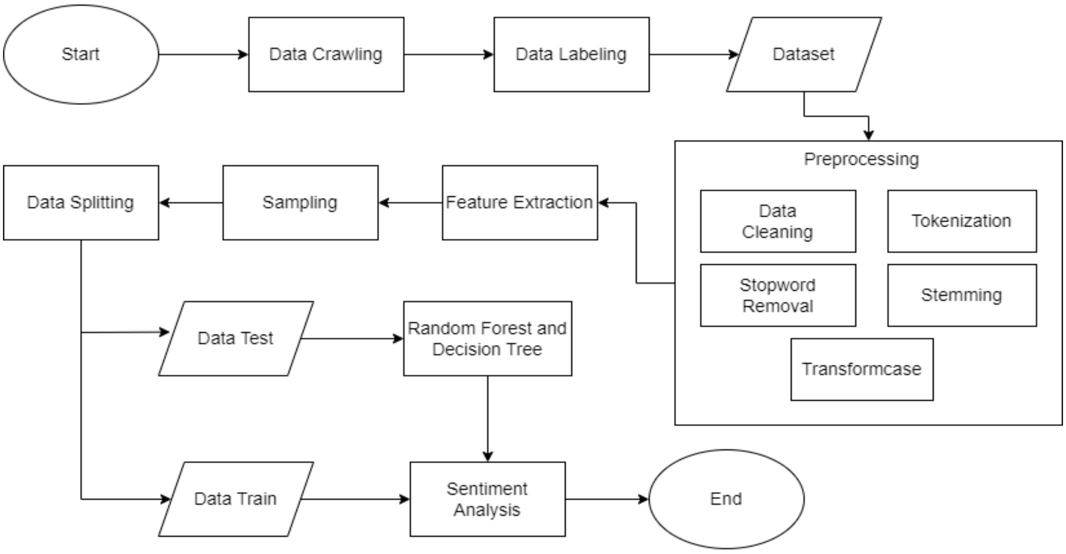

Figure 1. Flowchart

Based on the flowchart in figure 1, the application of sentiment analysis begins with data collection, which is achieved by performing a data crawling process on Twitter. Once the data are collected, data preprocessing is performed, which includes data cleaning, case transformation, tolerization, stopwords removal, and stemming. After the preprocessing stage, the next step is feature extraction. This step involves identifying the aspects of comments and predicting sentiment to determine whether the text contains sentiment or opinion. To prevent class imbalance, data sampling is performed. The data are then split into two categories: training data and test data. The next step is to implement the random forest and decision tree algorithms. This process involves calculating the accuracy, precision, recall, and F1-Score values to evaluate the performance of the algorithms.

2.2 Data Crawling

This study utilizes a dataset collected from Twitter using a crawling technique. As a popular social media platform, Twitter offers data widely employed for various analyses. The dataset comprises tweets expressing public opinions and views on the closure of the Indonesian TikTok Shop, all written in Indonesian. It includes 6000 entries with attributes such as usernames and hashtags.

2.3 Data Labeling

Data labeling entails categorizing text by assigning specific labels for the purpose of model evaluation. This process also helps classify the sentiment expressed in a sentence. Data is categorized into three types: positive, neutral, and negative. The results of the labeling process are detailed in Table 1.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

Table 1. Data Labeling

Data Review and Feature Extraction	Label
Tiktok shop ditutup ini akan banyak bawa korban: - staff Tiktok yg ngurusi shop ini - bisnis eCommerce enabler seperti Sirclo - UMKM yg berinovasi dgn live shopping - agency yg menawarkan jasa live streaming - KOL yg endorse brand/product sbg affiliate Teten/Zulhas paham gak? (The closure of the TikTok shop indeed affects various parties, including TikTok staff, eCommerce enabler businesses like UMKM that innovate with live shopping, agencies providing live streaming services, and KOL who endorse brands or products as affiliates. However, whether Teten or Zulhas understand the implications remains?)	Positive
Bukan soal drama nya, pasti ada sedihnya dri tiktok shop ditutup. Orang yg biasa jdi host live jdi bingung bres ini gmna, trus karyawan yg biasa nerima orderan beberapa e-commerce ilang ini satu pasti ada efek nya juga (It's not about the drama, there is definitely sadness from the TikTok shop being closed. People who usually host live are confused about what this is all about, then employees who usually take orders for several e-commerce sites have disappeared, this is definitely an effect too)	Negative
Dulu orderan dr tiktokshop bisa sampe 500 perhari, sekarang cuman 100 an aja. Kenapa ngga malah ngadain workshop umkm aja biar bisa ikut jualan di tiktok shop? Gimana pak? (Previously orders from TikTok could reach 500 per day, now it's only 100 or so. Why not just hold a workshop for umkm so they can sell on the TikTok shop? How are you, sir?)	Neutral

Based on the table one provided above, it illustrates a labeled dataset containing sentiments categorized as positive, negative, and neutral.

2.4 Dataset

The dataset for this research was gathered from Twitter using the data crawling technique. It consists of Indonesian language tweets that express the opinions of Indonesians about the closure of the TikTok Shop. The dataset includes 6000 tweets, with the selected data attributes being Tweet, Polarity, and Label.

2.5 Preprocessing

Data preprocessing involves several steps. This preprocessing includes data preparation to ensure that the data are processed properly during the modeling process. Preprocessing activities involve data refinement and removal of irrelevant data so that it can be directed to the next steps. The following are the steps in data preprocessing:

a. Data Cleaning

Data Cleaning is the first step in data processing. At this stage, errors, inaccuracies, and inconsistencies in the data are corrected or removed. The goal is to ensure that the data conform to the established modeling design. The data cleaning process also includes handling data by eliminating duplicates and improving the format or structure of the data to ensure its accuracy, so that the data are ready to be processed in the next stage [10] .

Table 2. Data Cleaning

Actual Data	Data Cleaning
Mana boleh, wong salah satu problem kenapa TikTok Shop di tutup kemarin tu ya karena barang2 luar negeri macam shopee. Bahayanya buat UMKM dalam negeri (OK guys, one of the problems why the TikTok Shop closed yesterday was because of foreign goods such as Shopee. The danger is for domestic UMKM)	Mana boleh wong salah satu problem kenapa TikTok Shop di tutup kemarin tu ya karena barang luar negeri macam shopee Bahayanya buat UMKM dalam negeri (One of the problems why TikTok Shop closed yesterday was because of foreign goods such as Shopee. The danger is for domestic UMKM)

Based on the table two provided above, In the data cleaning stage, the dataset undergoes a process to eliminate anomalies and potential errors. This includes identifying and addressing missing values, outliers, duplicates, and other data entry errors.

b. Tokenization

Tokenization is the process of extracting all the words from each document and removing characters such as symbols, numbers, punctuation marks, and anything that is not letters[11] . Then, the words are separated from these characters to form meaningful words that will be analyzed.

Table 3. Process Tokenization

Data Cleaning	Tokenization
Mana boleh wong salah satu problem kenapa TikTok Shop di tutup kemarin tu ya karena barang luar negeri macam shopee Bahayanya buat UMKM dalam negeri	[‘Mana’,’boleh’,’wong’,’salah’,’satu’,’problem’, ’kenapa’,’Tik’,’Tok’.’Shop’,’di’,’tutup’,’kemarin’, ’itu’,‘ya’,’karena’,’barang’,’luar’,’negeri’,’macam’,

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

(One of the problems why the TikTok Shop was closed yesterday was because of foreign goods like Shopee The danger is for domestic UMKM)	'shopee','Bahayanya','buat','UMKM','dalam', 'negeri]

On the basis of Table three, the sentences are broken down into individual words. This tolerization process is performed using the NLTK library, specifically the `nltk.tokenize` module.

c. Stopword Removal

Stopword Removal: At this stage, words that fall into the stopwords category are eliminated. Stopwords refer to words that frequently appear but are considered meaningless [10]. The importance of removing stopwords is to avoid words that are considered less relevant in the classification model.

Table 4. Stopword Removal

Tokenization	Stopword Removal
['Mana','boleh','wong','salah','satu','problem', 'kenapa','Tik','Tok'.'Shop','di','tutup','kemarin','itu' ,'ya','karena','barang','luar','negeri','macam','shopee ','Bahayanya','buat','UMKM','dalam','negeri]	['Mana','boleh','~~wong~~','salah','satu','~~problem~~', 'kenapa','~~Tik~~','~~Tok~~'.'~~Shop~~','~~di~~','tutup','kemarin', '~~itu~~','ya','karena','barang','luar','negeri','~~macam~~', '~~shopee~~','Bahaya','buat','UMKM','dalam','negeri]

Based on Table Four above, during the stopword removal stage, common words that typically do not significantly contribute to the meaning or context of a text are eliminated.

d. Stemming

At this stage, the base word is identified by removing all affixes attached to the word [10]. The stemming process allows us to treat nouns, verbs, and adverbs that have the same root in a uniform manner.

Table 5. Process Stemming

Stopword Removal	Stemming
['Mana','boleh','salah','satu','kenapa', 'tutup','kemarin','ya','karena','barang', 'luar','negeri','Bahayanya','buat','UMKM', 'dalam','negeri]	['Mana','boleh','salah','satu','kenapa', 'tutup','kemarin','ya','karena','barang', 'luar','negeri',**'Bahaya'**,'buat',**'UMKM'**, 'dalam','negeri]

Table five shows the stemming process in which words are changed back to their basic forms. For example, the word "happen" is stemmed to its base form "so." This stemming process uses a stemmer from the Python Sastrawi library with the `StemmerFactory` method.

e. Transform Case

Transform Case is used to change uppercase letters to lowercase in the text. The aim is to ensure the consistency of letters during the classification process and to avoid errors during tolerization [3].

Table 6. Transform Case

Stemming	Transform Case
['Mana','boleh','salah','satu','kenapa', 'tutup','kemarin','ya','karena','barang', 'luar','negeri','Bahaya','buat','UMKM', 'dalam','negeri]	[**'mana'**,'boleh','salah','satu','kenapa', 'tutup','kemarin','ya','karena','barang', 'luar','negeri',**'bahaya'**,'buat',**'umkm'**, 'dalam','negeri]

In Table Six, during the transform case stage, all text in the dataset is standardized to a consistent format, either uppercase or lowercase.

2.6 Feature Extraction

In this research, feature extraction is used to analyze comments, predict sentiment by identifying text that contains opinions, and determine the sentiment's polarity positive, negative, or neutral. The output of feature extraction is a set of features derived from the input text in various forms .[12] Count Vectorizer, a technique in natural language processing (NLP), is employed to transform a collection of text documents into a feature matrix, where each element represents the frequency of word occurrences in the document.

2.7 Sampling

This research carried out data sampling to prevent class imbalance. The data sampling technique uses the SMOTE algorithm and Random Undersampling . The SMOTE technique performs heuristic oversampling, generating artificial samples from minority classes by interpolating surrounding instances. The Random Undersampling technique removes instances from the majority class to compress the class imbalance ratio[13].

2.8 Data Splitting

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

Data Splitting involves dividing the dataset into two parts: training data and test data. The training data are used to develop the model, while the test data are used to evaluate its performance. This evaluation measures the model's accuracy in making predictions.[14].

2.9 Random Forest

Random Forest, an ensemble method, employs decision trees as its base classifiers, which are built and then aggregated. Certain critical aspects of the Random Forest technique involve employing bootstrap sampling to construct prediction trees, where each decision tree makes predictions using a random subset of predictors. This methodology enhances the model's performance by reducing variance and guarding against overfitting. In a study led by Schenbly and his research team, they focused on using a classification system based on the Random Forest method to detect bots and misinformation. Random Forest was selected because of its resilience and capability to mitigate overfitting. When a Random Forest generates predictions, the outcomes from each decision tree are combined through majority voting for classification and averaging for regression.[4], [3].

The process of building a Random Forest involves creating bootstrap samples by randomly selecting data with a replacement for each decision tree. Approximately one-third of the data samples are not used in this process, and these unused samples are referred to as Out-Of-Bag (OOB) data. Decision trees in the forest have their own OOB data, which are used to estimate individual error rates, known as OOB error estimation. Random Forest also has the capability to calculate importance levels and variable estimates, which are useful for identifying variables that significantly contribute to prediction results [15].

2.10 Decision Tree

A Decision Tree generates a model structured like a tree, comprising elements such as root, internal, internal nodes and terminal nodes. Root, nodes and internal nodes depict variables or features, whereas terminal nodes signify class labels. During classification, the data to be categorized traverses the path from the root and internal nodes until it reaches the terminal node. Class labels for query data are determined on the basis of the labels found in the internal nodes. Traditional Decision Trees typically use data with specified feature values[16]. Through attribute selection, Decision Tree, another recursive tree induction approach, improves accuracy by eliminating noisy branches from the data [17].

In a Decision Tree, the top node is called the root node. The root node can be determined by finding the attribute that provides the highest gain or has the lowest entropy value. This usually involves calculating the entropy value using the formula in equation one. Next, the gain value is calculated using the formula given in equation two [8]. Decision Trees can be used to understand the relationship between variables in a dataset. They can find hidden correlations between input and target variables, even if the correlation is not directly visible[13]. The formula for equation two is as follows:

$$Entropy\ (S)^n = \sum_{i=0}^{n} - pi * log^2 \text{pi} \quad (1)$$

$$Gain\ (S, A) = Entropy\ (S) = \sum_{i}^{n} \frac{|si|}{|s|} * Entropy\ (Si) \quad (2)$$

2.11 Evalution

Evaluation is carried out to gauge the effectiveness of the constructed model, which is a critical step in assessing the efficacy of sentiment analysis. System evaluation entails employing a confusion matrix that furnishes a comprehensive depiction of the model's performance with the tested data [12]. This matrix typically includes classes such as

a. True Positive (TP) refers to positive data that is correctly predicted to be positive.
b. True Negative (TN) refers to negative data that is correctly predicted to be negative.
c. False Positive (FP) refers to negative data that is incorrectly predicted to be positive.
d. False Negative (FN) refers to positive data that is incorrectly predicted to be negative.

Values such as accuracy, precision, recall, and F1-Score are calculated on the basis of the confusion matrix, providing the final assessment of the model's performance[13].

a. Accuracy

A statistical measure known as accuracy is employed to assess the effectiveness of predictions in classifying a document as either positive or negative. The accuracy value provides insights into the overall accuracy of data categorization by indicating the proportion of correctly identified examples to the total number of instances [17].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

b. Precision

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

Precision measures the model's accuracy in predicting the correct class overall. It indicates how often the model accurately predicts positive classes. High precision means the model makes few false positive errors, reliably identifying positive instances[18].

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

c. Recall

Recall, also known as sensitivity or the true positive rate, represents the proportion of correctly predicted positive instances compared to the total actual positive data. It assesses the model's capability to identify positive cases within all existing positive instances [19].

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

d. F1-Score

The F1-score is a metric that combines precision and recall, providing a balanced measure of a model's performance. It is often used as an alternative to accuracy, especially when there is a high cost associated with false negatives. The F1-score is calculated as the harmonic mean of precision and recall, which means it gives equal weight to both metrics. This balance helps to mitigate the trade-off between precision and recall, making it a more robust measure of model performance [20].

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision \times Recall} \quad (6)$$

3. RESULT AND DISCUSSION

3.1 Data

This study uses a dataset collected from Twitter using web crawling techniques. Twitter, a prominent social media platform, offers widely accessible data. The dataset comprises tweets expressing public opinions and perspectives on the matter of closing the Indonesian TikTok Shop, with all content in Indonesian. It encompasses 6000 entries and features attributes such as usernames and hashtags. The contents of these Tweets were classified into two sentiment categories: positive and negative.

Table 7. Data Label

Sentiment	All Data
Positive	3644
Negative	1926

Table 6 shows 3644 reviews labeled positivee and 1926 reviews labeled negative. This stage has gone through two labeling phases, done automatically with a few lines of Python code.

3.2 Testing Results and Scenarios

The research conducted an evaluation following the modeling of the Random Forest and Decision Tree algorithms, comparing six proportions of data splitting, with a ratio of 80:20. Data splitting involves dividing the dataset into two subsets. The first subset is used for training the model (Train Data), and the second subset is employed to test or evaluate the performance of the trained model (Test Data). The following data splitting, three data sampling techniques were applied:

a. Baseline: Refers to the original dataset and is used for the analysis process or model training without any modifications to address class imbalance.
b. Oversampling: This method addresses class imbalance by augmenting the number of samples from the minority class. It can involve duplicating existing samples or generating artificial samples to achieve a balanced representation of classes within the dataset.
c. Undersampling: This technique is also used to address class imbalance but in the opposite manner. It reduces the number of samples from the majority class to balance it with the minority class.

Data splitting plays a pivotal role in assessing performance and ensuring the generalization of the classification model. The outcomes of this process involve assessing accuracy, precision, recall, and the F1-score, with the goal of attaining optimal results in the classification model. The results of the testing and result are in table 8.

Table 8. Results from the Data Splitting

Model	Sampling Type	Accuracy	Precision	Recall	F1-Score
Random Forest	Baseline	74.01%	78.98%	82.10%	80.51%
	Oversampling	75.24%	80.18%	67.06%	73.04%

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

	Undersampling	71.85%	72.62%	70.06%	71.32%
Decision Tree	Baseline	68.84%	76.94%	74.75%	75.83%
	Oversampling	70.91%	75.98%	61.14%	67.76%
	Undersampling	65.93%	65.39%	67.56%	66.46%

The test results presented in the table eight show that an 80:20 data split ratio yields the best performance. Specifically, the Random Forest algorithm (Oversampling) achieved the highest scores across all metrics: an accuracy of 75.24%, precision of 80.18%, recall of 67.06%, and an F1 score of 73.04%. These findings indicate that the Random Forest algorithm is highly effective in accurately predicting sentiment classification while maintaining balanced performance across all evaluation metrics. The consistently high scores affirm the robustness of the Random Forest model when applied to a specific dataset.

The success of the Random Forest (Oversampling) model can be attributed to several factors. Random Forest is an ensemble method that constructs multiple decision trees and combines their predictions, reducing the risk of overfitting and enhancing the model's generalization capability. During training, Random Forest creates a large number of decision trees using bootstrapped samples of the data, and each split considers a random subset of features. This approach generates a diverse set of decision trees, and their combined predictions result in superior performance compared to individual decision trees. The ensemble method works by aggregating the mode of the classes (for classification) or the average prediction (for regression) from each tree, inherently enhancing the model's stability and accuracy. The combination of Random Forest and oversampling helps the model achieve high precision scores, indicating fewer false positives, and high recall scores, indicating fewer false negatives, which balance into a strong F1 score. High precision means that when the model predicts a positive sentiment, it is most likely correct, while high recall means that the model successfully identifies most of the actual positive sentiments. The F1 score, which is the harmonic mean of precision and recall, provides a single metric that balances both concerns, and in this case, the score of 73.04% indicates balanced performance.

Conversely, the lowest scores were obtained with the Decision Tree algorithm (using undersampling), which achieved an accuracy of 65.93%, precision of 65.39%, recall of 67.56%, and an F1 score of 66.46%. This lower performance highlights the challenges and limitations associated with the Decision Tree algorithm in this context, especially when undersampling techniques are applied. The undersampling method aims to balance the dataset by reducing the number of examples from the majority class. While this can help reduce bias toward the majority class, it also results in the loss of valuable information, thus reducing the overall amount of data available for training. This reduction in data can lead to a less effective model, as seen with the Decision Tree algorithm. Undersampling also affects the structure of the decision tree itself. Decision Trees split the data at various points to create nodes, with each split designed to maximize class separation based on criteria such as Gini impurity or information gain. When the dataset is undersampled, the reduced number of samples can lead to splits that are not as representative or optimal. The tree may become shallower or have fewer branches, limiting its ability to capture the complexity of the data.

3.3 Test Result

Testing in this research using Random Forest and Decision Tree algorithms requires Data Splitting. By performing Data Splitting, you obtain three Data Sampling, namely Baseline, Oversampling, and Undersampling, to observe differences in the model. This research yielded quite effective results in terms of accuracy, precision, recall, and F1-score with the data proportion being 80:20. Here, 80% of the testing Train Data is used to study the data and features in the dataset, while 20% of the test data is allocated for Random Forest and Decision Tree algorithm testing.

Results obtained from the Baseline using Random Forest and Decision Tree algorithms show scores of 75.24% accuracy, 80.18% precision, 67.06% recall, and 73.04% F1 score. The Confusion Matrix is a table of test data that predicts positive and negative classification by models. This method is used for multiple classifiers. The results of the Confusion Matrix are in table 9.

Table 9. Confusion Matrix: Random Forest

		Confusion Matrix: Random Forest (Oversampling)	
		Negative	Positive
Predicted Value	Negative	760	151
	Positive	300	611

The results from Table Nine, which presents the Confusion Matrix for Random Forest (Oversampling) in identifying public sentiment toward the closure of TikTok Shop with positive and negative labels, are as follows:

a. True Positive (TP) = 611: The model correctly predicts actual and true positive sentiments .
b. True Negative (TN) = 760: The model correctly predicts actual and true negative sentiments.
c. False Positive (FP) = 151: The model incorrectly predicts positive sentiments when the actual sentiment is negative.
d. False Negative (FN) = 300: The model incorrectly predicts negative sentiments when the actual sentiment is positive.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

The results obtained from the Decision Tree (Undersampling) with the lowest score show an accuracy score of 65.93%, precision of 65.39%, recall of 67.56%, and F1 score of 66.46%.

Table 10. Confusion Matrix: Decision Tree (Undersampling)

		Confusion Matrix: Decision Tree (Undersampling)	
		Negative	Positive
Predicted Value	Negative	310	172
	Positive	156	325

Table Ten displays the Confusion Matrix results for Decision Tree (Undersampling) in identifying public sentiment toward the closure of the TikTok Shop with positive and negative labels:

a. True Positive (TP) = 325: The model correctly predicts actual and true positive sentiments .
b. True Negative (TN) = 310: The model correctly predicts actual and true negative sentiments.
c. False Positive (FP) = 172: The model incorrectly predicts positive sentiments when the actual sentiment is negative.
d. False Negative (FN) = 156: The model incorrectly predicts negative sentiments when the actual sentiment is positive.

4. CONCLUSION

In this study, sentiment analysis was conducted using both Random Forest and Decision Tree algorithms to investigate the closure of the TikTok Shop, with a particular emphasis on comparing their performance. The findings revealed that the Random Forest (Oversampling) algorithm achieved the highest scores, boasting an accuracy of 75.24%, precision of 80.18%, recall of 67.06%, and F1 score of 73.04%. Following closely, the Decision Tree (Undersampling) algorithm yielded the second highest results, with respective accuracy rates of 65.93%, precision of 65.39%, recall of 67.56%, and F1 score of 66.46%. Consequently, it can be inferred that the Random Forest algorithm, employing Data Splitting, Oversampling, or Undersampling techniques, outperforms other approaches, particularly excelling in Oversampling. These outcomes underscore the Random Forest's robustness, capable of handling fundamental datasets without the necessity of additional sampling methodologies. The Random Forest algorithm's superior performance stems from its adeptness in managing multiple features and its ensemble nature, mitigating the risk of overfitting and enhancing generalization capabilities. Conversely, Decision Tree algorithms, even with Undersampling, demonstrate inefficacy, shedding light on the constraints of single tree models in dealing with imbalanced data. This study underscores the efficacy of ensemble methods like Random Forest over individual models such as Decision Trees for sentiment analysis tasks, especially those involving intricate datasets. Future research avenues could explore the integration of ensemble techniques or hybrid models to further enhance the accuracy and dependability of sentiment analysis outcomes.

REFERENCES

[1] S. C. Nandaresta and C. Warman, “A Analisis Sentimen Tanggapan Masyarakat Terhadap TikTok Shop Dan Shopee Di Twitter Menggunakan Metode Naive Bayes Dan KNN (K-Nearest Neighbor): Indonesia,” in Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra, 2023, pp. 119–127.
[2] A. N. Sa’adah, A. Rosma, and D. Aulia, “Persepsi generasi Z terhadap fitur Tiktok Shop pada aplikasi Tiktok,” Transekonomika: Akuntansi, Bisnis Dan Keuangan, vol. 2, no. 5, pp. 131–140, 2022.
[3] F. F. Rachman and S. Pramana, “Analisis sentimen pro dan kontra masyarakat Indonesia tentang vaksin COVID-19 pada media sosial Twitter,” Indonesian of Health Information Management Journal (INOHIM), vol. 8, no. 2, pp. 100–109, 2020.
[4] R. G. Ramli and Y. Sibaroni, “Klasifikasi topik twitter menggunakan metode random forest dan fitur ekspansi word2vec,” eProceedings of Engineering, vol. 9, no. 1, 2022.
[5] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, “Random forests and decision trees,” International Journal of Computer Science Issues (IJCSI), vol. 9, no. 5, p. 272, 2012.
[6] A. Priyam, G. R. Abhijeeta, A. Rathee, and S. Srivastava, “Comparative analysis of decision tree classification algorithms,” International Journal of current engineering and technology, vol. 3, no. 2, pp. 334–337, 2013.
[7] N. T. Romadloni, I. Santoso, and S. Budilaksono, “Perbandingan Metode Naive Bayes, Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl Commuter Line,” IKRA-ITH Informatika: Jurnal Komputer dan Informatika, vol. 3, no. 2, pp. 1–9, 2019.
[8] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiasari, “Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE,” AITI, vol. 18, no. 2, pp. 173–184, 2021.
[9] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. S. T. Ayu, and W. F. Dicka, “Dataset Indonesia untuk Analisis Sentimen,” Jurnal Nasional Teknik Elektro dan Teknologi Informasi, vol. 8, no. 4, pp. 334–339, 2019.
[10] M. A. A. Jihad, A. Adiwijaya, and W. Astut, “Analisis sentimen terhadap ulasan film menggunakan algoritma random forest,” eProceedings of Engineering, vol. 8, no. 5, 2021.
[11] T. W. Putra, A. Triayudi, and A. Andrianingsih, “Analisis Sentimen Pembelajaran Daring menggunakan Metode Na\"\ive Bayes, KNN, dan Decision Tree,” Jurnal Teknologi Informasi dan Komunikasi, vol. 6, no. 1, p. 2022, 2022.
[12] M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, “A review of feature extraction in sentiment analysis,” Journal of Basic and Applied Scientific Research, vol. 4, no. 3, pp. 181–186, 2014.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 378-386
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5285

[13] D. N. I. Huda, C. Prianto, and R. M. Awangga, “Analisis Sentimen Perbandingan Layanan Jasa Pengiriman Kurir Pada Ulasan Play Store Menggunakan Metode Decision Tree Dan Random Forest,” JURNAL ILMIAH INFORMATIKA, vol. 11, no. 02, pp. 150–158, 2023.
[14] B. Vrigazova, “The proportion for splitting data into training and test set for the bootstrap in classification problems,” Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy, vol. 12, no. 1, pp. 228–242, 2021.
[15] L. Ratnawati and D. R. Sulistyaningrum, “Penerapan random forest untuk mengukur tingkat keparahan penyakit pada daun apel,” Jurnal Sains dan Seni ITS, vol. 8, no. 2, pp. A71–A77, 2020.
[16] E. Fitri, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” Jurnal Transformatika, vol. 18, no. 1, pp. 71–80, 2020.
[17] K. A. Rokhman, B. Berlilana, and P. Arsi, “Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online,” Journal of Information System Management (JOISM), vol. 2, no. 2, pp. 1–7, 2021.
[18] V. A. Permadi, “Analisis sentimen menggunakan algoritma Na\"\ive Bayes terhadap review restoran di Singapura,” Jurnal Buana Informatika, vol. 11, no. 2, pp. 141–151, 2020.
[19] T. F. Basar, D. E. Ratnawati, and I. Arwani, “Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan Shopeepay dengan Algoritma Random Forest,” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 6, no. 3, pp. 1426–1433, 2022.
[20] M. Hudha, E. Supriyati, and T. Listyorini, “Analisis Sentimen Pengguna YouTube Terhadap Tayangan #MataNajwaMenantiTerawan Dengan Metode Naïve Bayes Classifier,” JIKO (Jurnal Informatika dan Komputer), vol. 5, no. 1, pp. 1–6, 2022.

