International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Sentiment Analysis Of Online Learning Based On Tiktok Comments Using The Naive Bayes Algorithm

Aripun[1], Ujang Juhardi[2], Dandi Sunardi[3], Muhammad Husni Rifqo[4]

[1,2,3,4]Universitas Muhammadiyah Bengkulu

aripunripun@gmail.com, ujangjuhardi@umb.ac.id , dandisunardi@umb.ac.id , mhrifqo@umb.ac.id

Abstract

Online learning is a form of information technology utilization in education that enables the learning process to be conducted without limitations of space and time. The development of online learning has generated various public responses, which can be found through user comments on social media platforms, particularly TikTok. This study aims to analyze public sentiment toward online learning based on TikTok comments using the Naive Bayes algorithm. The research data consisted of 1,193 comments collected through web scraping. The research stages included preprocessing, consisting of cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. The data were then labeled using the Lexicon-Based method and feature weighting was performed using TF-IDF. Subsequently, the data were classified into three sentiment categories, namely positive, negative, and neutral, using the Naive Bayes algorithm. Model evaluation was conducted using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The results showed that the Naive Bayes algorithm achieved an accuracy of 60.37%. These findings indicate that the Naive Bayes algorithm can be used to identify public sentiment toward online learning based on TikTok comments.

Keywords: Sentiment Analysis, TikTok, Online Learning, Naive Bayes, TF-IDF, Text Mining.

INTRODUCTION

The development of information and communication technology has driven significant transformation in the education sector, particularly through the implementation of online learning systems. This learning model became a primary solution during the COVID-19 pandemic and continues to be widely used because it provides flexibility in accessing learning materials

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

without being restricted by space and time. Nevertheless, the effectiveness of online learning remains a subject of debate, particularly regarding material comprehension, interaction between teachers and students, and students' learning motivation (Samsir et al., 2021).

In 2026, online learning has once again attracted attention in relation to discussions on energy-use efficiency. In April 2026, Liputan6.com reported the implementation of an online learning scheme for students in their fifth semester and above as part of an effort to promote a more efficient work culture. The implementation of online learning is considered capable of reducing student mobility as well as energy consumption within the campus environment, including the use of air conditioning and classroom lighting. This condition indicates that online learning is no longer associated solely with educational needs during the pandemic but has also re-emerged as an alternative mode of education related to resource efficiency.

The renewed discussion of online learning in 2026 has generated various responses and perspectives among the public. Differences in experiences and perceptions of online learning may lead to diverse opinions, including both support and criticism of its implementation. Social media has become a platform that enables people to express their opinions and respond directly to information they receive. One platform with such characteristics is TikTok, which allows users to respond to various types of content through its comment feature.

On the other hand, the emergence of social media platforms such as TikTok has created new opportunities in education as an alternative learning medium. TikTok is not only used for entertainment but is also utilized to deliver educational content in creative and interactive ways. Several studies have indicated that TikTok has the potential to serve as an effective learning medium because it can increase user engagement through attractive short-form videos (Apriani et al., 2021). This phenomenon reflects a paradigm shift from conventional learning methods toward approaches based on digital technology and social media.

Public perceptions of online learning presented through TikTok remain diverse. Comments posted by users reflect various opinions, attitudes, and experiences, which may be positive, negative, or neutral. Therefore, an analytical approach is required to systematically process and classify these opinions. One approach that can be employed is sentiment analysis, a text mining technique designed to identify and classify opinions based on specific polarities (Friska Aditia Indriyani et al., 2023).

In practice, the Naïve Bayes algorithm is one of the methods widely used in sentiment analysis. This is attributed to its simplicity, computational efficiency, and ability to classify textual data effectively. The algorithm is based on Bayes' theorem and assumes independence among features, making it suitable for processing large volumes of comment data, such as those found

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

on social media (Apriani et al., 2024). Furthermore, several studies have demonstrated that Naïve Bayes can achieve relatively high accuracy in social media-based sentiment analysis (Rahmadani et al., 2022).

Previous studies have examined sentiment analysis on social media, including Twitter and TikTok. For example, research on online learning on Twitter demonstrated that the Naïve Bayes method was effective in classifying public opinions during the pandemic (Samsir et al., 2021). In addition, another study analyzing TikTok comments using the same algorithm showed that the method was capable of providing a systematic overview of public perceptions (Pratama et al., 2026). However, studies specifically examining sentiment toward online learning based on TikTok comments remain relatively limited. Therefore, further investigation is necessary within this context.

TikTok was selected as the object of this study because it is one of the social media platforms experiencing rapid user growth and characterized by high levels of interaction through its comment feature. Unlike some other social media platforms, TikTok is frequently used to distribute educational content, allowing users to respond, share experiences, and provide direct evaluations of the learning materials presented. These comments constitute a valuable source of data for reflecting public perceptions of online learning. Therefore, sentiment analysis of TikTok comments is expected to provide relevant and up-to-date information that can serve as an evaluation resource for the development of digital-based learning.

Based on the discussion above, this study aims to analyze sentiments toward online learning expressed in TikTok comments using the Naïve Bayes algorithm. This study is expected to provide a long-term contribution to educators, educational content creators, and researchers in understanding users' perceptions of social media-based learning. Furthermore, the findings are expected to serve as a basis for developing more effective digital learning strategies that are better aligned with users' needs.

LITERATURE REVIEW

Sentiment Analysis

Sentiment analysis is a method applied within the field of Natural Language Processing (NLP) to identify, process, and classify opinions, feelings, or attitudinal tendencies contained in textual data (Bedir, 2021). This method can be utilized to determine public opinion trends regarding various objects—such as specific topics, products, services, or information—conveyed through digital media (Ligthart, 2021). According to Ligthart, sentiment analysis is a computational study focusing on the analysis of opinions, sentiments, emotions, and attitudes

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

found in user-generated text. With the advancement of information technology and digital media, the volume of content containing public opinions and responses has increased, creating a need for a method capable of identifying and processing the sentiment trends present in such data. Sentiment analysis enables the processing of unstructured text data into more systematic information, thereby providing an overview of user opinion trends or responses regarding a specific object. Sentiment analysis is applied here to identify public opinion trends regarding news information published by Kompas.com on Instagram. The data analyzed consists of public comments on the news content, which are subsequently classified into three sentiment categories: positive, negative, and neutral. Positive sentiment refers to comments indicating support, agreement, appreciation, or responses with a positive inclination. Conversely, negative sentiment indicates criticism, disagreement, disappointment, or other responses with a negative inclination. Meanwhile, neutral sentiment refers to comments that do not clearly exhibit either a positive or negative inclination. The results of this classification are then used to illustrate public response trends regarding the news information serving as the research subject (Pratama, 2022).

Online Learning

Online learning is an educational activity that utilizes the internet, LAN, and WAN as methods for delivery, interaction, and learning support, complemented by various other learning services (Anugrahana, 2020). It is a learning process conducted online using internet technology, learning applications, and social networks to deliver material, facilitate communication and interaction, and carry out assessments. Online learning takes place without direct face-to-face contact between educators and learners, relying instead on various available learning platforms (Pratama et al., 2020).

Advancements in information technology have significantly impacted the education sector, particularly through the implementation of online learning. This shift was accelerated by the COVID-19 pandemic, which compelled educational institutions to transition to digital learning systems. However, the implementation of online learning has elicited diverse perspectives—both positive and negative—from users; consequently, systematic analysis is required to evaluate these views and inform the formulation of educational policies (Fajar Sidik et al., 2022).

Social media has emerged as a primary source for gauging public opinion firsthand. Platforms like TikTok serve not only as entertainment venues but have also transformed into channels for disseminating educational content. User comments on educational content on TikTok reflect views, experiences, and levels of acceptance regarding online learning, making them valuable data sources for analysis using text mining approaches (Apriani et al., 2024).

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Sentiment analysis is a Natural Language Processing (NLP) technique used to identify opinions, attitudes, or emotions within text. This method enables the classification of opinions into specific categories, such as positive, negative, or neutral. In the context of online learning, sentiment analysis serves to measure satisfaction levels and identify user-facing issues, thereby providing a data-driven foundation for decision-making (Sarasvananda et al., 2022).

One algorithm widely used in sentiment analysis is Naive Bayes. This algorithm is probability-based and offers advantages such as simplicity, computational efficiency, and strong performance in text classification tasks. The working principle of Naive Bayes involves calculating the probability that a document belongs to a specific category based on the distribution of words contained within that document (Ernamia & Herliana, 2022).

Previous studies indicate that the Naive Bayes algorithm achieves competitive accuracy levels in social media sentiment analysis. For instance, research by Friska Aditia Indriyani et al. (2023) demonstrates that Naive Bayes can yield effective sentiment classification, even when compared to other methods like Support Vector Machines (SVM). This suggests that Naive Bayes is a relevant method for analyzing large volumes of text data.

A number of studies show that applying sentiment analysis based on classification algorithms—such as Naive Bayes—is not only effective for identifying opinion polarity but also capable of representing aggregate user perception trends within the context of online learning. This aligns with research findings stating that opinion data sourced from social media can be processed into strategic information to support the evaluation of educational policies, particularly in understanding satisfaction levels and the various challenges users face during digital learning processes (Saputri et al., 2024).

Research on sentiment analysis in the context of online learning typically utilizes data from platforms like Twitter. However, the use of data from TikTok remains relatively limited, despite the platform's unique characteristics, such as short-video-based interactions and more expressive comments. Therefore, this study proposes a novel approach that uses TikTok comments as the primary data source for sentiment analysis regarding online learning (Dina et al., 2025).

However, social media comments tend to be brief and unstructured, employing colloquial language and carrying meanings that vary according to context. These characteristics present analytical challenges, necessitating a computational approach capable of delving deeper into linguistic context. In this study, sentiment analysis is employed to assess public sentiment regarding online learning based on emotional expressions found in TikTok comments (Pratama et al., 2026).

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Tiktok

TikTok is a short-form video social media platform that enables users to create, upload, share, and enjoy a wide range of audiovisual content. The platform offers various supporting features—such as music, filters, and other creative tools—that allow users to express their creativity and produce engaging content. Its short-form video format makes it user-friendly and facilitates the rapid dissemination and consumption of information. Over time, TikTok has evolved beyond mere entertainment to serve as a medium for communication, self-expression, creativity, and the sharing of information with the public. (Muhamad Nur Fitrianto 2025)

TikTok also features interactive capabilities that allow users to respond to content by liking, commenting on, and sharing videos. The platform employs artificial intelligence-driven algorithms to analyze user preferences based on interaction patterns and subsequently presents content deemed relevant to the user's interests. Consequently, TikTok can be viewed as a digital medium that functions not only as a channel for information consumption but also as a space for interaction and the exchange of feedback among users. This makes TikTok a relevant platform for observing various user responses and opinion trends regarding specific information or topics. (Wicaksono et al. 2024)

Text Mining

Text mining is the process of analyzing text data to discover useful patterns or information using computational methods. It is a crucial aspect of sentiment analysis because the data being analyzed typically consists of unstructured text. The primary focus of text mining is to identify patterns, relationships, or new knowledge hidden within the text, which can then serve as a basis for decision-making.

Natural Language Processing

Natural Language Processing (NLP) is a subfield of artificial intelligence that focuses on the interaction between computers and human language. NLP enables systems to capture, interpret, process, and generate language that closely resembles human usage, whether in text or speech (Sawicki, Ganzha & Paprzycki, 2023).

In sentiment analysis, NLP plays a significant role because the information being examined consists of user-generated text. The NLP process typically involves data preparation stages such as cleaning, converting uppercase to lowercase, word normalization, tokenization, removing common words (stop words), and reducing words to their base forms (stemming or lemmatization). These steps aim to improve data quality, making it easier for machine learning

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

and deep learning algorithms to process.

In this study, NLP is used to prepare TikTok comment data into a consistent format prior to sentiment labeling and classification using the Naive Bayes algorithm.

Machine Learning

Machine learning is a branch of artificial intelligence that enables computers to learn patterns from data in order to make predictions or decisions without being explicitly programmed. In machine learning, models are built using training data, allowing them to recognize patterns that are subsequently used to classify or predict new data.

Within the field of Natural Language Processing (NLP), machine learning is widely utilized for various text processing tasks, such as document classification, spam detection, sentiment analysis, and topic recognition. Machine learning algorithms are capable of learning relationships between features in text data, allowing them to group documents into specific categories based on the characteristics of the words they contain.

Term Frequency–Inverse Document Frequency (TF-IDF)

Term Frequency–Inverse Document Frequency (TF-IDF) is a word-weighting method widely used in text mining and Natural Language Processing (NLP). This method aims to convert text data into a numerical representation so that it can be processed by machine learning algorithms. TF-IDF assigns higher weights to words that appear frequently in a specific document but rarely across the entire collection of documents; consequently, such words are considered more representative for distinguishing one document from another. The formula for Term Frequency is as follows:

$$TF(t,d) = \frac{f(t,d)}{\sum_{t' \in d} f(t',d)}$$

1. TF(t,d) = Term Frequency value of the word ttt in the ddd document
2. f(t,d)f(t,d)f(t,d) = number of occurrences of the word ttt in the document
3. ∑f(t′,d)\sum f(t',d)∑f(t′,d) = number of all words in the document

Next, Inverse Document Frequency (IDF) is used to reduce the weight of words that appear too frequently across all documents, as such words are considered less effective at distinguishing the characteristics of individual documents. The IDF value is calculated using the following equation:

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

$$IDF(t) = \log(\frac{N}{df(t)})$$

Notes:

1. N = total number of documents
2. df(t) = number of documents containing the word

The final weight of each word is obtained by multiplying the TF and IDF values, resulting in the TF-IDF value as follows.

$$\text{TF-IDF}(t,d)=\text{TF}(t,d)\times\text{IDF}(t)$$

Through this weighting process, words that contribute significantly to distinguishing a document receive higher values compared to common words that appear frequently across all documents. In this study, the TF-IDF method is employed to convert TikTok user comments regarding online learning into numerical features prior to classification using the Multinomial Naive Bayes algorithm. This TF-IDF feature representation is expected to enhance input data quality, thereby enabling the model to better recognize patterns of positive, negative, and neutral sentiment (Zhang, 2025).

Naive Bayes Algorithm

Naive Bayes is a probability-based classification algorithm that relies on Bayes' Theorem, operating under the assumption that features are independent of one another (Systems et al., 2025). This algorithm is frequently used in sentiment analysis due to its simplicity and satisfactory performance. Mathematically, the Naive Bayes formula can be expressed as follows:

$$P(H \mid X) = \frac{P(X \mid H) \cdot P(H)}{P(X)}$$

Here, P(H|X) represents the probability of hypothesis H given data X, P(X|H) is the probability of the data based on the hypothesis, P(H) is the prior probability of the hypothesis, and P(X) is the probability of the data. Based on this, Naïve Bayes was selected as the primary method for this study to classify TikTok comments regarding online learning into positive, negative, and neutral categories, while prioritizing computational efficiency and system reliability.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

RESEARCH METHOD

This study employs the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology to analyze the sentiment of TikTok user comments regarding online learning. The research stages include business understanding, data understanding, data preparation, modeling, evaluation, and deployment. During the modeling stage, the Multinomial Naïve Bayes algorithm is used to classify comments into three sentiment categories: positive, negative, and neutral.

The study focuses on TikTok user comments on videos discussing online learning. The data consists of Indonesian-language comments obtained via web scraping using Python. Videos were selected based on their relevance to the research topic, view count, and connection to information about online learning. The collected dataset ranges from 300 to 1,000 comments. Comments containing spam, duplicates, or irrelevant content were excluded from the analysis.

Preprocessing was performed to clean and standardize the comment data prior to classification. The steps included cleaning (removing URLs, emojis, numbers, symbols, and irrelevant characters), case folding (converting all text to lowercase), normalization (converting non-standard words and abbreviations into standard forms), tokenization (splitting sentences into individual words), stopword removal (eliminating words that carry little meaning for classification), and stemming using the Sastrawi library (reducing words with affixes to their root forms).

The preprocessed data was then labeled for sentiment using a lexicon-based approach. Each word in a comment was matched against an Indonesian sentiment dictionary and assigned a polarity value. The values for all words were summed to determine the sentiment category: a score greater than 0 indicated a positive comment, a score less than 0 indicated a negative comment, and a score of 0 indicated a neutral comment.

The labeled dataset was split into training data and testing data using an 80:20 ratio. The training data was used to build the classification model, while the testing data was used to evaluate the model's ability to classify new data.

Multinomial Naïve Bayes Algorithm

Sentiment classification is performed using the Multinomial Naïve Bayes algorithm. This algorithm applies Bayes' Theorem under the assumption that each feature or word is independent. The method is well-suited for text data because it can process features based on either word occurrence frequency or TF-IDF weights. The basic Naïve Bayes equation is:

$$P(H|X) = [P(X|H) \times P(H)] / P(X)$$

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Note: P(H|X) represents the probability of a data point belonging to a specific class, P(X|H) represents the probability of the data given a specific class, P(H) represents the prior probability of a class, and P(X) represents the overall probability of the data. Based on the obtained probability values, the model selects the sentiment class with the highest probability as the prediction result.

Model Evaluation

Evaluation is conducted to assess the model's ability to classify sentiment. Metrics used include the confusion matrix, accuracy, precision, recall, and F1-score. Accuracy measures overall prediction correctness; precision measures prediction accuracy for a specific class; recall measures the model's ability to identify correct data within a class; and the F1-score assesses the balance between precision and recall. The evaluation formulas used are:

Accuracy = (TP + TN) / (TP + TN + FP + FN) × 100%

Precision = TP / (TP + FP) × 100%

Recall = TP / (TP + FN) × 100%

F1-score = 2 × (Precision × Recall) / (Precision + Recall)

Experimental Parameters and Specifications

The model employs the following configuration: lowercase set to True, Indonesian stop words, an n-gram range of (1,1) (unigrams), a min_df of 2, and a max_df of 0.9. The dataset comprises 300–1,000 Indonesian-language TikTok comments across three sentiment categories: positive, negative, and neutral. Processing was performed using Python, utilizing the Sastrawi library for stemming and Scikit-learn for implementing the Multinomial Naïve Bayes algorithm and evaluating the model. The dataset was split using an 80/20 ratio for training and testing data.

RESULTS AND DISCUSSION

Data Description

The research data consists of 1,193 comments from a Liputan6 TikTok account post discussing plans for online learning aimed at energy efficiency. These comments were collected to analyze public sentiment trends regarding online learning using the Naive Bayes algorithm. The collected comments were analyzed based on their characteristics and content to classify them into three sentiment categories: positive, negative, and neutral. The comments were predominantly in Indonesian, featuring non-standard language, abbreviations, and slang

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

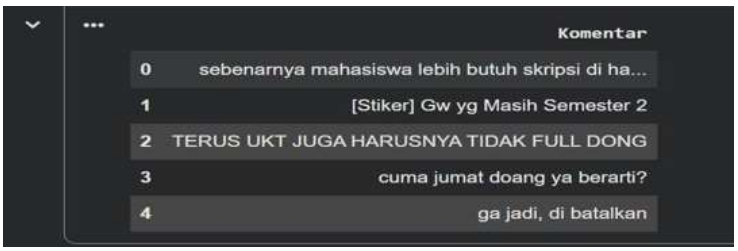
commonly used in social media communication.

Dataset

Prior to the analysis phase, the dataset underwent a filtering process to ensure that only comments relevant to online learning were included. Comments unrelated to the research topic, duplicates, or spam were excluded. Following the selection process, the data underwent preprocessing stages comprising cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. These steps were performed to produce cleaner, structured data ready for subsequent processing. Subsequently, each comment was assigned a sentiment label using a lexicon-based method, resulting in positive, negative, and neutral categories to serve as classification targets.

The research dataset comprises several attributes supporting the analysis process: date, comment, cleaned comment, normalized comment, tokenized comment, stopwords, stemming, stemming result, and sentiment label. The date attribute records when the comment was published, while the comment attribute contains the original text retrieved from TikTok. The sentiment label attribute indicates the labeling result used as a reference for building the classification model. Meanwhile, the other attributes represent the outputs of the various preprocessing stages conducted to prepare the data for feature extraction via the Term Frequency–Inverse Document Frequency (TF-IDF) method and subsequent processing using the Naive Bayes algorithm.

To provide an overview of the dataset used, Table 4.1 displays the first few entries from the research dataset.



	Komentar
0	sebenarnya mahasiswa lebih butuh skripsi di ha...
1	[Stiker] Gw yg Masih Semester 2
2	TERUS UKT JUGA HARUSNYA TIDAK FULL DONG
3	cuma jumat doang ya berarti?
4	ga jadi, di batalkan

Figure 4.1 Dataset

The next stage of this research involves compiling the dataset to serve as the primary input for the sentiment analysis process. The dataset consists of user comments from TikTok regarding online learning, obtained through web scraping. Each collected comment underwent a selection process to ensure that only data relevant to the research topic was

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

utilized. Subsequently, the data underwent preprocessing and sentiment labeling, assigning each comment a sentiment category—positive, negative, or neutral—based on the results of the Lexicon-Based labeling method. Once the dataset compilation was complete, the data was divided into training data and testing data using the train-test split method. The training data was used to build a classification model employing the Naive Bayes algorithm, while the testing data was used to evaluate the model's ability to classify sentiments in previously unseen comments. Finally, the model's performance was assessed using a Confusion Matrix by calculating Accuracy, Precision, Recall, and F1-Score. These evaluation results serve to determine the effectiveness of the Naive Bayes algorithm in classifying the sentiment of TikTok user comments regarding online learning. An example of the dataset used in this study is presented in Table IV.1.

Table IV.1 Dataset Results

No.	Comment
1	actually students need theses more in ha...
2	[Sticker] I'm still in Semester 2
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER
4	so it's only Friday, right?
5	no, it's canceled

Data Collection

The data collection phase is the initial step in this research, aimed at acquiring the data to be used in the sentiment analysis process. The research data consists of user comments from the TikTok platform responding to content regarding online learning. Data acquisition was performed using web scraping techniques—an automated process for extracting data from TikTok pages using the Python programming language.

During the scraping process, specific TikTok videos related to the topic of online learning were selected. The collected comments were then stored in a CSV-formatted dataset to facilitate data processing in subsequent stages. The acquired data was initially raw, containing various elements such as emojis, punctuation marks, uppercase letters, symbols, URLs, and words irrelevant to sentiment analysis. Each collected data entry primarily contained the text of the comment written by the TikTok user. All comments were subsequently reviewed to remove data unsuitable for the research objectives—such as comments unrelated to online learning, duplicates, or spam. This selection phase was conducted to ensure higher dataset quality and to enable the creation of a more accurate classification model.

The collected data served as the research dataset, undergoing subsequent stages of preprocessing, sentiment labeling, feature extraction (using TF-IDF), and classification (using the Naive Bayes algorithm). Examples of the collected TikTok comments can be seen in Table 4.1.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

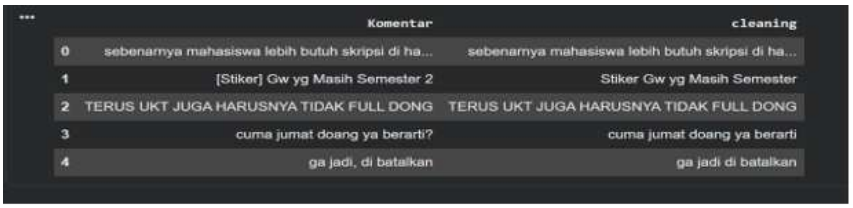
Data Preprocessing Phase

The preprocessing phase is the initial stage of data processing, aimed at improving dataset quality prior to sentiment analysis. Data obtained via web scraping is raw, containing various elements such as punctuation, numbers, emojis, URLs, symbols, uppercase letters, and non-standard vocabulary. These factors can impact the classification process; therefore, preprocessing is necessary to produce cleaner, more structured data.

In this study, preprocessing was carried out in stages using the Python programming language within the Google Colab environment. The preprocessing process included cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. Each stage serves a distinct function yet is interconnected, collectively producing data ready for TF-IDF weighting and classification using the Naive Bayes algorithm.

Through the preprocessing stage, each comment is transformed into a simplified form, allowing words with similar meanings to be represented identically. This results in more consistent data and reduced noise, thereby enhancing the quality of the features used in the sentiment classification process. The outcomes of each preprocessing stage are detailed in the following subsection.

Cleaning



	Komentar	cleaning
0	sebenarnya mahasiswa lebih butuh skripsi di ha...	sebenarnya mahasiswa lebih butuh skripsi di ha...
1	[Stiker] Gw yg Masih Semester 2	Stiker Gw yg Masih Semester
2	TERUS UKT JUGA HARUSNYA TIDAK FULL DONG	TERUS UKT JUGA HARUSNYA TIDAK FULL DONG
3	cuma jumat doang ya berarti?	cuma jumat doang ya berarti
4	ga jadi, di batalkan	ga jadi di batalkan

Figure 4.3 Cleaning process

The cleaning stage is the initial step in preprocessing, aimed at removing characters from comment data that are unnecessary for sentiment analysis. In this study, the cleaning process involved removing punctuation marks, special characters, symbols, emojis or stickers, and other characters that hold no informational value for the classification process. Data cleaning was performed to ensure the comment text was cleaner and easier to process in subsequent stages.

The cleaning results show that the core content of the comments was preserved, while irrelevant characters were successfully removed. For example, in the comment "[Stiker] Gw yg Masih Semester 2", the "[Stiker]" tag and the number "2" were removed, resulting in "Stiker Gw yg Masih Semester". Similarly, for the comment "cuma jumat doang ya berarti?", the question mark was removed, yielding the text "cuma jumat doang ya berarti". This cleaning process produced cleaner data, thereby enhancing the quality of data processing in the subsequent

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

stages of case folding, word normalization, tokenization, stopword removal, and stemming. Examples of the cleaning process results are presented in Table IV.3.

Table IV.3 Cleaning Results

No.	Original Comment	Cleaning Result
1	actually students need theses more in ha...	actually students need theses more in ha...
2	[Sticker] I'm still in Semester 2	Sticker I'm still in Semester
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER
4	so it's only Friday, right?	so it's only Friday, right

Sumber: Hasil preprocessing data menggunakan Python, diolah oleh peneliti (2026).

Case Folding



Figure 4.4 Case folding process

The case folding stage is a subsequent step following data cleaning, aimed at standardizing all comment text to lowercase. This process is performed to prevent discrepancies in word recognition caused by the use of uppercase versus lowercase letters. Consequently, words that share the same meaning but differ in capitalization are treated as a single word by the system during sentiment analysis.

The results of the case folding process show that comments originally containing uppercase letters—whether fully or partially capitalized—have been converted to lowercase without altering their content or meaning. For instance, the comment "TERUS UKT JUGA HARUSNYA TIDAK FULL DONG" was converted to "terus ukt juga harusnya tidak full dong," while "Stiker Gw yg Masih Semester" became "stiker gw yg masih semester." Meanwhile, comments that were already in lowercase remained unchanged. This stage yields a more consistent writing format, thereby enhancing data quality prior to the subsequent processes of word normalization, tokenization, stopword removal, and stemming. Examples of case folding results are presented

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

in Table IV.4.

Table IV.4 Case Folding Results

No.	Comment	Cleaning Result	Case Folding Result
1	actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...
2	[Sticker] I'm still in Semester 2	Sticker I'm still in Semester	sticker i'm still in semester
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	then the tuition fee should not be full either
4	so it's only Friday, right?	so it's only Friday, right	so it's only friday, right
5	no, it's canceled	no it's canceled	no it's canceled

Source: Results of preprocessing using Python, processed by the researcher (2026).

Normalisasi Kata



Figure 4.5 Normalization Process

The word normalization stage is a process carried out after case folding to convert non-standard words, abbreviations, slang, and spelling variations into standard forms in accordance with Indonesian language rules. This process aims to standardize word forms so that words sharing the same meaning are not treated as distinct terms during sentiment analysis. The normalization results show that several words originally in non-standard forms were successfully converted to their standard equivalents. For instance, "gw" was changed to "gue," "yg" to "yang," and "ga" to "tidak." Conversely, comments already consisting entirely of standard forms remained unchanged. This normalization process yields more consistent data, thereby improving the quality of text representation prior to the stages of tokenization, stopword removal,

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

stemming, TF-IDF weighting, and classification using the Naive Bayes algorithm. Examples of word normalization results are presented in Table IV.5.

Table IV.5 Normalization Results

No.	Comment	Cleaning	Case Folding	Normalization Result
1	actually students need theses more in ha…	actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...
2	[Sticker] I'm still in Semester 2	Sticker I'm still in Semester	sticker i'm still in semester	sticker i am still in semester
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	then the tuition fee should not be full either	then the tuition fee should not be full either
4	so it's only Friday, right?	so it's only Friday, right	so it's only friday, right	so it's only friday, right
5	no, it's canceled	no it's canceled	no it's canceled	it is not canceled

Source: Results of preprocessing using Python, processed by the researcher (2026).

Tokenizing



Figure 4.6 Tokenization process

The tokenization stage follows word normalization and aims to break down each comment into a collection of words or tokens. At this stage, each sentence is split based on spaces so that every word is recognized as a standalone unit. This process is crucial because classification algorithms do not process sentences in their entirety; instead, they utilize individual words as features for sentiment analysis. Based on the tokenization results, comments that were originally sentences are successfully converted into lists of words. For instance, the comment "stiker gue yang masih semester" is transformed into [stiker, gue, yang, masih, semester], while the

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

comment "terus ukt juga harusnya tidak full dong" is broken down into [terus, ukt, juga, harusnya, tidak, full, dong]. This process is carried out without altering the comment's meaning; it merely changes the data's presentation format to allow for the individual processing of each word. The tokenization output is subsequently used in the stopword removal stage to eliminate words lacking significant informational value, followed by stemming, TF-IDF weighting, and classification using the Naive Bayes algorithm. An example of the tokenization results is presented in Table IV.6.

Table IV.6 Tokenization Results

Comment	Cleaning	Case Folding	Normalization Result	Tokenize
actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...	[actually, students, need, theses, more, in, ha, ...]
[Sticker] I'm still in Semester 2	Sticker I'm still in Semester	sticker i'm still in semester	sticker i am still in semester	[sticker, i, am, still, in, semester]
THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	then the tuition fee should not be full either	then the tuition fee should not be full either	[then, the, tuition, fee, should, not, be, full, either]
so it's only Friday, right?	so it's only Friday, right	so it's only friday, right	so it's only friday, right	[so, it's, only, friday, right]
no, it's canceled	no it's canceled	no it's canceled	it is not happening, it is canceled	[it, is, not, happening, it, is, canceled]

Source: Results of preprocessing using Python, processed by the researcher (2026).

Stopword Removal



Figure 4.7 Stopword Removal Process

The stopword removal stage is a subsequent process following tokenization, aimed at

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

eliminating common words (stopwords) that do not provide significant information for sentiment analysis. Words such as *di*, *yang*, *lebih*, *masih*, *harusnya*, *juga*, and other common terms frequently appear in sentences but contribute little to determining sentiment polarity. Therefore, these words are removed so that the classification process focuses more on words that carry meaning and sentiment value. As a result of stopword removal, the number of words in each token list is reduced because words included in the stopword list are eliminated. For example, the comment "stiker gue yang masih semester," which was initially converted into the tokens [stiker, gue, yang, masih, semester], becomes [stiker, gue, semester] after the stopword removal process. In another example, the tokens [terus, ukt, juga, harusnya, tidak, full, dong] change to [ukt, full] because the words *terus*, *juga*, *harusnya*, *tidak*, and *dong* are included in the stopword list. Meanwhile, the tokens [tidak, jadi, di, batalkan] are reduced to just the word [batalkan] after the words lacking informational value are removed. An example of the stopword removal process results can be seen in Table IV.7.

Table IV.7 Stopword Removal Results

No.	Comment	Cleaning	Case Folding	Normalization Result	Tokenize	Stopword Removal
1	actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...	actually students need theses more in ha...	[actually, students, need, theses, more, in, ha, ...]	[students, need, theses, ha]
2	[Sticker] I'm still in Semester 2	Sticker I'm still in Semester	sticker i'm still in semester	sticker i am still in semester	[sticker, i, am, still, in, semester]	[sticker, semester]
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	then the tuition fee should not be full either	then the tuition fee should not be full either	[then, the, tuition, fee, should, not, be, full, either]	[tuition, fee, full]
4	so it's only Friday, right?	so it's only Friday, right	so it's only friday, right	so it's only friday, right	[so, it's, only, friday, right]	[friday]
5	no, it's	no it's	no it's	it is not	[it, is, not,	[happening,

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

	canceled	canceled	canceled	happening, it is canceled	happening, it, is, canceled]	canceled]

Stemming



Figure 4.8 Stemming Process

The stemming stage is the final step in the preprocessing process, aimed at converting each word into its root form (root word). This process is carried out using the Sastrawi library, allowing words with affixes (prefixes or suffixes) or non-standard forms to be reverted to their root word forms. Consequently, words that share the same meaning but are written in different forms are represented as a single, uniform word.

Based on the stemming results, it is evident that some tokens are converted into root forms, while tokens that were already root words remain unchanged. For example, the tokens [mahasiswa, butuh, skripsi, hapus, sih] remain "mahasiswa butuh skripsi hapus sih" after stemming, as all the words were already in their root forms. Similarly, for the comment "stiker gue semester," the result remains "stiker gue semester" because none of the words contain affixes. Likewise, the tokens [ukt, full] remain "ukt full" as neither word undergoes a change in form. The output of the stemming process serves as the final data for the preprocessing stage. This data is then used as input for the TF-IDF weighting process to convert the text data into a numerical representation prior to classification using the Naive Bayes algorithm. Examples of stemming process results are shown in Table IV.8.

Table IV.8 Stemming Results

No.	Comment	Cleaning	Case Folding	Normalization Result	Tokenize	Stopword Removal	Stemming
1	actually students need	actually students need	actually students need	actually students need theses more in	[actually, students, need,	[students, need, theses,	students need theses

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

	theses more in ha...	theses more in ha...	theses more in ha...	ha...	theses, ...]	ha]	remove
2	[Sticker] I'm still in Semester 2	Sticker I'm still Semester	sticker i'm still semester	sticker i am still semester	[sticker, i, am, still, semester]	[sticker, semester]	sticker semester
3	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	THEN THE TUITION FEE SHOULD NOT BE FULL EITHER	then the tuition fee should not be full either	then the tuition fee should not be full either	[then, tuition, fee, should, not, full, either]	[tuition, full]	tuition full
4	so it's only Friday, right?	so it's only Friday, right	so it's only friday, right	so it's only friday, right	[so, it's, only, friday, right]	[friday, right]	friday right
5	no, it's canceled	no it's canceled	no it's canceled	it is not happening, it is canceled	[it, is, not, happening, it, is, canceled]	[canceled]	cancel

Source: Results of preprocessing using Python, processed by the researcher (2026)

Sentiment Labeling



Figure 4.9 Sentiment Labeling

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

The sentiment labeling stage is carried out after the entire comment dataset has undergone preprocessing. At this stage, each comment—now in its stemmed form—is assigned a sentiment label using a lexicon-based method. This method works by matching each word in the comment against an Indonesian sentiment dictionary containing pre-assigned positive and negative polarity values. The polarity values of the individual words are then summed to produce a sentiment score for each comment.

Based on the calculated scores, the system automatically determines the sentiment category for each comment. If the total score is greater than zero (score > 0), the comment is categorized as having a positive sentiment. Conversely, if the total score is less than zero (score < 0), the comment is classified as having a negative sentiment. Meanwhile, if the total score equals zero (score = 0), the comment is labeled as neutral. This approach enables the labeling process to be performed automatically and consistently across the entire research dataset.

The labeling results shown in the figure indicate that each comment has a distinct score and sentiment classification based on its word content. For example, the comment "mahasiswa butuh skripsi hapus sih" (students need the thesis abolished) received a score of -8, resulting in a negative sentiment classification, whereas the comment "jumat doang ya" (just Fridays, right?) received a score of 6, placing it in the positive sentiment category. Meanwhile, the comments "stiker gue semester" and "ukt full" received a score of 0, classifying both as neutral.

Distribution of Sentiment Labeling Results.

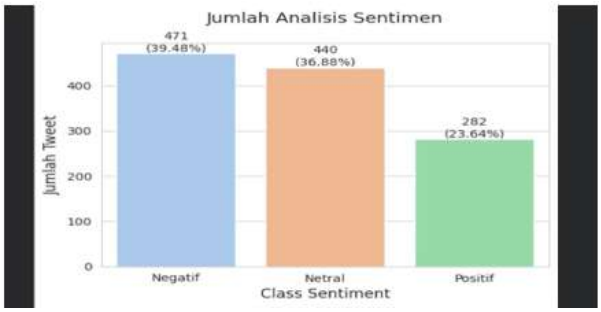


Figure 4.10 Sentiment Labeling Diagram

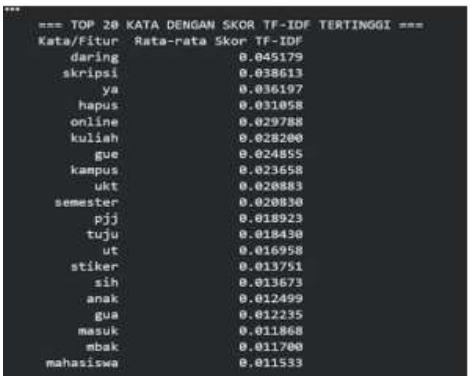
After the comment data underwent preprocessing and labeling using the Lexicon-Based method, an analysis was conducted on the distribution of data across the various sentiment categories. The purpose of this analysis was to determine the prevailing opinions of TikTok users regarding online learning based on the labeling results obtained. Figure 4.10 shows that the

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

research dataset comprises three sentiment categories: negative, neutral, and positive. The negative sentiment category contains the highest volume of data, with 471 comments, representing approximately 39.48% of the total dataset.

The neutral category follows with 440 comments (approximately 36.88%), while the positive category has the fewest entries, with 282 comments (approximately 23.64%). The predominance of comments in the negative category indicates that the majority of TikTok users hold an unfavorable view of online learning. These comments generally consist of complaints, criticisms, or accounts of unsatisfactory experiences during online classes. Conversely, the presence of neutral sentiment comments suggests that some users express their opinions objectively, without displaying a positive or negative emotional bias. Meanwhile, comments reflecting positive sentiment indicate that some users appreciate or view the implementation of online learning favorably, even though this group is smaller than the other two categories.

Proses TF-IDF



```
=== TOP 20 KATA DENGAN SKOR TF-IDF TERTINGGI ===
Kata/Fitur  Rata-rata Skor TF-IDF
     daring                0.045179
    skripsi                0.038613
         ya                0.036197
      hapus                0.031058
     online                0.029788
     kuliah                0.028200
        gue                0.024855
     kampus                0.023658
        ukt                0.020883
   semester                0.020830
        pjj                0.018923
       tuju                0.018430
         ut                0.016958
     stiker                0.013751
        sih                0.013673
       anak                0.012499
        gua                0.012235
      masuk                0.011868
       mbak                0.011700
  mahasiswa                0.011533
```

Figure 4.11 TF-IDF Process

After all comment data has undergone preprocessing, the next step is feature extraction using the Term Frequency–Inverse Document Frequency (TF-IDF) method. This method aims to convert text data into a numerical representation suitable for processing by the Naive Bayes algorithm. TF-IDF assigns a weight to each word based on its frequency within a specific document relative to its frequency across the entire dataset. A higher TF-IDF value indicates that a word contributes more significantly to distinguishing the characteristics of a specific comment from others.

Based on the TF-IDF weighting results, the 20 words with the highest average weights—considered the most representative within the research dataset—were identified. The word "daring" (online/remote) had the highest TF-IDF value at 0.045179, followed by "skripsi" (thesis)

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

at 0.038613, "ya" (yes) at 0.036197, "hapus" (delete) at 0.031058, and "online" at 0.029788. The high weights assigned to these words indicate that while they appear frequently in specific comments, they do not appear excessively across the entire dataset, thereby demonstrating a strong ability to differentiate the characteristics of individual comments.

Additionally, other words with relatively high TF-IDF values included "kuliah" (college/class), "gue" (I/me—informal), "kampus" (campus), "ukt" (tuition fee), "semester", "pjj" (remote learning), "tuju" (agree/aim), "ut" (Open University), "stiker" (sticker), "sih" (particle for emphasis), "anak" (student/child), "gua" (I/me—informal), "masuk" (enter/attend), "mbak" (Ms./female address), and "mahasiswa" (university student). These words illustrate that the majority of user comments relate to academic activities, online learning, tuition costs, and student experiences during the learning process.

These TF-IDF weighting results serve as the primary input for the classification process using the Naive Bayes algorithm. Each comment is represented as a numerical vector based on the TF-IDF values of its constituent words, enabling the model to learn word occurrence patterns across different sentiment categories. With this numerical representation, the classification process becomes more effective because the algorithm no longer processes the text directly, but instead processes the weights of each generated feature

Splitting Training and Testing Data

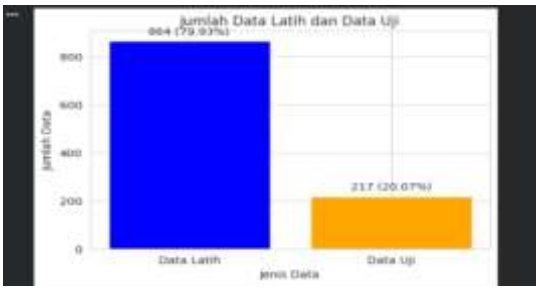


Figure 4.12 Data splitting process

Once all data has undergone the stages of preprocessing, sentiment labeling, and weighting using the TF-IDF method, the next step is to split the dataset into training data and testing data. This split is performed using the train-test split method with an 80:20 ratio. The objective of this division is to enable the Naive Bayes model to learn data patterns from the training set and subsequently be tested on the testing set to assess its ability to classify previously unseen data.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

As shown in Figure 4.12, the training set consists of 864 data points, representing approximately 79.93% of the total dataset, while the testing set comprises 217 data points, or about 20.07%. This proportion aligns with ratios commonly used in classification research, as it provides a sufficient volume of data for model training while ensuring adequate data for the evaluation process. The training data is used to build the classification model via the Naive Bayes algorithm by learning the relationships between the TF-IDF weighted features and the pre-determined sentiment labels.

Following the training phase, the model is tested using the testing data to evaluate its capability to classify comments into positive, negative, or neutral categories. Splitting the dataset into training and testing sets is a crucial stage in model development; it helps mitigate the risk of overfitting—a condition where the model recognizes training data well but performs poorly when classifying new data. Consequently, the evaluation results derived from the testing data provide insight into the model's performance in classifying the sentiment of previously unseen data.

Klasifikasi Naïve Bayes

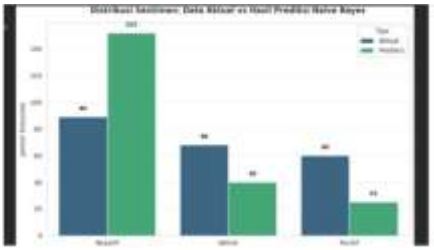


Figure 4.13 Naïve Bayes Classification Process

After the Naïve Bayes model has been trained using the training data, the next step is to perform the prediction process on the test data. The prediction results are then compared against the actual sentiment labels to determine the extent to which the model can classify comments into negative, neutral, and positive categories. This comparison is presented as a graph showing the sentiment distribution between the actual data and the prediction results.

Figure 4.13 reveals a discrepancy in the number of data points across sentiment categories when comparing the actual data to the model's predictions. In the negative category, the actual data consists of 89 documents, whereas the model predicted 152 documents as having negative sentiment. This indicates a tendency for the model to classify more comments as negative than the actual count.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

For the neutral category, there are 68 actual documents, while the model predicted only 40. Similarly, in the positive category, there are 60 actual documents, but the model predicted only 25 as positive. This situation suggests that some comments actually belonging to the neutral or positive categories were predicted by the model as negative.

The difference in distribution between the actual data and the predictions indicates that the Naïve Bayes model still makes classification errors on some test data. This may be attributed to the characteristics of social media comments—which are typically brief and often employ non-standard language, abbreviations, or words with ambiguous meanings—making it difficult for the model to accurately distinguish sentiment polarity. Furthermore, an imbalanced data distribution across classes can influence the model's tendency to predict specific classes. Nevertheless, this graph provides an overview of the model's initial capability to recognize sentiment patterns based on the provided training data. To assess the model's performance level more accurately, an evaluation using a Confusion Matrix—along with metrics such as Accuracy, Precision, Recall, and F1-Score—is required. The results of this evaluation will be discussed in the following subsection.

Model Evaluation

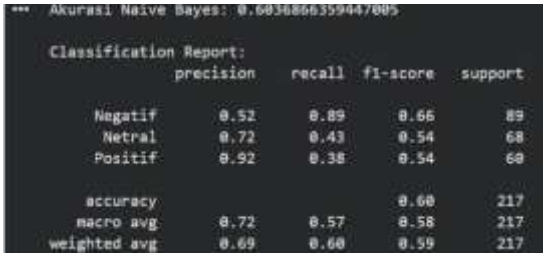


Figure 4.14 Model Evaluation Process

After the classification process using the Naive Bayes algorithm was completed, the next step was to evaluate the model's performance. Evaluation was conducted using a Classification Report, which displays accuracy, precision, recall, and F1-score values for each sentiment category. These metrics serve to assess the model's ability to classify comments into negative, neutral, and positive classes.

Based on the test results, the Naive Bayes model achieved an accuracy of 0.6037, or approximately 60.37%. This figure indicates that the model correctly classified about 60% of the test data out of the total 217 data points used in the testing process. Although the model successfully recognized sentiment patterns within the data, some instances were misclassified, thereby affecting the overall accuracy level.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

For the negative sentiment category, the model yielded a precision of 0.52, a recall of 0.89, and an F1-score of 0.66, with a sample size (support) of 89. The high recall value indicates that the model successfully identified the majority of comments that actually belonged to the negative category. However, the lower precision value suggests that some comments from other categories were incorrectly predicted as negative sentiment. For the neutral sentiment category, the results showed a precision of 0.72, a recall of 0.43, and an F1-score of 0.54, with a sample size of 68. These values indicate that while the model performed reasonably well in predicting neutral sentiment, it was unable to identify all the neutral comments actually present in the test data.

Meanwhile, for the positive sentiment category, the model achieved a precision of 0.92, a recall of 0.38, and an F1-score of 0.54, with a sample size of 60. The high precision value indicates that the positive predictions made by the model were largely correct. However, the relatively low recall value suggests that many positive comments went unrecognized and were instead predicted as belonging to other sentiment categories. Overall, the macro-average values obtained were 0.72 for precision, 0.57 for recall, and 0.58 for the F1-score. Meanwhile, the weighted average yielded a precision of 0.69, a recall of 0.60, and an F1-score of 0.59. These results indicate that the model's performance falls into the "reasonably good" category, yet there remains room to improve classification accuracy, particularly in identifying neutral and positive sentiments.

Based on these evaluation results, it can be concluded that the Naive Bayes algorithm is suitable for classifying the sentiment of TikTok user comments regarding online learning. However, the model's performance is influenced by factors such as the diverse nature of the text data, the use of non-standard language, and the distribution of data across the various sentiment classes. Therefore, future research could consider improving preprocessing quality, balancing the data, or employing alternative classification methods to achieve more optimal results.

Analisis *Wordcloud*

In addition to performing sentiment classification using the Naive Bayes algorithm, this study utilizes Wordcloud visualization to provide an overview of the words that appear most frequently in each sentiment category. A Wordcloud is a text data visualization technique that displays words based on their frequency of occurrence; the more often a word appears in the collection of comments, the larger it appears in the Wordcloud. Consequently, Wordclouds help identify the topics or issues most frequently discussed by TikTok users regarding online learning.

The Wordclouds were generated based on preprocessing results that underwent

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. These steps ensure that the displayed words accurately represent the content of the comments, free from the influence of common words (stopwords) or variations in writing style.

In the negative sentiment Wordcloud, the largest words represent terms that appear most frequently in comments expressing negative sentiment. These words generally relate to complaints, criticisms, or unpleasant experiences during online learning. This information highlights the aspects that most frequently trigger user concern or dissatisfaction regarding online learning.

Meanwhile, the neutral sentiment Wordcloud displays words frequently used in comments that are informative or lack a specific emotional tone. Comments in this category typically consist of questions, statements, or information that conveys neither positive nor negative assessments of online learning.

Finally, the positive sentiment Wordcloud displays words that appear most frequently in comments expressing appreciation or a favorable assessment of online learning. These words reflect positive experiences, support, and user satisfaction regarding the online learning process. Overall, Wordcloud visualization provides an easily understandable overview of the characteristics of each sentiment category. These visualizations also support the classification analysis performed using the Naive Bayes algorithm by highlighting the dominant words that contribute to the formation of each sentiment category.



Figure 4.15 Negative sentiment word cloud

The word cloud in Figure 4.15 visualizes the most frequently occurring words in comments categorized as having negative sentiment. This visualization is based on preprocessing results that underwent cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. The larger a word appears in the word cloud, the higher its frequency of occurrence within the set of comments labeled with negative sentiment.

Figure 4.15 shows that the word "daring" (online/remote learning) is the most dominant. This indicates that the majority of negative comments focus on the implementation of online

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

learning as the primary topic. Additionally, several other words appear prominently, such as "online," "skripsi" (thesis), "kampus" (campus), "bayar" (pay), "ukt" (tuition fee), "semester," "offline," "anak" (child/student), "mahasiswa" (university student), "hapus" (delete/cancel), "tuju" (aim/agree), and "gue" (informal "I"). The dominance of these words suggests that many users are expressing negative feedback regarding the online learning process, tuition fees (UKT), academic activities, and the conduct of classes on campus.

Beyond the large-sized words, there are others—such as "praktikum" (practicum), "kelas" (class), "tugas" (assignment), "fasilitas" (facilities), "sidang" (thesis defense), "mahal" (expensive), "susah" (difficult), "hilang" (lost/missing), and "bingung" (confused)—which, despite lower frequencies, still illustrate the various issues users face during online learning. These words indicate that negative comments relate not only to teaching methods but also to supporting facilities, academic workloads, and education costs.

Overall, the word cloud visualization demonstrates that negative sentiment is dominated by discussions regarding online learning, the conduct of remote classes, tuition fees (UKT), and various academic activities associated with student life. This visualization supports the earlier sentiment analysis findings, confirming that the majority of negative comments contain criticism, complaints, or expressions of user dissatisfaction regarding the implementation of online learning and its associated aspects.



Figure 4.16 WordCloud Positive sentiment

The word cloud in Figure 4.16 visualizes the words that appear most frequently in comments categorized as having positive sentiment. This visualization is based on preprocessing results that underwent cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. The larger a word appears in the word cloud, the higher its frequency of occurrence in comments expressing positive sentiment.

Figure 4.16 shows that the word "daring" (online) is the most dominant term. Additionally, several other words appear prominently, such as "ya" (yes), "online," "kuliah" (lectures/classes), "efektif" (effective), "banget" (very/really), "kampus" (campus), "lulus"

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

(graduate), "mahasiswa" (student), "langsung" (direct/in-person), "ajar" (teach), "tatap muka" (face-to-face), "hemat" (save/economical), "libur" (holiday/break), "semester," "jalan" (run/proceed), and "pjj" (distance learning). The dominance of these words indicates that many users responded positively to the implementation of online learning and various academic activities related to university studies.

Other words that appear include "praktikum" (practicum/lab work), "wisuda" (graduation), "bimbing" (guidance/mentoring), "efisien" (efficient), "gratis" (free), "diskon" (discount), "sekolah" (school), "materi" (material/content), and "sehat" (healthy). These words suggest that some users perceive online learning as offering benefits such as ease of attending classes, savings in time and cost, and the ability to maintain an effective learning process. The presence of words like "efektif," "hemat," and "lulus" reflects a positive perception regarding the flexibility of online learning in supporting academic activities.

Overall, the word cloud visualization reveals that comments with positive sentiment are dominated by words associated with the convenience, efficiency, and benefits of online learning. This visualization reinforces the sentiment analysis results, showing that despite various criticisms regarding online learning, many TikTok users appreciate and view its implementation positively. Thus, the word cloud provides a clearer picture of the topics and dominant words that shape the positive s



Figure 4.17 Neutral sentiment word cloud

The word cloud in Figure 4.17 displays the words that appear most frequently in comments categorized as having neutral sentiment. This visualization is based on preprocessing results that underwent cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. The size of each word in the word cloud indicates its frequency of occurrence; thus, larger words represent those most frequently used by users in comments labeled as neutral.

Figure 4.17 shows that the word "daring" (online) is the most dominant. Additionally,

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

several other words appear in relatively large sizes, such as "online," "skripsi" (thesis), "kampus" (campus), "semester," "gue" (informal "I"), "stiker" (sticker), "mahasiswa" (student), "bimbing" (guidance/supervision), "praktikum" (practicum), "pjj" (distance learning), "ukt" (tuition fee), and "mbak" (miss/ma'am). The dominance of these words indicates that comments with neutral sentiment generally discuss various academic activities related to online learning without displaying a clear emotional inclination—whether positive or negative.

Beyond the dominant words, there are others—such as "materi" (material), "dosen" (lecturer), "magang" (internship), "kuliah" (class/lecture), "luring" (offline/in-person), "minggu" (week), "jalan" (running/going), "masuk" (entering/starting), "program," and "mahal" (expensive)—that appear in smaller sizes. These words illustrate that most neutral comments consist of questions, information, or opinions regarding the learning process, coursework, and student activities, without expressing strong satisfaction or dissatisfaction.

Confusion Matrix Analysis

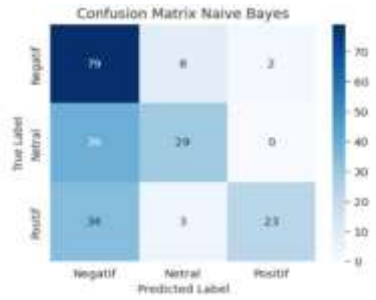


Figure 4.18 Confusion Matrix

After the Naive Bayes model performs the classification process on the test data, the next step is to evaluate the prediction results using a Confusion Matrix. The Confusion Matrix is an evaluation method used to compare the model's prediction results with the actual sentiment labels (true labels). Through the Confusion Matrix, one can determine the number of data points correctly classified as well as those misclassified across each sentiment category: negative, neutral, and positive.

Based on Figure 4.18, it is evident that the model successfully classified 79 negative comments correctly, whereas 8 negative comments were predicted as neutral and 2 were predicted as positive. These results indicate that the model possesses a reasonably good ability to recognize comments that truly belong to the negative sentiment category.

In the neutral sentiment category, 29 comments were correctly predicted. However, 39 neutral comments were predicted as negative, while no neutral comments were predicted as

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

positive. This indicates that the model still struggles to distinguish between neutral comments and those with a tendency toward negative sentiment. Meanwhile, in the positive sentiment category, the model was able to correctly classify 23 comments. However, 34 positive comments were predicted as negative, and 3 were predicted as neutral. This situation suggests that some positive comments share word characteristics similar to negative comments, causing the model to lean toward predicting them as the negative class. Overall, the Confusion Matrix reveals that the Naive Bayes model performs reasonably well in recognizing negative sentiment, yet its performance is suboptimal when distinguishing between neutral and positive comments. This is evident from the large number of data points in both categories being predicted as negative sentiment. This finding aligns with the results of the Classification Report, where the recall value for negative sentiment reached 0.89, whereas the recall values for neutral and positive sentiments were only 0.43 and 0.38, respectively. Consequently, the model shows a stronger tendency to classify comments as negative sentiment compared to the other two categories.

Based on these evaluation results, it can be concluded that the Naive Bayes algorithm successfully classified the sentiment of TikTok user comments regarding online learning with an accuracy rate of approximately 60.37%. However, significant misclassification errors persisted within the neutral and positive classes. Therefore, future research could consider employing data balancing techniques, optimizing model parameters, or applying other classification algorithms to improve classification performance.

Model Evaluation

Based on the three-class confusion matrix (Negative, Positive, and Neutral), the model's predictions across the 217 test data points were mapped out. Regarding the Recall performance for each class, the distribution of correct classification predictions is as follows:

- Correctly classified Negative data = 79 instances
- Correctly classified Neutral data = 29 instances
- Correctly classified Positive data = 23 instances

Overall accuracy calculation:

$$\boldsymbol{Akurasi} = (Number\ of\ Correct\ Classifications)\ /\ (Total\ Number\ of\ Data\ Points)$$

$$= \frac{79 + 29 + 23}{89 + 68 + 60} = \frac{131}{217} = 0{,}6037 = 60{,}37\%$$

$$\boldsymbol{Precision} = \frac{79}{79 + 73} = \frac{79}{152} = 51{,}97\%$$

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

$$\boldsymbol{Sensitifitas} = \frac{79}{79+10} + \frac{79}{89} = 88{,}76\%$$

$$\boldsymbol{Spesifitas} = \frac{55}{55+73} + \frac{55}{128} = 42{,}97\%$$

Analysis of Results

An accuracy rate of 60.37% indicates that the Naive Bayes model demonstrates a reasonably good ability to classify TikTok user comments regarding online learning into negative, neutral, and positive sentiments. Misclassifications in the remaining data were primarily caused by:

- Ambiguous sentences, where comments contained words capable of conveying different sentiment meanings.
- Non-standard language, including abbreviations, slang, and typos used by TikTok users.
- Comments with mixed opinions, containing more than one sentiment tendency within a single comment.
- Imbalanced data distribution across the sentiment classes.

Overall, the Naive Bayes model is capable of performing sentiment classification on TikTok comments, although it still encounters errors in distinguishing between neutral and positive sentiments. The research results show that applying the Naive Bayes algorithm to sentiment analysis of comments regarding online learning yields an accuracy rate of 60.37%. While the model can identify public sentiment trends, further development is required to improve classification performance across all sentiment categories.

CONCLUSION AND RECOMMENDATIONS

Based on the research conducted, it can be concluded that the Naive Bayes algorithm was successfully applied to perform sentiment analysis on TikTok user comments regarding online learning. The study began with data collection via web scraping, followed by preprocessing stages—including cleaning, case folding, word normalization, tokenization, stopword removal, and stemming—to produce cleaner, more structured data. Subsequently, the data was labeled for sentiment using a lexicon-based method and converted into a numerical representation via TF-IDF before the classification process.Testing results indicate that the Naive Bayes model achieved an accuracy of 60.37% in classifying comments into three sentiment categories: negative, neutral, and positive. Evaluation using a confusion matrix and classification report revealed that the model performed reasonably well in identifying negative sentiment comments

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

but struggled to distinguish between neutral and positive categories. This was evidenced by a number of comments in both categories being misclassified as negative. Additionally, Wordcloud visualization showed that terms such as *daring* (online), *online*, *skripsi* (thesis), *kampus* (campus), *semester*, *UKT* (tuition fees), and *mahasiswa* (university students) appeared most frequently in TikTok user comments. These words indicate that discussions regarding online learning primarily revolve around coursework activities, academic assignments, education costs, and students' learning experiences. Overall, this study demonstrates that the Naive Bayes algorithm is suitable for sentiment analysis of TikTok comments, although further development is required to improve the model's ability to classify all sentiment categories with greater accuracy.

REFERENCES

Anugrahana, Andri. "Hambatan, Solusi dan Harapan: Pembelajaran Daring Selama Masa Pandemi Covid-19 Oleh Guru Sekolah Dasar." 2020, 282–289.

Apriani, E., I. F. Hanif, and F. Oktavianalisti. "Analisis Sentimen Penggunaan TikTok Sebagai Media Pembelajaran Menggunakan Algoritma Naïve Bayes Classifier." MALCOM: Indonesian Journal of Machine Learning and Computer Science 4, no. 3 (2024).

Bedir, Ligthart, Alexander, Cagatay Catal, and Tekinerdogan. "Systematic Reviews in Sentiment Analysis: A Tertiary Study." Information and Software Technology 54 (2021). Springer Netherlands.

Dina, D. F. M., T. Haryanti, and M. A. Haq. "Analisis Sentimen Terhadap Komentar pada Media Sosial TikTok yang Berpotensi Menyebabkan Depresi Menggunakan Metode Naive Bayes." Computing Insight: Journal of Computer Science 7, no. 1 (2025): 1–9. https://doi.org/10.30651/comp_insight.v7i1.26327.

Ernamia, E. M. A., and A. Herliana. "Analisis Sentimen Kuliah Daring dengan Algoritma Naïve Bayes, K-NN dan Decision Tree." Jurnal Responsif: Riset Sains dan Informatika 4, no. 1 (2022): 1–11. https://doi.org/10.51977/jti.v4i1.614.

Fajar Sidik, Ibnu Suhada, Azhar Haikal Anwar, and Firman Noor Hasan. "Analisis Sentimen terhadap Pembelajaran Daring dengan Algoritma Naïve Bayes Classifier." Jurnal Linguistik Komputasional (JLK) 5, no. 1 (2022): 1–10.

Harahap, E. P., P. Herwanto, and A. Sulaiman. Evaluating Preprocessing Strategies and Twitter Sentiment Analysis for the 2024 Simultaneous Regional Elections. 2024.

Indriyani, Friska Aditia, Ahmad Fauzi, and Sutan Faisal. "Analisis Sentimen Aplikasi TikTok Menggunakan Algoritma Naïve Bayes dan Support Vector Machine." TEKNOSAINS: Jurnal

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Sains, Teknologi dan Informatika 10, no. 2 (2023): 1–9. https://doi.org/10.37373/tekno.v10i2.419.

Lestari, V. B., and C. A. Hutagalung. "Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes." 021: 36–44.

Palomino, M. A. "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis." Applied Sciences (2022).

Pratama, Fadhil Rifgi. "Terhadap Berita Sepakbola Liga Indonesia Logistic." 13, no. 3 (2022): 1338–1345.

Pratama, P. W., and E. W. Pamungkas. "Analisis Sentimen terhadap Komentar pada Video Viral (FYP) TikTok Menggunakan Metode Naïve Bayes." SKANIKA: Sistem Komputer dan Teknik Informatika 9, no. 1 (2026): 35–46.

Pratama, Rio Erwan, Sri Mulyati, Ogan Komering Ulu Timur, Dinas Pendidikan Kabupaten Ogan Komering Ulu, and Pembelajaran Luring. "Pembelajaran Daring dan Luring pada Masa Pandemi Covid-19." Gagasan Pendidikan Indonesia 1, no. 2 (2020): 49–59. https://doi.org/10.30870/gpi.v1i2.9405.

Rahmadani, P. S., F. C. Tampubolon, A. N. Jannah, N. L. H. Hutabarat, and A. M. Simarmata. "TikTok Social Media Sentiment Analysis Using the Naive Bayes Classifier Algorithm." SinkrOn 7, no. 3 (2022): 1–6. https://doi.org/10.33395/sinkron.v7i3.11579.

Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos. "Analisis Sentimen Penggunaan TikTok Sebagai Media Pembelajaran Menggunakan Algoritma Naïve Bayes Classifier." Jurnal Media Informatika Budidarma 5, no. 1 (2021): 1–7. https://doi.org/10.30865/mib.v5i1.2604.

Saputri, N. K. T. A., I. G. A. Gunadi, and I. M. G. Sunarya. "Analisis Sentimen Pelayanan Daring di Fakultas Teknik dan Kejuruan Universitas Pendidikan Ganesha Menggunakan Algoritma Naïve Bayes dan LSTM." MALCOM: Indonesian Journal of Machine Learning and Computer Science 4, no. 3 (2024): 1–10. https://doi.org/10.57152/malcom.v4i3.1336.

Sarasvananda, I. B. G., D. Selivan, M. L. Radhitya, and I. N. T. A. Putra. "Analisis Sentimen pada Pembelajaran Daring di Indonesia melalui Twitter Menggunakan Naïve Bayes Classifier." SINTECH (Science and Information Technology) Journal 5, no. 2 (2022): 1–7. https://doi.org/10.31598/sintechjournal.v5i2.1241.

Sawicki, J., M. Ganzha, and M. Paprzycki. "The State of the Art of Natural Language Processing—A Systematic Automated Review of NLP Literature Using." Data Intelligence (2023). https://doi.org/10.1162/dint.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Systems, I., Y. A. Wahid, and D. D. Andayani. “Performance Comparison of SVM and Naïve Bayes for Indonesian-Language Sentiment Analysis on Free Fire Reviews Using TF-IDF and SMOTE.” 6, no. 4 (2025): 674–689.

Wicaksono, Thomas Ari, Viki Love Reformasianto, Diffani Mufidah, Elsa Anggelina, Supra Wimbarti, and Sri Kusrohmaniah. “The Impact of TikTok Social Media on Users: A Neuropsychological Perspective.” Psikostudia: Jurnal Psikologi 13, no. 3 (2024): 455. https://doi.org/10.30872/psikostudia.v13i3.15920.

Zhang, L. “Features Extraction Based on Naive Bayes Algorithm and TF-IDF for News.” 2025: 1–17. https://doi.org/10.1371/journal.pone.0327347.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.