Jurnal Galaksi (Global Knowledge, Artificial Inteligent and Information System) E-ISSN: 3048-2399
Vol 1, No 2, August 2024 , pp.70-80
https://doi.org/10.70103/galaksi.v1i2.15

Sentiment Analysis of YouTube Comments on the Closure of TikTok Shop Using Naïve Bayes and Decision Tree Method Comparison

Putu Pebri Armaeni[1], I Komang Arya Ganda Wiguna[2*], Wayan Gede Suka Parwita[3]

[1,2*]Informatika, Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia
[3] Politeknik Negeri Bali, Denpasar, Indonesia
[1] pebriarmaeni@gmail.com , [2*] kmaryagw@instiki.ac.id , [3] gede.suka@gmail.com
* Corresponding Author

ARTICLE INFO

Article history:
Received 04 August 2024
Revised 29 August 2024
Accepted 30 August 2024
Available Online 31 August 2024

***Keywords**:*
Closure TikTok Shop;
Decision Tree;
Naïve Bayes;
Python;
Sentiment Analysis;

ABSTRACT

As technology advances, YouTube has become a social media platform that allows users to watch, broadcast, and share videos. One of the videos that has garnered a lot of comments from the public is about the closure of TikTok Shop. This research uses two methods: Decision Tree and Naïve Bayes. The aim of this study is to compare the Naïve Bayes and Decision Tree methods in analyzing public sentiment regarding the closure of TikTok Shop. The test results for both methods are not significantly different. Each method is divided into three research scenarios. In Scenario 1, with an 80:20 data split, the Decision Tree method achieved an accuracy of 74.71%, a precision of 57%, a recall of 57%, and an F1-score of 57%, while Naïve Bayes had an accuracy of 73.96%, a precision of 58%, a recall of 34%, and an F1-score of 29%. In Scenario 2, with a 70:30 data split, the Decision Tree method achieved an accuracy of 73.27%, while Naïve Bayes achieved an accuracy of 73.99%. In Scenario 3, with a 60:40 data split, the Decision Tree method achieved an accuracy of 71.78%, while Naïve Bayes achieved an accuracy of 74.02%. The evaluation results indicate that the Decision Tree method using an 80:20 data split has superior accuracy compared to the Naïve Bayes method.

Copyright © 2023 Galaksi Journal.All rights reserved.
is Licensed under a Creative Commons Atribusi-NonKomersial 4.0 Internasional Lisensi (CC BY-NC 4.0)

1. Introduction

Along with technological advancements, social media has become a means for people to communicate with each othser in various ways, one of which is through YouTube. YouTube is a social media platform that allows its users to watch, broadcast, and share videos. YouTube features advertising tools with like and dislike options. YouTube users who watch videos on YouTube channels can express their opinions about the videos in the comment section (Deori et al., 2023). The information on YouTube is mostly in the form of opinions and written statements in unstructured text (Jelodar et al., 2021; Kurniawan et al., 2024).

To see public response to the topic of TikTok Shop closure, various media can be used, one of which is social media YouTube. The closure of TikTok Shop has become a trending topic on YouTube. One of the videos that received many comments from the public is about the closure of TikTok Shop. TikTok decided to close the buying and selling service on its platform, TikTok Shop, starting Wednesday, October 4, 2023, at 5:00 PM WIB. This began with Minister of Trade Zulkifli Hasan's statement that social commerce platforms may only promote goods or services but are prohibited from facilitating transactions or sales for users. The Ministry of Trade revised Minister of Trade Regulation (Permendag) Number 50 of 2020 with Permendag Number 31 of 2023 concerning Business Licensing, Advertising, Development, and Supervision of Business Actors in Electronic Trading Systems (Puwaningwulan et al., 2024). Social media like TikTok should function according to their permits. The impact of TikTok Shop has led to a decline in sales and production within the scope of micro, small, and medium enterprises (MSMEs) and conventional markets, affecting MSMEs (Rismauli & Suherman, 2024). This issue has not only appeared on television but has also been widely discussed on social media like YouTube. In a short time, hundreds or even thousands of public responses filled the comment sections on CNN Indonesia's video upload with 4.1K likes and 5.1K comments, tvOneNews with 1.7K likes and 1.7K comments, CNBC Indonesia with 327 likes and 719 comments, and dr. Richard Lee, MARS's podcast with 3.9K likes and 1.7K comments regarding the closure of TikTok Shop. This research can contribute to the development of knowledge, particularly in the field of sentiment analysis and natural language processing. By comparing the Naïve Bayes and Decision Tree methods (Anwar et al., 2022; Arya et al., 2022; Fitri et al., 2019), this research will provide a deeper understanding of the effectiveness of each method in the context of the TikTok Shop closure. The results of this study will enrich our knowledge of sentiment analysis on social media data and offer new insights for future research.

2. Literature Review

Sentiment analysis of YouTube comments, particularly in the context of the closure of TikTok Shop, can be effectively approached using machine learning methods such as Naïve Bayes and Decision Trees. These methods are widely recognized for their applicability in text classification tasks, including sentiment analysis, where the goal is to classify comments as positive, negative, or neutral based on the expressed sentiments.

The Naïve Bayes classifier has been shown to perform well in sentiment analysis tasks due to its simplicity and efficiency. For instance, demonstrated that Naïve Bayes achieved an accuracy of 63.21% in their sentiment analysis of YouTube comments, outperforming other methods like K-Nearest Neighbors (KNN), which only reached 58.10% accuracy (Udayana et al., 2023; Wiratama et al., 2022). This suggests that Naïve Bayes is a robust choice for analyzing sentiments in user-generated content on platforms like YouTube.

In contrast, Decision Trees offer a different approach by creating a model that predicts the value of a target variable based on several input variables. They are particularly useful for their interpretability and ability to handle both numerical and categorical data. The hybrid approach combining Decision Trees with other machine learning techniques has been explored in various studies, indicating that ensemble methods can enhance sentiment classification performance (Aribowo et al., 2020). For example, the use of tree-based ensemble methods has been shown to improve sentiment analysis accuracy on YouTube comments significantly.

Moreover, the effectiveness of sentiment analysis on YouTube comments is supported by various studies that highlight the emotional content shared by users. emphasized the importance of sentiment analysis in validating the relationship between comment sentiment and user engagement metrics, such as subscriber growth (Danda, 2021). This indicates that understanding the sentiment behind comments can provide insights into user behavior and preferences, which is crucial for platforms like YouTube.

Additionally, the emotional complexity of comments on YouTube is noteworthy. Research by revealed that users often express mixed sentiments, which complicates the analysis (Bozkurt & Aras, 2020). This aligns with findings from other studies that suggest a need for sophisticated sentiment analysis techniques capable of discerning nuanced emotional expressions within comments (Naz et al., 2023; Septiani et al., 2024). The use of hybrid models that incorporate various algorithms, including Naïve Bayes and Decision Trees, can address this complexity by improving the classification of sentiments in a more granular manner.

These studies have not extensively compared the Naïve Bayes and Decision Tree methods in sentiment analysis. The novelty of this research is comparing the Naïve Bayes and Decision Tree methods. Both methods are chosen for their respective advantages: Naïve Bayes excels in classification performance and has a lower error rate when dealing with large datasets, requiring only

OPEN ACCESS

a small amount of training data to estimate the parameters used in probabilistic testing (Supriyadi, 2023; Suryawan et al., 2023). In contrast, the Decision Tree method is flexible, enhancing decision outcomes, and is effective in estimating missing data (Damanik et al., 2022; Tangkelayuk, 2022).

In summary, both Naïve Bayes and Decision Trees present viable methods for conducting sentiment analysis on YouTube comments regarding the closure of TikTok Shop. Naïve Bayes is advantageous due to its proven accuracy and efficiency, while Decision Trees offer interpretability and flexibility. The combination of these methods, possibly through ensemble techniques, could yield even more accurate and insightful results in understanding public sentiment on this topic.

Based on previous research, the authors will conduct a study on the closure of TikTok Shop. The difference between past research and the current study is that the current study collects sentiment data from YouTube through a crawling process using Google Colab with Python programming language, which is then processed using Python applications. This study also employs two methods to compare accuracy the Naïve Bayes method and the Decision Tree method. Additionally, this study's advantage is that it addresses the most recent issue in Indonesia. Therefore, this research can also assist Indonesia in resolving the issue at hand. Previous research has used various algorithms such as SVM and Naïve Bayes for sentiment analysis on different topics (lobster seeds, online lectures, ITE Law), but no research has directly compared Naïve Bayes and Decision Tree methods in sentiment analysis. In addition, there are no studies that highlight the current issue, namely the closure of TikTok Shop, or that use the YouTube platform as a data source. This research has the novelty of comparing the two methods and focusing on a more contextual and recent issue in Indonesia, namely the closure of TikTok Shop, using crawling techniques from YouTube.

3. Research Methods

Research Flow

In the data collection stage, the process involves using literature studies, which include observation and library methods, as well as the Crawling stage. This data collection is carried out to meet the needs of the objects used during the research process. Crawling is performed by retrieving YouTube comments. During the research process, several steps are taken in such a way as to obtain the conclusions of the research conducted by the authors. The research flow is as follows:

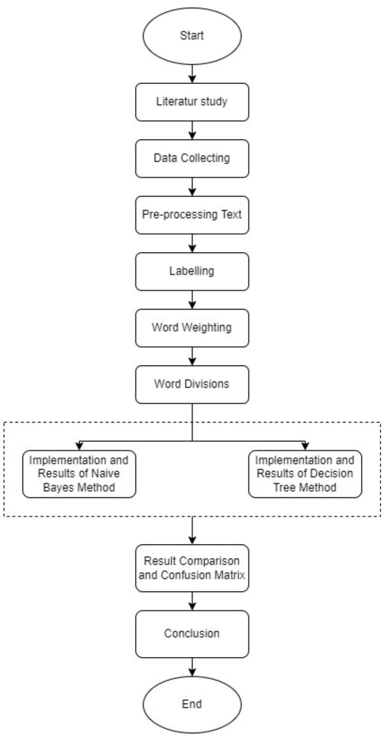

Fig. 1. Research Flow

The following is an explanation of the research flow. At this stage, the authors conduct a YouTube crawling process to gather the data needed throughout the research. The data collected consists of public comments related to the topic of TikTok Shop closure. Data collection is carried out using Google Colab with the Python programming language. This data collection process utilizes the YouTube Application Programming Interface (API). In this stage, video URL requests are used to access comment data by calling "commentThreads," which are then saved in a CSV file. The data crawling period is from October 6, 2023, to December 10, 2023. This research uses 7,000 raw data entries. The data used in this study are Indonesian-language comments that discuss the topic of TikTok Shop closure, which will be processed in the subsequent steps.

Preprocessing

Preprocesing is a step to obtain clean data for the next processes. This step is done manually and involves the following stages:

- Cleansing
  In this stage, unnecessary words, characters, and symbols are removed from the data. These include HTML characters (e.g., &), emojis (:), :0), hashtags (#), URLs, emails, punctuation marks, etc.
- Case Folding
  This process standardizes the text in YouTube comments to lowercase. The role of case folding is to ensure uniformity in the use of uppercase letters. The role of case folding is to standardize the use of uppercase letters. For example, if the text data we receive is written as "DaTA SCIence," case folding means converting all letters to lowercase. Meanwhile, other characters that are not letters or numbers, such as punctuation marks and spaces, are considered delimiters.
- Tokenizing
  Tokenizing, or tokenization, is the process of separating words in a document into independent words. Tokenizing is done to obtain tokens or word fragments that will become entities with value in the construction of the document matrix in the next process. Here is an example of a sentence after tokenization, input. "Please, the government should implement that regulation as soon as possible because it is truly very harmful. I support the government's steps.", Tokenized Output: "please", "the", "government", "should", "implement", "that", "regulation", "as", "soon", "as", "possible", "because", "it", "is", "truly", "very", "harmful", "I", "support", "the", "government's", "steps".
- Filtering
  Filtering is a process to remove frequently occurring words that do not have significant impact or meaning (stopwords).
- Stemming
  Stemming is a procedure to reduce all words to their common root form. Usually, affixes are removed to obtain the root word. Here is an example of a sentence after stemming, Input "closing tiktok shop to make tanah abang market busy seems to be a futile nonsense, people's habits have changed, all sellers have to follow whether they like it or not", output "closing tiktokshop to make tanah abang market busy seems to be a futile nonsense people's habits have changed all sellers have to follow whether they like it or not"

Data Labeling

The data that has undergone preprocessing will then be labeled or assigned attribute classes. The attribute class determination involves creating a lexicon dictionary of positive and negative Indonesian words in .txt format. This allows for labeling sentences that contain positive, negative, or neutral words as found in the dictionary (Suryadana & Sarasvananda, 2024). The attribute classes in this study are divided into positive, negative, and neutral. Here is an example of a document with its attribute class determined:

Table 1. Data Labeling

Text	Positive	Negative	Sentiment	Label
agree that the economic war until taxes come in destroys UMKM	1	2	-1	Negative
Thank you, the government has moved quickly to regulate and create healthy trade regulations so that Indonesian UMKM do not collapse.	6	1	5	Positive

Word Weighting

The word weighting stage is a mechanism for assigning values based on the frequency of a word's occurrence in a text document. One of the word weighting methods is TF-IDF (Term Frequency Inverse Document Frequency) (Lin, 2024). In TF-IDF weighting, pure TF is used, which counts how many times a term appears in the text; for example, if a term appears three times, it will have a score of three. Then, pure TF is normalized by dividing the pure TF score by the number of terms present in the document.

Data Splitting

After the data has been processed through preprocessing and labeled with sentiment classes, and word weighting (TF-IDF) has been applied, the comment data is divided into two types: training data and testing data. The training data is used to build the model, which is a representation of knowledge that will be used to predict the class of new, unseen data. The larger the training data set, the better the machine can understand data patterns (Radhitya et al., 2024). The testing data is used to evaluate how well the classifier performs in correctly classifying the data.

Table 2. Data Splitting

Training Data	Test Data
60%	40%
70%	30%
80%	20%

4. Result and Discussions

Implementation of "Sentiment Analysis of YouTube Comments on the Closure of TikTok Shop Using a Comparison of Naïve Bayes and Decision Tree Methods." This comparison is conducted to determine the accuracy of both methods. Thus, the results of this study will evaluate the performance of Naïve Bayes and Decision Tree methods in performing sentiment analysis on the topic of TikTok Shop closure. The research will provide an overview to the public on whether the topic of TikTok Shop closure leans towards positive, negative, or neutral opinions.

Accuracy Comparison

This process displays the df_eval DataFrame to view the accuracy comparison. By examining the accuracy for each formation and algorithm, we can see that the Decision Tree algorithm has superior accuracy compared to the Naive Bayes algorithm. Additionally, the data split can also affect the evaluation results, even if the difference is only a few percentage points.

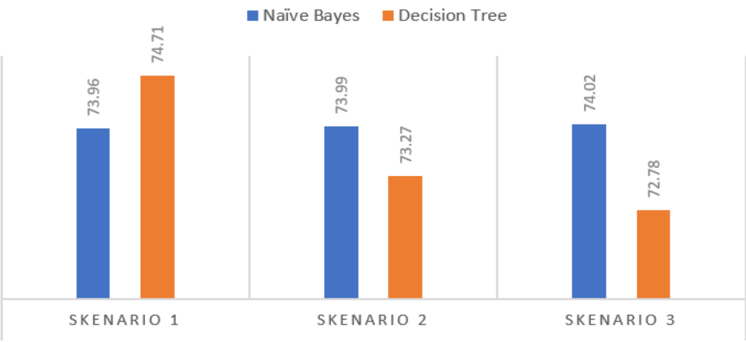

Fig. 2. Accuracy Comparison

Precision Comparison

Table 3. Precision Comparison

Methods	Scenario 1	Scenario 2	Scenario 3
Naïve Bayes	0.58	0.58	0.58
Decision Tree	0.57	0.55	0.55

The overall comparison represents the overall value of each aspect obtained from each method by calculating the percentage of positive, negative, and neutral classes and then dividing by three. The precision figures are as follows: for the Naïve Bayes method, in scenario 1, the precision is 0.58, in scenario 2, it is 0.58, and in scenario 3, it is 0.58. In contrast, for the Decision Tree method, the precision is 0.57 in scenario 1, and 0.55 in scenarios 2 and 3.

From the overall precision comparison using the Confusion Matrix across the three scenarios and two methods, it is evident that the Naïve Bayes method is superior with a precision of 0.58. The reason Naïve Bayes performs better than Decision Tree is that precision is more influenced by the amount of positive data (which is to be predicted correctly). Naïve Bayes generally achieves higher precision. However, the model performs poorly in classifying "neutral" and "positive" classes, as indicated by the low recall and F1-score for these classes. This suggests the model tends to classify most samples as "negative," with many misclassifications for "neutral" and "positive" classes. Therefore, the model needs improvement to enhance performance on "neutral" and "positive" classes.

On the other hand, the Decision Tree method shows relatively constant precision around 0.57 to 0.55 across scenarios, indicating that the Decision Tree is less effective in handling variability and noise in the data. The Decision Tree method excels in achieving better accuracy, particularly with balanced class data.

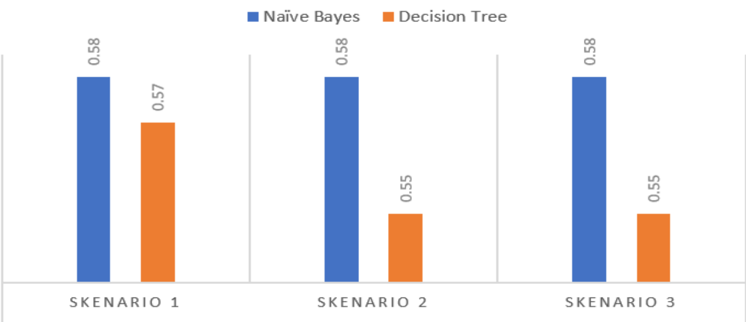

Fig. 3. Precision Comparison

Recall Comparison

Table 4. Recall Comparison

Methods	Scenario 1	Scenario 2	Scenario 3
Naïve Bayes	0.33	0.33	0.34
Decision Tree	0.57	0.55	0.55

OPEN ACCESS

The overall comparison represents the overall value of each aspect obtained from each method by calculating the percentage of positive, negative, and neutral classes and then dividing by three. The recall figures indicate that for the Naïve Bayes method, in all scenarios (1, 2, and 3), the recall values are 0.33, 0.33, and 0.34. This shows that Naïve Bayes has a relatively stable ability to retrieve positive information across all tested scenarios, albeit with a low success rate.

In scenario 1, the Decision Tree achieves the highest recall value of 0.57. In scenarios 2 and 3, the recall for the Decision Tree is 0.55. This indicates that the Decision Tree has a better capability to correctly identify actual positive classes, especially in the first scenario. Based on the recall analysis, the Decision Tree demonstrates better performance compared to Naïve Bayes in retrieving actual positive classes, particularly shown by the higher recall value in scenario 1. The Decision Tree's advantage in recall may be due to its ability to adapt the model well to training data, allowing it to better identify minority or outlier classes. Conversely, Naïve Bayes, with its assumption of feature independence and simpler model, tends to be less effective in capturing all actual positive examples. These recall results provide important insights into how Naïve Bayes and Decision Tree classify data, with the Decision Tree showing a better ability to recognize actual positive cases, which can be particularly advantageous in situations where identifying all positive cases is crucial.

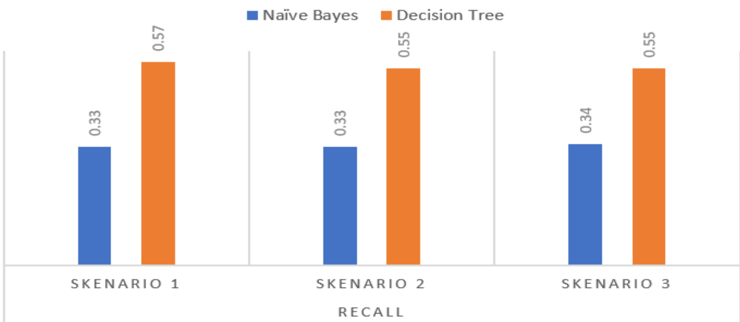

Fig. 4. Recall Comparison

F1-Score Comparison

Table 5. Recall Comparison

Methods	Scenario 1	Scenario 2	Scenario 3
Naïve Bayes	0.29	0.29	0.29
Decision Tree	0.57	0.55	0.55

The overall comparison represents the overall value of each aspect obtained from each method by calculating the percentage of positive, negative, and neutral classes and then dividing by three. The F1 Score figures for the Naïve Bayes method are 0.29 across all scenarios (1, 2, and 3). This indicates that Naïve Bayes has a relatively stable performance in retrieving information across the tested scenarios, but with a low success rate.

In scenario 1, the Decision Tree achieves the highest F1 Score of 0.57. In scenarios 2 and 3, the F1 Score for the Decision Tree is slightly lower but still stable at 0.55. This shows that the Decision Tree performs better in retrieving information compared to Naïve Bayes in this case. From the F1 Score analysis, it is evident that the Decision Tree, particularly in scenario 1, demonstrates the most superior performance in retrieving information compared to Naïve Bayes. This is consistent with the Decision Tree's more flexible and adaptive nature in handling complex data, while Naïve Bayes is often limited by its simple feature independence assumptions. Thus, the F1 Score results provide a clear view of the relative performance of Naïve Bayes and Decision Tree in the given scenarios and explain why the Decision Tree is considered superior in this case.

OPEN ACCESS

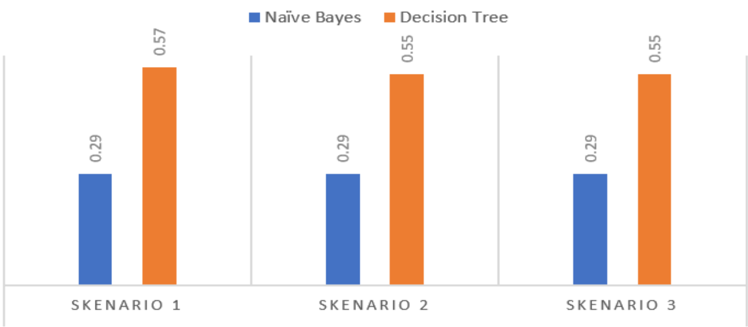

Fig. 5. F1-Score Comparison

Results of Testing the Decision Tree Method

The figure above summarizes the results of testing the Decision Tree method. Accuracy measures how well the model predicts overall, calculated as the percentage of correct predictions out of all predictions made. In Scenario 1, the Decision Tree achieved an accuracy of 74.71%. In Scenario 2, the accuracy was 73.27%. In Scenario 3, the accuracy was 72.78%. The highest accuracy was achieved in Scenario 1, indicating that the Decision Tree provided more accurate predictions in this scenario compared to the others.

Precision measures how many of the positive predictions are correct, calculated as the ratio of true positives (TP) to the total positive predictions (TP + false positives (FP)). In Scenario 1, the Precision was 57%. In Scenario 2, it was 55%. In Scenario 3, it was 55%. The highest Precision was also achieved in Scenario 1, showing that the Decision Tree had a better ability to accurately identify positive predictions in this scenario.

Recall (also known as sensitivity) measures how many of the actual positive class instances are correctly predicted by the model. In Scenario 1, the Recall was 57%. In Scenario 2, it was 55%. In Scenario 3, it was 55%. The highest Recall was achieved in Scenario 1, indicating that the Decision Tree was able to maintain the capability to recognize most of the actual positive class instances in this scenario.

The F1 Score is a combination of Precision and Recall, providing a single measure of how well the model identifies the positive class. In Scenario 1, the F1 Score was 57%. In Scenario 2, it was 55%. In Scenario 3, it was 55%. The highest F1 Score was also achieved in Scenario 1, reflecting that the Decision Tree had a better balance between Precision and Recall in this scenario.

Based on the results, the Decision Tree showed better performance in Scenario 1 compared to the other scenarios, with higher Accuracy, Precision, Recall, and F1 Score values. This indicates that the Decision Tree method is suitable for this specific test case, providing more accurate and consistent predictions for the data used. The Decision Tree method is appropriate for handling datasets with imbalanced class distributions and performs better than several other algorithms for structured data.

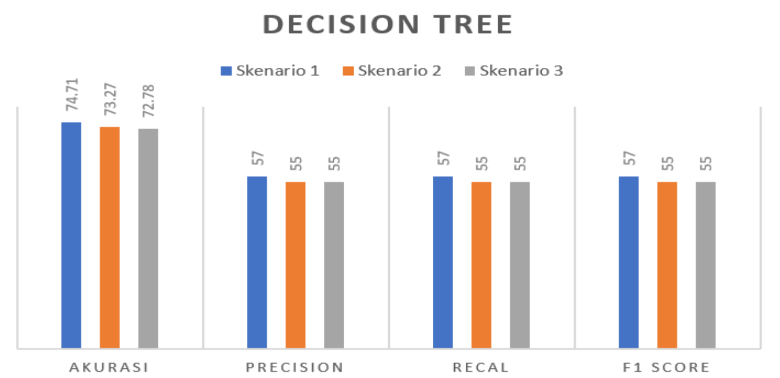

Fig. 6. Results of Testing the Decision Tree Method

Results of Testing the Naïve Bayes Method

The figure above shows a summary of the testing results for the Naïve Bayes method. Accuracy measures how well the model predicts overall, calculated as the percentage of correct predictions out of all predictions made. In Scenario 1, the Accuracy was 73.96%. In Scenario 2, the Accuracy was 73.99%. In Scenario 3, the Accuracy was 74.02%. The highest Accuracy was achieved in Scenario 3, indicating the consistency and stable performance of Naïve Bayes in making correct predictions in this scenario.

Precision measures how many of the positive predictions are correct, calculated as the ratio of true positives (TP) to the total positive predictions (TP + false positives (FP)). In Scenario 1, the Precision was 58%. In Scenario 2, the Precision was 58%. In Scenario 3, the Precision was 58%. Precision was consistent across all scenarios, showing that Naïve Bayes has a good ability to accurately identify positive predictions in these scenarios.

Recall (also known as sensitivity) measures how many of the actual positive class instances are correctly predicted by the model. In Scenario 1, the Recall was 33%. In Scenario 2, the Recall was 33%. In Scenario 3, the Recall was 34%. Recall was consistent across all scenarios, indicating that Naïve Bayes is able to maintain the ability to recognize most of the actual positive class instances.The F1 Score is a combination of Precision and Recall. In Scenario 1, the F1 Score was 29%. In Scenario 2, the F1 Score was 29%. In Scenario 3, the F1 Score was 29%. The F1 Score was the same across all scenarios, showing that Naïve Bayes has a uniform balance between Precision and Recall in the tested scenarios.Based on the results, Naïve Bayes shows stable Accuracy in identifying positive classes, particularly evident in all three scenarios with consistent Precision. However, the same F1 Score across all scenarios indicates that the model faces challenges in optimizing both Precision and Recall simultaneously. This suggests that Naïve Bayes is suitable for cases where the data in each class is balanced and performs well in terms of speed and efficiency in computation, both in training and prediction. Therefore, it is well-suited for very large datasets and cases where the classification problem is relatively simple and the assumption of feature independence is nearly met.

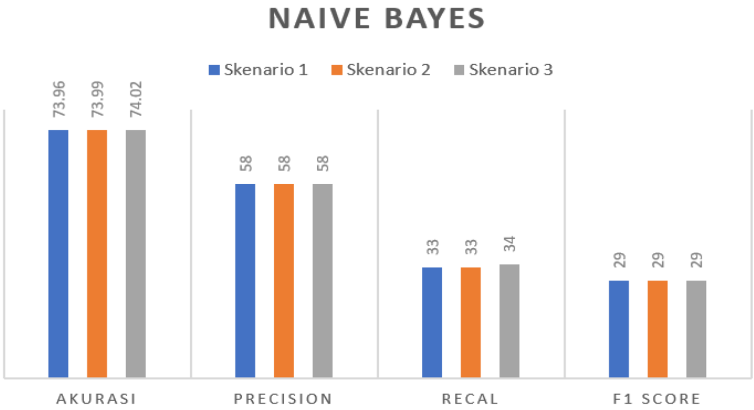

Fig. 7. Results of Testing the Naïve Bayes Method

5. Conclusion

Conclusions obtained from the research on sentiment analysis of YouTube comments about the closure of TikTok Shop using a comparison of Naïve Bayes and Decision Tree methods are: 1) From the research conducted, the results show: 4,889 negative sentiments, 1,429 neutral sentiments, and 1,537 positive sentiments. 2) From the testing results, it is evident that the Decision Tree method has superior accuracy compared to Naïve Bayes. In the data split, the Decision Tree method in Scenario 1 achieved an accuracy of 74.71%, while Naïve Bayes in Scenario 1 achieved an accuracy of 73.96%. In Scenario 2, the Decision Tree method achieved an accuracy of 73.27%, whereas Naïve Bayes achieved an accuracy of 73.99%. In Scenario 3, the Decision Tree method achieved an accuracy of 71.78%, while Naïve Bayes achieved an accuracy of 74.02%. Thus, it can be seen that the 80:20 data split using the Decision Tree method in Scenario 1 produced the highest accuracy compared to other scenarios. 3) From the evaluation results, the Decision Tree method achieved the highest precision at 57%, recall at 57%, and F1-score at 57%, while the Naïve Bayes method achieved a precision of 58%, recall of 34%, and F1-score of 29%. These evaluation results indicate that the Decision Tree method performs better compared to the Naïve Bayes method. 4) From all the processes involving Confusion Matrices, all methods and data splits indicate that the sentiment of the public regarding the closure of TikTok Shop leans more towards negative, as there are many comments opposing the closure of TikTok Shop.

References

Anwar, M. K. M. K., Yusoff, M., & Kassim, M. (2022). Decision tree and naïve bayes for sentiment analysis in smoking perception. *2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 294–299. https://doi.org/10.1109/ICSECC.2019.8907228

Aribowo, A. S., Basiron, H., Suryana, N., & Khomsah, S. (2020). An Evaluation of Preprocessing Steps and Tree-Based Ensemble Machine Learning for Analysing Sentiment on Indonesian YouTube Comments. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(5), 7078–7086. https://doi.org/10.30534/ijatcse/2020/29952020

Arya, I. K. A. G. W., Wiguna, G. W., & Sudipa, I. G. I. S. I. (2022). Sentiment Analysis Using Backpropagation Method To Recognize The Public Opinion. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *16*(4). https://doi.org/10.22146/ijccs.78664

Bozkurt, A. P., & Aras, I. (2020). Cleft Lip and Palate YouTube Videos: Content Usefulness and Sentiment Analysis. *The Cleft Palate-Craniofacial Journal*, *58*(3), 362–368. https://doi.org/10.1177/1055665620948722

Damanik, S. F., Wanto, A., & Gunawan, I. (2022). Penerapan Algoritma Decision Tree C4. 5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok. *Jurnal Krisnadana*, *1*(2), 21–32. https://doi.org/10.58982/krisnadana.v1i2.108

Danda, A. (2021). Gaming Sentiment: The Relationship of Comment Sentiment and Subscriber Growth Rate. *Journal of Student Research*, *10*(2). https://doi.org/10.47611/jsrhs.v10i2.1722

Deori, M., Kumar, V., & Verma, M. K. (2023). Analysis of YouTube video contents on Koha and DSpace, and sentiment analysis of viewers' comments. *Library Hi Tech*, *41*(3), 711–728. https://doi.org/10.1108/LHT-12-2020-0323

Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment 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*, 765–772. https://doi.org/10.1016/j.procs.2019.11.181

Jelodar, H., Wang, Y., Rabbani, M., Ahmadi, S. B. B., Boukela, L., Zhao, R., & Larik, R. S. A. (2021). A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments. *Multimedia Tools and Applications*, *80*, 4155–4181. https://doi.org/10.1007/s11042-020-09755-z

Kurniawan, A., Pattiasina, P. J., Rahman, A., Lestari, N. C., & Al Haddar, G. (2024). Utilization of Youtube as a Problem Solving-Based Learning Media. *TECHNOVATE: Journal of Information Technology and Strategic Innovation Management*, *1*(2), 62–68. https://doi.org/10.52432/technovate.1.2.2024.62-68

Lin, A. K. (2024). The AI Revolution in Financial Services: Emerging Methods for Fraud Detection and Prevention. *Jurnal Galaksi*, *1*(1), 43–51. https://doi.org/10.70103/galaksi.v1i1.5

Naz, S., Hina, S., Fatima, U., & Tabassum, H. (2023). A Hybrid Approach to Measure Students' Satisfaction on YouTube Educational Videos. *International Journal of Emerging Technologies in Learning (Ijet)*, *18*(09), 131–147. https://doi.org/10.3991/ijet.v18i09.38473

Puwaningwulan, M. M., Wulandari, T. A., & Anggaswari, N. A. (2024). The Impact of Tiktok Shop Policy Dynamics to The Micro-Small-Medium-Sized Enterprises Sustainability. *International Conference on Business, Economics, Social Sciences, and Humanities-Humanities and Social Sciences Track (ICOBEST-HSS 2024)*, 113–132. https://doi.org/10.2991/978-2-38476-269-9_11

Radhitya, M. L., Widiantari, N. K. M., Asana, M. D. P., Wijaya, B. K., & Sudipa, I. G. I. (2024). Product Layout Analysis Based on Consumer Purchasing Patterns Using Apriori Algorithm.

Journal of Computer Networks, Architecture and High Performance Computing, *6*(3), 1701–1711. https://doi.org/10.47709/cnahpc.v6i3.4400

Rismauli, E. I., & Suherman, S. (2024). Market Domination Through Social Media and E-commerce Merger in Business Competition Law's Perspective. *Journal of Law, Politic and Humanities*, *4*(4), 752–765. https://doi.org10.38035/jlph.v4i4.435

Septiani, A. T. D., Kuncoro, A. P., Subarkah, P., & Riyanto, R. (2024). Perbandingan Kinerja Metode Naïve Bayes Classifier dan K-Nearest Neighbor pada Analisis Sentimen Ulasan Mobile Banking Jenius. *Jurnal Krisnadana*, *3*(2), 67–77. https://doi.org/10.58982/krisnadana.v3i2.516

Supriyadi, A. (2023). Perbandingan Algoritma Naive Bayes dan Decision Tree(C4.5) dalam Klasifikasi Dosen Berprestasi. *Generation Journal*, *7*(1), 39–49. https://doi.org/10.29407/gj.v7i1.19797

Suryadana, K., & Sarasvananda, I. B. G. (2024). Streamlining Inventory Forecasting with Weighted Moving Average Method at Parta Trading Companies. *Jurnal Galaksi*, *1*(1), 12–21. https://doi.org/10.70103/galaksi.v1i1.2

Suryawan, I. G. T., Arimbawa, I. K. S., & Sudipa, I. G. I. (2023). Implementation of Naive Bayes Method for Granting Fisherman Business Credit. *Jurnal Info Sains: Informatika Dan Sains*, *13*(01), 24–32. https://doi.org/10.54209/infosains.v13i01

Tangkelayuk, A. (2022). The Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes, dan Decision Tree. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *9*(2), 1109–1119. https://doi.org/10.35957/jatisi.v9i2.2048

Udayana, I. P. A. E. D., Indrawan, I. G. A., & Putra, I. P. D. G. A. (2023). Decision Support System for Sentiment Analysis of Youtube Comments on Government Policies. *Journal of Computer Networks, Architecture and High Performance Computing*, *5*(1), 27–37. https://doi.org/10.47709/cnahpc.v5i1.1999

Wiratama, I. K., Welda, W., Permana, I. P. H., Aristana, M. D. W., & Sudipa, I. G. I. (2022). RECOMMENDATION FOR HIGH SCHOOL DETERMINATION BASED ON ACADEMIC POTENTIAL USING NAÃ VE BAYES METHOD. *JIKO (Jurnal Informatika Dan Komputer)*, *5*(2), 108–117. https://doi.org/10.33387/jiko.v5i2.4668

OPEN ACCESS