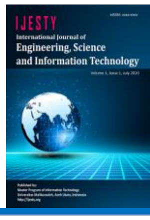
**International Journal of Engineering, Science & InformationTechnology**
Volume 4, No. 2 (2024) pp. 31-37
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v4i1.498
**Research Paper, Short Communication, Review, Technical Paper**




# Cosmetic Shop Sentiment Analysis on TikTok Shop Using the Support Vector Machine Method

**Rahmawati*, Wahyu Fuadi, Yesy Afrillia**

*Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia*

**Corresponding author Email: rahmawati.rahmabirn@gmail.com*

*The manuscript was received on 1 January 2024, revised on 30 March 2024, and accepted on 19 June 2024, date of publication 7 July 2024*

**Abstract**

User reviews are crucial in today's digital world for determining a product's quality. Nevertheless, these remarks are frequently disorganized and erratic, which confuses people and makes it challenging for them to make wise purchases. The erratic character of these reviews breeds uncertainty and makes determining a product's actual value more difficult. To help consumers more effectively evaluate and select products on platforms such as TikTok Shop, this study uses sentiment analysis tools. It hopes to accomplish this by improving the overall shopping experience and empowering customers to make more confident and informed selections. This research aims to assist consumers in evaluating and selecting products on TikTok Shop, an online shopping platform, by employing sentiment analysis techniques that help consumers make more informed decisions. In this study, a total of 500 comments from TikTok Shop users were collected as data. 350 comments have been set aside for training and 150 comments were set aside for testing. Data was gathered employing scraping, an automated process that makes use of the Python library's Selenium module to retrieve data from the internet. We employed the Support Vector Machine approach, an efficient machine learning tool for text classification, to assess the comments. 121 comments were categorized as having positive sentiment and 29 as having negative sentiment based on the test results. The system successfully recommended the "Ourluxbeauty" cosmetics store as a shop with many positive sentiments, indicating a recommendation level of 0.7 on the positive sentiment scale. The system's accuracy was measured using a Confusion Matrix, resulting in an accuracy rate of 78% and an inaccuracy rate of 22%. This demonstrates that the system can accurately classify comment sentiments and has significant potential for application in e-commerce practices to enhance the online shopping experience.

***Keywords***: *Analysis, Sentiment, TikTok Shop, Cosmetic Store, Support Vector Machine.*

## 1. Introduction

The proliferation of e-commerce platforms has revolutionized the way consumers engage in shopping activities, offering unparalleled convenience and accessibility. Among these platforms, Tik Tok has emerged as a prominent player with its innovative feature, Tik Tok Shop. This platform allows users and content creators to market and sell products directly through the app. The interactive nature of TikTok Shop, characterized by user-generated content and real-time engagement, has contributed to its rapid popularity. However, the unstructured and voluminous nature of user comments presents a significant challenge for consumers in discerning the quality and reliability of products and sellers.

Sentiment analysis, which provides a methodical way to assess and categorize user attitudes expressed in comments, has emerged as a crucial tool in tackling this problem. Sentiment analysis can classify comments as either positive or negative by using machine learning techniques, particularly the Support Vector Machine (SVM) algorithm, to convert unstructured textual data into meaningful insights. Through a clearer grasp of the general perception of products and sellers in the marketplace, this classification helps customers make educated selections.

This study's goal is to create a reliable sentiment analysis system specifically for TikTok Shop by analyzing user comments using the SVM algorithm. With reliable sentiment classification, the technology hopes to improve consumers' ability to make decisions. By using web scraping methods to obtain data.

Light increasing reliance on online reviews and the potential impact of consumer feedback on purchasing behavior, this research underscores the importance of advanced analytical methods in e-commerce. The findings from this study are expected to contribute to the broader field of sentiment analysis, offering practical implications for improving user experience and trust in online shopping platforms. By utilizing the Python library and web scraping techniques using Selenium available in the Python library. Many developers and major companies utilize Python to develop desktop, web, and mobile applications [1]. Moreover, Selenium can also be utilized for web scraping [2].

## 2. Literature Review

### 2.1. Tik Tok Shop

Through online platforms, e-commerce facilitates buying and selling transactions between sellers and buyers [3]. By utilizing online platforms, e-commerce enables easier exchanges in which sellers offer goods or services and buyers make purchases [4]. Due to its large user base, TikTok offers opportunities for promotion[5]. The TikTok application has now become a highly popular social media platform among users over the age of 16. Developed by Zhang Yiming from China, this program was officially released in 2016. Due to its ease and practicality of use [6].TikTok has become the most frequently downloaded software, reaching 45.8 million downloads, surpassing other social networking platforms such as WhatsApp, YouTube, Facebook, and even Instagram. As a social media platform, TikTok provides facilities for sharing a wide variety of content, including innovations, video challenges, lip-syncing, songs, dances, singing, and more. With a large user base, TikTok offers significant potential as a marketing platform [5]. One of the popular functions of the current social media platform is the TikTok Shop.

### 2.2. Sentiment Analysis

Sentiment analysis is a component of natural language processing (NLP) that is oriented towards studying opinions, sentiments, and emotions contained within the text [7]. Sentiment analysis can be utilized to categorize reviews into positive or negative sentiments [8]. By conducting sentiment analysis, we can automatically obtain opinions about products, brands, or emotional descriptions from netizens. The negative designation is represented by the number 0, which is applied to unfavorable ratings. Reviews that receive a positive rating in the interim are given the number 1, which stands for the positive label. The review data is then divided into two categories: training data and test data [9].

### 2.3. Data Mining

Data mining is a term used to describe the process of discovering knowledge within databases [10]. This process aims to extract and identify useful information and related insights from extensive databases. The branch of data mining known as text mining focuses on analyzing data in text format [11] [12].

### 2.4. Support Vektor Machine (SVM)

The best hyperplane to divide data points into N classes is found by SVM. In SVM, a dividing line, or hyperplane, is used to distinguish between classes (dimensions), and the distance between this hyperplane and the closest data points is called the margin. The support vectors are the closest data points to the hyperplane, coming from different classes, and are used as references to define the hyperplane. The decision function for classifying a new data point x is given by [13] [14]:

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \quad (1)$$

Definition:
$\mathbf{w}$ is the weight vector,
$b$ is the term,
$\mathbf{x}_i$ is the feature vector of the $i$-th training example

### 2.4. Implementation Step

1. **Data Collection :** The first step involves collecting the comments from TikTok Shop. This is done using web scraping techniques with Python's Selenium library, resulting in a dataset of 500 comments.

2. **Text Preprocessing:** In this context, crucial to ensure that the data is cleaner and more accurate before undergoing further processing [15]. In process, the unstructured text data is cleaned and converted a format that can be analyzed. Among the preprocessing actions are [16]:
   a. Removing stopwords
   b. Tokenization
   c. Stemming or lemmatization
   d. Converting text to lowercase
   e. Removing punctuation

3. **Feature Extraction:** The next step involves applying the Term Frequenc - Inverse Document Frequency (TF-IDF) to convert the cleaned text data into numerical characteristics. A word's significance inside a document is assessed using TF-IDF [17]. Transforming words into numerical values [18]. The TF-IDF formula is:

$$\mathrm{TF-IDF}(t,d) = \mathrm{TF}(t,d) \times \mathrm{IDF}(t) \quad (2)$$

   **Definition:**
   a. $\mathrm{TF}(t,d)$ is the term frequency of term $t$ in document $d$,
   b. $\mathrm{IDF}(t)$ is the inverse document frequency of term $t$,
   calculated as :

$$\mathrm{IDF}(t) = \log\left(\frac{N}{\mathrm{DF}(t)}\right) \quad (3)$$

**Definition:**

a. $N$ is the total number of documents,

b. $DF(t)$ is the number of documents containing term $t$.

4. **Model Training:** A 70:30 split of the dataset is made up of training and testing sets. The SVM model is trained using the training set. To find the ideal *w* and *b* that maximize the margin between the classes, the optimization issue is solved.
5. **Model Testing:** Using the testing set, the trained SVM model is evaluated. Several performance metrics are calculated to assess the efficacy of the model, such as accuracy, precision, recall, and F1-score.
6. **Evaluation using Confusion Matrix:** By comparing the anticipated and real labels, the confusion matri is used to evaluate the performance of classification It provides the metrics listed below [19].

**Table 2.** ***Confusion Matriks***

| Actual Class | Prediction Result Class | |
|---|---|---|
| | Positive | Negative |
| Positive | TP (True Positive) | FN (False Negative) |
| Negative | FP (False Positive) | TN (True Negative) |

The formulas for accuracy, precision, recall, and F1-score are[18]:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

**Definition:**

a. True Positives (TP)

b. True Negatives (TN)

c. False Positives (FP)

d. False Negatives (FN)

7. **Word Cloud:** A word cloud is a visual representation from a document that shows words that frequently appears by plotting those words in two dimensions so that they resemble an awan group [20].

## 3. Methods

The study was conducted on the TikTok Shop platform, specifically targeting cosmetic stores. User comments on TikTok Shop in Indonesia, a rapidly growing e-commerce platform, were collected as data. The data collection period ranged from August 2023 to October 2023. This study used a sentiment analysis model that made use of the Support Vector Machine (SVM) technique, which has been shown to have a high degree of text categorization accuracy in earlier research. Positive and negative user comment sentiments were categorized using the Support Vector Machine (SVM) technique. Web scraping techniques were employed in the data collecting procedure to obtain user comments from TikTok Shop. Selenium, a module from the Python library, was employed to automatically scrape the comments. A total of 500 comments were collected, with 350 comments designated as training data and 150 comments as testing data. Preprocessing steps included removing stop words, punctuation, case folding, stemming, and tokenization to transform the unstructured text into one that is ready for examination.

**Table 2. Data from Online Cosmetic Stores on TikTok Shop**

| Store Name | Number of Comments Collected |
|---|---|
| @kimberlybeauty88 | 494 |
| @ourluxbeauty | 41 |
| @Pinkflash.Beauty | 371 |
| @Jelitacosmetic | 56 |
| @Theoriginote | 105 |

The study targeted various cosmetic stores on TikTok Shop. These stores included popular ones like "kimberlybeauty88," "ourluxbeauty," "Pinkflash. Beauty," "Jelitacosmetic," and "Theoriginote." The comments on these stores were analyzed to determine the sentiment expressed by customers.

The following table presents the results after applying text preprocessing to user comments extracted from various cosmetic stores on TikTok Shop. The preprocessing steps have transformed the raw, unstructured text into a more structured and analyzable format.

## 4. Result And Discussion

Preprocessing is the process of converting unstructured text input into a more streamlined and organized state that is appropriate for sentiment analysis. There are multiple steps in this process:

a. Stopword Removal: Eliminating popular terms from the analysis that don't significantly advance its meaning.

b. Remove Punctuation: Get rid of any punctuation that might interfere with text processing.

c. Case Folding: To maintain uniformity, convert all text to lowercase.

d. Stemming: A process that unifies related phrases by reducing them to their basic forms.

e. Tokenization: Dividing the text into discrete words or tokens to make analysis simpler.

These comments come from customer reviews on various cosmetic stores on TikTok Shop. By preprocessing these comments, the data is prepared for sentiment analysis using the Support Vector Machine(SVM) method This analysis will help determine whether the comments express positive or negative sentiments.

## 4.1. Preprocessing Results

**Table 3. Text Processing Sample**

| No. | Before Preprocessing | After Preprocessing (Context Translated) |
|---|---|---|
| 1 | Sudah lama ak pake produk lamiela... Bagus2 bnyk macamnya | I use lamiela products, good, many types |
| 2 | Ini produk merkuri semua | Mercury products |
| 3 | Lamiela ini bagus bngt mnrut qczq juga pake sumpah .jujur bngt q pake bb nya pokoknya bagus glow di wjahkena keringt tambah glow ???? | Very good, use it honestly, it makes the face glow |
| 4 | Kak produk yg di kirim ke saya kurang | The product sent is lacking |
| 5 | Wah ternyata bedakya sangat bagus sangking bagusnya wajahku jadi merah merah | The powder is very good, it makes my face red |
| 6 | Produknya bagus sekali sehingga wajah saya merah glowing | Very good product, makes my face red and glowing |
| 7 | Awalnya pake sih bagus tapi udah pemakaian ke 3 kali bikin wajah saya merah apalagi pas wajah lagi ada jerawatnya malah bikin tambah iritasi | Initially good, but after 3 uses, it made my face red and irritated |
| 8 | Produknya sih merkuri tapi hasilnya bikin wajah tambah glow | Mercury product but makes the face glow more |

Explanations:

a. Before Preprocessing: This column shows the original comments extracted from TikTok Shop.

b. After Preprocessing: This column shows the comments after various preprocessing steps, such as removing unnecessary words, and punctuation, and converting to a uniform case, making the text ready for sentiment analysis.

## 4.2. Research Result

1. **Weighting Process (Sample Lexicon Data)**

The process of assigning weights to terms (words) in a document involves giving each term a numerical value that represents its importance in sentiment analysis. This step is performed before data labeling to ensure the terms are properly prepared for accurate classification. Here's a detailed explanation of the process:

a. **Collection of Terms**: Terms are collected and categorized as either positive or negative based on their sentiment.

b. **Assigning Weights**: A weight is given to each phrase. Generally, positive terms are given a weight (e.g., +1), while negative terms are given a weight (e.g., -1). The sentiment strength of each term is shown by the weights.

c. **Lexicon Creation**: A lexicon or dictionary of terms with their corresponding weights is created. This lexicon is used to evaluate the sentiment of each comment during the labeling process.

**Table 4. Indonesian Data Dictionary Sample**

| No. | Positive Words | Weight | Negative Words | Weight |
|---|---|---|---|---|
| 1 | Good | +1 | Bad | -1 |
| 2 | Great | +1 | Poor | -1 |
| 3 | Honest | +1 | Fake | -1 |
| 4 | Increase | +1 | Mercury | -1 |
| 5 | Product | +1 | Irritate | -1 |
| 6 | Many | +1 | Few | -1 |
| 7 | Glow | +1 | Acne | -1 |
| 8 | Very | +1 | Different | -1 |

2. **Steps in the Weighting Process:**

a. Identify Key Terms: Identify key terms that frequently appear in the comments and determine their sentiment orientation (positive or negative).

b. Assign Weights: Assign a weight of +1 to positive terms and -1 to negative terms.

c. Compile Lexicon: Compile the identified terms and their weights into a lexicon for use in sentiment analysis.

d. Prepare for Labeling: Use the lexicon to evaluate each comment's sentiment by summing the weights of the terms present in the comment.

3. **Word Labeling**

The data labeling process in the document involves assigning a sentiment label (positive or negative) to each comment based on specific keywords. Here's a brief explanation:

a. **Collection of Comments**: Comments are collected from various cosmetic stores on TikTok Shop.

b. **Preprocessing**: Comments undergo preprocessing steps such as stopword removal, punctuation removal, case folding, stemming, and tokenization to clean and structure the text.

c. **Keyword Identification**: Keywords associated with positive and negative sentiments are identified. For example:

1. Positive keywords: "good," "great," "love," etc.
2. Negative keywords: "bad," "poor," "mercury," etc.

d. **Label Assignment**: Each comment is analyzed for the presence of these keywords:

1. If a comment contains predominantly positive keywords, it is labeled as "positive."

2. If a comment contains predominantly negative keywords, it is labeled as "negative."
e. **Manual Verification**: The labeled data is manually verified to ensure accuracy.

**Table 5. Labeling Dataset**

| No. | Comment (Preprocessed) | Label |
|---|---|---|
| 1 | I use lamiela products, good, many types | Positive |
| 2 | Mercury products | Negative |
| 3 | Very good, use it honestly, it makes the face glow | Positive |
| 4 | The product sent is lacking | Negative |
| 5 | The powder is very good, it makes my face red | Positive |
| 6 | A very good product makes my face red and glowing | Positive |
| 7 | Initially good, but after 3 uses, it made my face red and irritated | Negative |
| 8 | Mercury product but makes the face glow more | Positive |

Finding a term's relative significance within a document and throughout a corpus of documents can be done using the TF-IDF (Term Frequency-Inverse Document Frequency) weighting approach.

**4. Table TF-IDF Weighting**

This table shows the term frequency (TF) and document frequency (DF) for each term across different documents.

**Table 6. TF-IDF weighting**

| No. | Term | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | DF |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Good | 1 | 0 | 2 | 0 | 2 | 1 | 1 | 0 | 7 |
| 2 | Great | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 |
| 3 | Honest | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 4 | Increase | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 3 |
| 5 | Product | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 5 |
| 6 | Mercury | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 7 | Many | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 8 | Glow | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 3 |
| 9 | Few | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 10 | Very | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 11 | Different | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

The final TF-IDF scores for each phrase across several documents are displayed in this table. These scores are obtained by multiplying the term frequency (TF) by the inverse document frequency (IDF).

**Table 7. Results of TF-IDF**

| No. | Term | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | Total Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Good | 0.058 | 0 | 0.116 | 0 | 0.116 | 0.058 | 0.058 | 0 | 0.406 |
| 2 | Great | 0 | 0 | 1.204 | 0 | 0 | 0 | 0 | 0 | 0.602 |
| 3 | Honest | 0 | 0 | 0.903 | 0 | 0 | 0 | 0 | 0 | 0.903 |
| 4 | Increase | 0 | 0 | 0.426 | 0 | 0 | 0 | 0.544 | 0.426 | 1.396 |
| 5 | Product | 0.204 | 0.204 | 0 | 0.204 | 0 | 0.204 | 0 | 0.204 | 1.020 |
| 6 | Mercury | 0 | 0.602 | 0 | 0 | 0 | 0 | 0 | 0.602 | 1.204 |
| 7 | Many | 0.903 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.903 |
| 8 | Glow | 0 | 0 | 0.852 | 0 | 0 | 0 | 0 | 0.426 | 1.278 |
| 9 | Few | 0 | 0 | 0 | 0.903 | 0 | 0 | 0 | 0 | 0.903 |
| 10 | Very | 0 | 0 | 0 | 0 | 0.903 | 0 | 0 | 0 | 0.903 |
| 11 | Different | 0 | 0 | 0 | 0 | 0.903 | 0 | 0 | 0 | 0.903 |

Steps for Calculating TF-IDF:

1. **Term Frequency (TF):** The number of times a term appear i a document.
2. **Document Frequency (DF):** The quantity of the term appearing in documents.
3. **The formula for calculating the inverse document frequency (IDF) is IDF** = log(total number of documents / number of documents containing the phrase). This indicates the importance of terms that appear in fewer texts by giving them more weight.
4. **TF-IDF Score:** Determined by multiplying TF by IDF. This combines the term's occurrences within a document with its overall corpus infrequency.

**Calculation of the SVM Method**

After acquiring the TF-IDF weights, the subsequent step is to use the Support Vector Machine (SVM) method to classify the sentiments of the comments. The SVM method is employed to determine the optimal hyperplane that separates the data points into different classes (positive and negative sentiments).

**Steps of Calculation**

1. **Preparation of Data**:
   a. The dataset includes comments that have been preprocessed and assigned labels.
   b. TF-IDF weights for each term in the comments are calculated and serve as input features for the SVM model
2. **Setting Up the SVM**:
   a. Define the SVM model with a linear.
   b. This sounds good the way it is
   c. The labels are the sentiment classes (positive or negative).

3. **Formulation of the SVM Objective:**
   a. The objective of the SVM function is to minimize the margin between the two classes.
   b. The decision function can be written as: f(x) = w·x + b, where w represents the weight vector, x denotes the feature vector, and b is the bias term.
4. **Constraints:**

$w.x + b = 0$

5. **Solving for Bias (b):**
   Combine the first and second equations to solve for :

$1088 \times 15 + b = 0$

$169 \times 15 + b = 0$

Solving these gives $b = -1704$.

6. **Formulating the Decision Function:**
   Substitute the weight vector and bias into the decision function:

   For the first data point: $f(x_1) = 1165 \times 15 + (-1704) = 435$ (Positive)
   For the second data point: $f(x_2) = 0806 \times 15 + (-1704) = -495$ (Negative)
   Repeat the calculations for the remaining data points.

7. **Testing New Data:**
   Apply the SVM model to new data points to predict their sentiment:
   $f(x_5) = 1922 \times 15 + (-1704) = 1179$ (Positive)

   $f(x_6) = 2068 \times 15 + (-1704) = 1398$ (Positive)

   $f(x_7) = 1658 \times 15 + (-1704) = 783$ (Positive)

   $f(x_8) = 229 \times 15 + (-1704) = 3435$ (Positive)

The SVM method is applied to classify the sentiment of comments based on their TF-IDF features. By calculating the decision function for each data point, the model accurately identifies positive and negative sentiments. This method ensures a robust classification that can be used to improve the sentiment analysis of comments on TikTok Shop. A classification model's accuracy is assessed using a confusion matrix, which compares the anticipated and real labels. It consists of four components:
a. True Positive (TP): The number of instances where positive predictions are correct.
b. True Negative (TN): The number of instances where negative predictions are correct.
c. False Positive (FP): The number of instances mistakenly predicted as positive.
d. False Negative (FN): The number of instances mistakenly predicted as negative.

8. **Steps to Calculate Confusion Matrix**
   a. Collect Predicted and Actual Labels: Gather the predicted sentiment labels from the SVM model and the actual labels from the test data.
   b. Construct the Matrix: Populate the confusion matrix with the counts of TP, TN, FP, and FN.

**Table 7. Confusion Matrix**

| Actual \ Predicted | Positive (Predicted) | Negative (Predicted) |
|---|---|---|
| Positive (Actual) | TP = 93 | FN = 5 |
| Negative (Actual) | FP = 28 | TN = 24 |

9. **Accuracy Testing**
   Accuracy is calculated using the confusion matrix components. It calculates the percentage of cases that are correctly classified out of all the instances.
   Using the confusion matrix values from the example above:
   a. $TP = 93$
   b. $TN = 5$
   c. $FP = 28$
   d. $FN = 24$

$$\text{Accuracy} = \frac{93 + 24}{121 + 29} = \frac{115}{150} = 0,76 = 76\,\%$$

So, the accuracy of the model is $78\%$. The confusion matrix and accuracy testing provide a comprehensive evaluation of the classification model's performance. By calculating these metrics, we can understand how well the model distinguishes between positive and negative sentiments and identify areas for improvement.

## 5. Conclusion

Data Collection and Analysis: The study focused on cosmetic stores on TikTok Shop, utilizing web scraping with the Selenium module in Python to gather a dataset of 500 comments. Out of these, 350 comments were used for training and 150 for testing, maintaining a 70:30 ratio. Using the Support Vector Machine (SVM) method, it was found that out of 150 test data points, 121 comments were classified as positive and 29 as negative. Evaluation Metrics: The evaluation using the Confusion Matrix yielded the following results: an accuracy of 78%, precision of 76.85%, recall of 94.89%, and an F-1 score of 84.93%. The percentage of inaccurate data classification stood at 22%. Recommended Stores Based on Sentiment Analysis: The sentiment analysis identified "ourluxbeauty" as the most recommended store, with 70% positive and 30% negative sentiments. "Pinkflash.beauty" followed with 69% positive and 31% negative sentiments. "Kimberlybeauty88" was next with 65% positive and 35% negative sentiments. "Theorginote" had 62% positive and 38% negative sentiments, and "jelitacosmetic" also had 62% positive and 38% negative sentiments.

## References


[1] Muhammad Romzi and B. Kurniawan, "Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma," *JTIM J. Tek. Inform. Mahakarya*, vol. 03, no. 2, pp. 37–44, 2020.

[2] A. S. Yondra, D. Triyanto, and S. Bahri, "Implementasi Web Scraping untuk Mengumpulkan Informasi Produk dari Situs E-commerce dan Marketplace dengan Teknik Pemrosesan Paralel," *Coding J. Komput. dan Apl.*, vol. 10, no. 01, pp. 93–102, 2022.

[3] A. J. Putri, A. S. Syafira, and M. E. Purbaya, "Analisis Sentimen E-Commerce Lazada pada Jejaring Sosial Twitter Menggunakan Algoritma Support Vector Machine," vol. 01, no. 1, 2022.

[4] J. A. Zulqornain and P. P. Adikara, "Analisis Sentimen Tanggapan Masyarakat Aplikasi Tiktok Menggunakan Metode Naïve Bayes dan Categorial Propotional Difference ( CPD )," vol. 5, no. 7, pp. 2886–2890, 2021.

[5] C. B. Dewa and L. A. Safitri, "Pemanfaatan Media Sosial Tiktok Sebagai Media Promosi Industri Kuliner Di Yogyakarta Pada Masa Pandemi Covid-19 (Studi Kasus Akun TikTok Javafoodie)," *Khasanah Ilmu - J. Pariwisata Dan Budaya*, vol. 12, no. 1, pp. 65–71, 2021, doi: 10.31294/khi.v12i1.10132.

[6] A. Aldila Safitri, A. Rahmadhany, and I. Irwansyah, "Penerapan Teori Penetrasi Sosial pada Media Sosial: Pengaruh Pengungkapan Jati Diri melalui TikTok terhadap Penilaian Sosial," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 3, no. 1, pp. 1–9, Jan. 2021, doi: 10.47233/jteksis.v3i1.180.

[7] Z. Fitra Ramadhan and A. Benny Mutiara, "Sentiment Analysis of Honkai: Star Rail Indonesian Language Reviews on Google Play Store Using Bidirectional Encoder Representations from Transformers Method," *Int. J. Eng. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 1–6, 2023, doi: 10.52088/ijesty.v3i3.462.

[8] Y. Afrillia, L. Rosnita, and D. Siska, "Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine," *G-Tech J. Teknol. Terap.*, vol. 6, no. 2, pp. 387–394, 2022, doi: 10.33379/gtech.v6i2.1778.

[9] R. Refianti and N. Anggraeni, "Sentiment Analysis Using Convolutional Neural Network Method to Classify Reviews on Zoom Cloud Meetings Application Based on Reviews on Google Playstore," vol. 3, no. 3, pp. 7–16, 2023.

[10] M. Qamal *et al.*, "ANALISIS SENTIMEN TOKO ONLINE MENGGUNAKAN dilakukan oleh Mehdi Mursalat Ismail dan Kemas Muslim Lhaksamana dengan judul ' Sen timen Analisis Pada Media Online Mengenai Pemilihan Presisen 2019 dengan Menggunakan Metode Naive Bayes '," no. 1.

[11] Munirul, Ula, M. M. Alvanof, and R. Triandi, "Analisa Dan Deteksi Konten Hoax Pada Media Berita," *J. Teknol. Terap. Sains 4.0 Univ. Malikussaleh*, vol. 1, p. 2, 2020.

[12] H. Hartono *et al.*, "A New Diversity Technique for Imbalance Learning Ensembles," *Int. J. Eng. Technol.*, 2018, doi: 10.14419/ijet.v7i2.11251.

[13] Y. Afrillia, L. Rosnita, and D. Siska, "Analisis Sentimen Pengguna Twitter Terhadap Isu Kesetaraan Gender Dalam Penerapan Permendikbudristek Nomor 30 Tahun 2021," *J. Informatics …*, vol. 8, no. 2, pp. 93–98, 2022.

[14] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, 1998, doi: 10.1023/A:1009715923555.

[15] S. Shevira, I. M. A. D. Suarjaya, and P. W. Buana, "Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia," *JITTER J. Ilm. Teknol. dan Komput.*, vol. 3, no. 2, p. 1074, 2022, doi: 10.24843/jtrti.2022.v03.i02.p06.

[16] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *Jdmsi*, vol. 2, no. 1, pp. 31–37, 2021.

[17] A. Apriani, H. Zakiyudin, and K. Marzuki, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta," *J. Bumigora Inf. Technol.*, vol. 3, no. 1, pp. 19–27, 2021, doi: 10.30812/bite.v3i1.1110.

[18] Y. Kardila, "Analisis Sentimen Review Pengguna Website IMDB Menggunakan Klasifikasi Naïve Bayes."

[19] F. S. Jumeilah, "Klasifikasi Opini Masyarakat Terhadap Jasa Ekspedisi JNE dengan Naïve Bayes," *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 92, 2018, doi: 10.21456/vol8iss1pp92-98.

[20] D. A. Agustina, S. Subanti, and E. Zukhronah, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine," *Indones. J. Appl. Stat.*, vol. 3, no. 2, p. 109, 2021, doi: 10.13057/ijas.v3i2.44337.