*Journal of Renewable Energy, Electrical, and Computer Engineering*
Volume 5, Number 1, March 2025. 18-29
eISSN 2776-0049
*Research Original Article*
DOI: https://doi.org/10.29103/jreece.v5i1.21342

# Comparison of Support Vector Machine and Naïve Bayes Algorithms in Sentiment Analysis of Tiktokshop Application User Reviews

**Agus Maula Rizki✉1, Bustami2, Said Fadlan Anshari3**
[1]Informatics Engineering, Faculty of Engineering, Universitas Malikussaleh, Bukit Indah, Lhokseumawe, 24353, Indonesia, agus.200170225@mhs.unimal.ac.id
[2]Informatics Engineering, Faculty of Engineering, Universitas Malikussaleh, Bukit Indah, Lhokseumawe, 24353, Indonesia, bustami@unimal.ac.id
[3]Informatics Engineering, Faculty of Engineering, Universitas Malikussaleh, Bukit Indah, Lhokseumawe, 24353, Indonesia, saidfadlan@unimal.ac.id
✉Corresponding Author: **agus.200170225@mhs.unimal.ac.id | Phone: +6282240666423**

**Received: December 13, 2024 Revision: January 15, 2025 Accepted: March 10, 2025**

## Abstract
This study presents a comparative analysis of Support Vector Machine (SVM) and Naïve Bayes algorithms for sentiment analysis of TikTokShop application user reviews. As TikTokShop emerges as an innovative platform integrating social media with e-commerce, understanding user sentiments becomes crucial for both consumers and businesses. A balanced dataset of 3,000 user reviews (1,000 positive, 1,000 neutral, and 1,000 negative) was collected through web scraping from Google Play Store. Following comprehensive preprocessing including cleansing, case folding, normalization, tokenization, stopword removal, and stemming, the data was vectorized using TF-IDF. Performance evaluation utilized accuracy, precision, recall, F1-score, confusion matrix, and 10-fold cross-validation. Results demonstrate that SVM consistently outperformed Naïve Bayes with higher accuracy (68.86% vs. 64.48%), precision (68.43% vs. 64.19%), and F1-score (68.58% vs. 62.46%). SVM exhibited balanced classification across all sentiment categories, while Naïve Bayes excelled at identifying negative sentiments (94.1% accuracy) but struggled significantly with neutral reviews (38.5%). Despite SVM's superior performance, Naïve Bayes demonstrated remarkable computational efficiency, with training time 224 times faster than SVM. The study reveals complementary strengths between the algorithms, suggesting potential value in ensemble approaches. These findings contribute to the understanding of sentiment analysis in video-based e-commerce platforms and provide valuable insights for businesses seeking to leverage user feedback for improved decision-making.
**Keywords:** Sentiment Analysis, Support Vector Machine, Naïve Bayes, TikTokShop, Text Classification

## Introduction
In recent years, online shopping has become an integral part of modern consumer behavior, driven by the increasing accessibility of e-commerce platforms (Silva, 2025). The shift from traditional shopping to digital transactions is not merely a trend but a fundamental transformation in the way people purchase products and services. This transition has been accelerated by advancements in mobile technology, widespread internet usage, and the integration of social media with online commerce (Ahmed et al., 2024). One of the most innovative platforms that has emerged from this evolution is TikTokShop, a feature within the TikTok application that seamlessly combines entertainment and shopping (Wahyuni et al., 2024).

TikTokShop distinguishes itself from conventional e-commerce platforms by integrating short-form videos and live streaming into the shopping experience (Bray, 2024). Unlike traditional marketplaces that rely on static product listings and text-based descriptions, TikTokShop allows sellers and influencers to showcase products in an engaging, real-time format (Chodak, 2024). This approach has proven to be highly effective, as users are more likely to trust and purchase products demonstrated by creators they follow. The combination of interactive content and impulse-driven purchasing has contributed to TikTokShop's rapid adoption, particularly among younger demographics who prefer dynamic and visually engaging shopping experiences (Hasibuan, 2024).

Despite its success, TikTokShop presents unique challenges, especially in ensuring consumer trust and maintaining the quality of product information (Septiani et al., 2024). User reviews play a critical role in influencing purchase decisions, yet the platform’s format differs from traditional e-commerce sites, where structured reviews are readily available. Instead, TikTokShop relies heavily on user-generated comments and live-stream interactions, making it difficult for buyers to filter through thousands of opinions to gauge a product’s reliability (Praneswara & Cahyono, 2023). Additionally, the viral nature of TikTok content can lead to misleading promotions, exaggerated claims, or even fraudulent activities, further complicating the decision-making process for consumers (Zhang, 2024). As a result, there is an urgent need for efficient methods to analyze and categorize user opinions, helping consumers make informed purchases while enabling sellers to refine their marketing strategies.

To address this issue, sentiment analysis has emerged as an essential tool in e-commerce, enabling automated classification of user opinions into positive, neutral, or negative sentiments. Sentiment analysis helps businesses understand consumer perceptions, improve product offerings, and enhance customer experience (Bharadwaj, 2023). Various machine learning techniques have been employed for this purpose, with Naïve Bayes (NB) and Support Vector Machine (SVM) being two of the most widely used classification algorithms.

Naïve Bayes, a probabilistic classifier based on Bayes' theorem, is known for its efficiency in handling large datasets with minimal computational cost (Boyko & Boksho, 2020). Meanwhile, SVM is a powerful algorithm that constructs optimal decision boundaries, making it highly effective for complex text classification tasks (Gupta & Rattan, 2023). Both algorithms have demonstrated strong performance in sentiment analysis, yet their effectiveness can vary depending on the characteristics of the dataset and the nature of the text being analyzed (Syahputra et al., 2022). Previous studies, such as those by (Rahmawati & Santoso, 2023) and (Muttaqin & Kharisudin, 2021), have shown that Naïve Bayes performs well in large-scale textual data, while SVM achieves higher accuracy in certain classification problems. However, the comparison between these algorithms in the specific context of TikTokShop user reviews remains underexplored.

Therefore, this study aims to compare the performance of Naïve Bayes and SVM in sentiment analysis of TikTokShop user reviews. By evaluating key performance metrics such as accuracy, precision, recall, and F1-score, this research seeks to determine which algorithm is better suited for analyzing user sentiments in a rapidly growing e-commerce environment. The results of this study are expected to provide valuable insights for e-commerce businesses, data analysts, and platform developers in improving user experience and decision-making based on consumer feedback.

## Literature Review

### TikTokShop as an Emerging E-commerce Platform

The integration of social media and e-commerce has given rise to innovative platforms like TikTokShop, which combines short-form video content with seamless shopping features(hu,2024). Unlike traditional e-commerce platforms that rely on static product listings, TikTokShop leverages video demonstrations and live streaming to create an engaging shopping experience. This approach has proven particularly effective among younger demographics, who value interactive and visually-driven content (Sabila & Andni, 2023).

TikTokShop's reliance on unstructured user-generated content, such as video comments and live-stream interactions, poses challenges for consumers evaluating product quality. Unlike structured reviews on traditional platforms, TikTokShop's informal feedback can be inconsistent and difficult to navigate. Additionally, the viral nature of TikTok content may lead to misleading promotions, further complicating consumer decision-making.

Despite these challenges, TikTokShop's rapid growth highlights its potential as a major player in the e-commerce industry. This study aims to address the need for effective sentiment analysis methods on TikTokShop, contributing to the understanding of user behavior on emerging e-commerce platforms.

### Sentiment Analysis in E-commerce

Sentiment analysis is a computational technique used to classify textual data, such as user reviews, into positive, neutral, or negative sentiments (Mehta & Pandya, 2020). In e-commerce, it helps businesses understand consumer opinions and improve decision-making. With the rise of social media-based platforms like TikTokShop, sentiment analysis has become essential for analyzing unstructured and informal user-generated content, such as comments and live-stream interactions (Huang et al., 2023).

Challenges such as slang, mixed sentiments, and the dynamic nature of social media content can reduce the accuracy of sentiment analysis models. Despite these limitations, sentiment analysis remains a valuable tool for enhancing customer experience and refining marketing strategies in e-commerce.

### Crawling

Web crawling is a technique used to collect large-scale textual data from online sources, such as e-commerce platforms or app stores (Kunekar et al., 2024). In this study, data crawling is employed to gather TikTokShop user reviews from the Google Play Store. Tools like Python libraries are commonly used for scraping unstructured data, including reviews, ratings, and timestamps.

Crawling data from platforms like TikTokShop presents challenges, such as API limitations, dynamic content loading, and anti-scraping mechanisms. Additionally, the unstructured nature of user-generated content requires preprocessing to ensure data quality before analysis.

### Text Preprocessing

Text preprocessing is a fundamental step in sentiment analysis to prepare raw textual data for machine learning models. This process involves transforming unstructured text into a clean and standardized format, which improves the accuracy and efficiency of sentiment classification (Kaur & Sharma, 2023). For TikTokShop user reviews, preprocessing is particularly crucial due to the informal nature of the data, which often includes slang, abbreviations, emojis, and mixed languages (Sarina & Tanniewa, 2023). The preprocessing steps typically include :

1. Cleansing, This step involves removing irrelevant characters, such as punctuation, numbers, and special symbols, which do not contribute to sentiment analysis. Additionally, HTML tags, URLs, and mentions (@username) are eliminated to ensure the text contains only meaningful words (Kaur & Sharma, 2023).
2. Case Folding, All text is converted to lowercase to ensure uniformity. This step prevents the algorithm from treating the same word differently due to variations in capitalization (e.g., "Good" vs. "good").

3. Stopword Removal, Common words that do not carry significant meaning, such as "the," "and," or "is," are removed. This step reduces noise and focuses the analysis on words that contribute to sentiment (Kaur & Sharma, 2023).
4. Tokenization, The text is split into individual words or tokens. This step breaks down sentences into smaller units, making it easier for the algorithm to process and analyze each word separately.
5. Stemming, Words are reduced to their root forms using stemming algorithms like Porter or Snowball. For example, "running" becomes "run," and "happiness" becomes "happi." While stemming may not always produce linguistically accurate results, it helps group similar words together, reducing the dimensionality of the data (Gupta & Rattan, 2023).
6. Normalization, For TikTokShop reviews, additional preprocessing may be required to handle informal language, such as converting slang ("u" to "you") or expanding contractions ("can't" to "cannot"). Emojis and emoticons can also be translated into their textual equivalents to capture sentiment (Kaur & Sharma, 2023).

**Feature Extraction using TF-IDF (Term Frequency – Invers Document Frequency)**

Feature extraction transforms textual data into numerical representations for machine learning models. Term Frequency-Inverse Document Frequency (TF-IDF) is a widely used method that evaluates word importance by combining Term Frequency (TF) (how often a word appears in a document) and Inverse Document Frequency (IDF) (how rare a word is across documents) (Kabra & Nagar, 2023). The formula is :

$$\mathrm{TF}-\mathrm{IDF}\,(t,d) = TF(t,d)x\,log\frac{N}{1+DF(t)} \qquad (1)$$

TF-IDF assigns higher weights to unique, sentiment-bearing words (e.g., "excellent," "poor") while reducing the importance of common words (e.g., "the," "and"). This makes it more effective than Bag of Words (BoW), which treats all words equally (Premasudha & Rampalli, 2024). In this study, TF-IDF is applied to preprocessed TikTokShop user reviews to create numerical vectors for sentiment classification using Naïve Bayes and SVM.

**Support Vector Machine (SVM)**

Support Vector Machine (SVM) is a supervised machine learning algorithm widely used for classification tasks, including sentiment analysis (Bustami & Aryani, 2023). SVM works by finding the optimal hyperplane that separates data points of different classes with the maximum margin . In text classification, SVM maps textual data into a high-dimensional space using techniques like TF-IDF and then constructs a hyperplane to distinguish between positive, neutral, and negative sentiments (Abdullah & Abdulazeez, 2021).

SVM is particularly effective for handling high-dimensional data, such as text vectors, and performs well with small to medium-sized datasets. However, its computational complexity increases with larger datasets, and it may struggle with noisy data. Despite these limitations, SVM has demonstrated high accuracy in sentiment analysis tasks, making it a popular choice for e-commerce applications (Khurana & Verma, 2022). In this study, SVM is employed to classify sentiments in TikTokShop user reviews, leveraging its ability to handle complex decision boundaries and high-dimensional feature spaces.

$$f(x) = w.x + b \qquad (2)$$

**Naïve Bayes Algorithm**

Naïve Bayes is a probabilistic classification algorithm based on Bayes' Theorem. It is widely used in sentiment analysis due to its simplicity, efficiency, and ability to handle large datasets (Boyko & Boksho, 2020). The algorithm calculates the probability of a text belonging to a specific class (e.g., positive, neutral, negative) given its features (words). The formula for Bayes' Theorem is :

$$P(A|B) = \frac{P(B|A).\,P(A)}{P(B)} \qquad (3)$$

Naïve Bayes is particularly effective for text classification tasks because it handles high-dimensional data well and requires minimal computational resources (Suwanda et al., 2024).

**Confussion Matrix**

A confusion matrix is a performance evaluation tool used in classification tasks to visualize the accuracy of a model. It provides a detailed breakdown of the model's predictions by comparing them to the actual labels (Rustam et al., 2021). For a binary classification problem (e.g., positive vs. negative sentiment), the confusion matrix consists of four components :

- True Positive (TP): The number of correctly predicted positive instances.
- True Negative (TN): The number of correctly predicted negative instances.
- False Positive (FP): The number of negative instances incorrectly predicted as positive.
- False Negative (FN): The number of positive instances incorrectly predicted as negative.

The confusion matrix is particularly useful for calculating key performance metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive evaluation of the model's performance, especially in imbalanced datasets where one class dominates the other (Korkmaz, 2020).

1. Accuracy, Proportion of correct predictions out of total predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

2. Precision, The proportion of correct positive predictions out of total positive predictions, measures how precise the algorithm is in identifying positive classes.

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

3. Recall, The proportion of positive cases identified measures how complete the algorithm is in identifying the positive class.

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

4. F1-Score, The harmonic mean of precision and recall, provides a balance between the two metrics.

$$F1 - Score = 2 \times \frac{Precission \times Recall}{Precission + Recall} \qquad (7)$$

**K-Fold Cross Validation**

K-Fold Cross-Validation is a robust technique used to evaluate the performance of machine learning models, particularly in scenarios with limited data. It works by dividing the dataset into k$k$ equal-sized subsets (folds). The model is trained $k$ times, each time using $k-1$ folds for training and the remaining fold for validation. This process ensures that every data point is used for both training and validation, providing a more reliable estimate of the model's performance (Elkari et al., 2024).

The performance metrics (e.g., accuracy, precision, recall) from each fold are averaged to produce a final evaluation score. This approach reduces the risk of overfitting and ensures that the model generalizes well to unseen data (Yacob et al., 2024). For example, in sentiment analysis of TikTokShop user reviews, 10-fold cross-validation can be used to assess the consistency of Naïve Bayes and SVM across different subsets of the data.

**Previous Research**

Sentiment analysis has been widely applied in various contexts, including e-commerce and service-based applications. For example, (Rahmawati & Santoso, 2023) implemented the Naïve Bayes method to classify reviews of the Tokopedia e-commerce application on Google Playstore. Using a dataset of 5.000 reviews, their study achieved an accuracy of 83.9%, with precision, recall, and F1-score values of 85.1%, 83.9%, and 83.8%, respectively. This demonstrates the effectiveness of Naïve Bayes in handling large-scale textual data, particularly in structured e-commerce environments.

Similarly, (Muttaqin & Kharisudin, 2021) compared the performance of Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) for sentiment analysis of Gojek application reviews on Google Playstore. Their results showed that SVM with a linear kernel achieved higher accuracy (87.98%) compared to KNN (82.14%), along with better precision (88.55%) and recall (95.43%). This highlights SVM's strength in handling complex text classification tasks, especially in service-based applications.

These studies primarily focus on traditional e-commerce platforms like Tokopedia or service-based apps like Gojek, where reviews are text-based and relatively structured. There is limited research on emerging platforms like TikTokShop, which rely heavily on unstructured, user-generated content such as video comments and live-stream interactions. This gap is significant, as TikTokShop's unique format introduces new challenges for sentiment analysis, including the prevalence of slang, emojis, and mixed languages.

Additionally, while Naïve Bayes and SVM have been compared in various contexts, their performance in analyzing TikTokShop user reviews remains underexplored. This study addresses this gap by comparing the effectiveness of these algorithms in classifying TikTokShop reviews, with a focus on accuracy, precision, recall, and F1-score. The findings are expected to provide valuable insights for e-commerce businesses and platform developers in leveraging sentiment analysis to enhance user experience and decision-making.

## Materials & Methods

In this research, a comparative analysis between Support Vector Machine (SVM) and Naïve Bayes (NB) algorithms is conducted in classifying the sentiment of user reviews on the TikTokShop platform. The method used includes several main stages, from data collection and labelling, text pre-processing, machine learning model building, to model performance evaluation.

These stages are designed to ensure that the data used has gone through a standardisation process, so that the sentiment classification results can be interpreted more accurately. The research flowchart used in this study can be seen in Figure below.

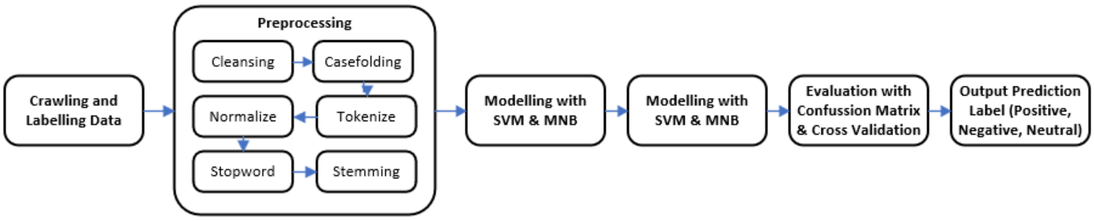


**Figure 1.** Schematic of Research

After understanding the general flow of the research, this section will describe in detail each stage carried out in this research, starting from data collection methods to model evaluation.

**Data Collection**

The dataset in this study was obtained through web scraping technique on user reviews of TikTokShop application on Google Play Store. A total of 3,000 reviews were collected, consisting of 1,000 positive, 1,000 neutral, and 1,000 negative reviews. Labelling was done manually based on review scores and text characteristic analysis, with reference to user expression patterns. The classified data was then used for the preprocessing stage before being fed into the machine learning model for sentiment analysis.

To ensure a balanced distribution of data, the collected reviews cover various aspects such as transaction experience, product quality, as well as delivery service. The data obtained was then saved in CSV format and used as input in the further processing stage. With a proportional composition of data in each sentiment category, this research aims to produce a model that is able to classify sentiment with optimal accuracy.

**Table 1**. Sample of Datasets

| User Review | Label | Amount of Data |
|---|---|---|
| "Sangat membantu kami mempromosikan produk, semoga ini jadi jalan rizki halal buat kami. aamiin" | Positive | 1000 Data |
| "Saldo saya di tahan Tampa alasan.. status saldo sudah ( / ) harus nya hari itu juga cair lah ini status ( / ) sudah 1 Minggu tidak di cair"" sudah ajukan tiket jawabannya itu itu aja .. emang gak niat balikin saldo ku .. dasar maling tik tok maling" | Negative | 1000 data |
| "Masih belajar. Tapi belum ngerti" | Neutral | 1000 Data |

**Data Preprocessing**

Once the data is collected, a preprocessing stage is performed to clean the text from irrelevant elements and create a more structured format for machine learning models to process. This process includes several main stages, starting with cleansing, which is the removal of special characters, punctuation marks, numbers, and emojis that have no contribution to sentiment analysis. Next, case folding is performed by converting the entire text to lowercase to equalise the format and avoid inconsistencies in the data.

In the next stage, normalisation is performed to convert nonstandard words and abbreviations into standard forms, for example the word 'gpp' is converted into 'it's okay'. After normalisation, the text is broken down into individual words through tokenisation, which aims to facilitate word-by-word analysis. Next, words that have no significant meaning in the analysis, such as 'in', 'to', and 'which', are removed using a stopword removal technique. The last process is stemming, which converts words into their base form using the Sastrawi library, for example the word 'buy' becomes 'buy'.

After the preprocessing process is applied, the amount of data is slightly reduced due to the removal of blank text or reviews that after processing no longer have meaningful words. From the initial 3,000 reviews, the usable data was reduced to 2,970 reviews, with a distribution of 990 positive, 990 neutral, and 990 negative reviews.

**Table 2.** Preprocessing Process

| Before Preprocessing | After Preprocessing |
|---|---|
| "Sangat membantu kami mempromosikan produk, semoga ini jadi jalan rizki halal buat kami. aamiin" | "sangat bantu promosi produk semoga jalan rezeki halal" |
| "Saldo saya di tahan Tampa alasan.. status saldo sudah ( / ) harus nya hari itu juga cair lah ini status ( / ) sudah 1 Minggu tidak di cair"" sudah ajukan tiket jawabannya itu itu aja .. emang gak niat balikin saldo ku .. dasar maling tik tok maling" | "saldo tahan tanpa alasan status saldo harus cair status minggu tidak cair ajukan tiket jawaban itu niat balikin saldo dasar maling tiktok maling" |
| "Masih belajar. Tapi belum ngerti" | "masih belajar belum ngerti" |

**Feature Extraction with TF-IDF (Term Frequency – Invers Document Frequency)**

After going through the preprocessing stage, the cleaned text is converted into numerical representation so that it can be processed by machine learning models. In this research, the Term Frequency-Inverse Document Frequency (TF-IDF) method is used to extract features from user review texts. TF-IDF calculates the importance of a word in a document based on how often the word appears (Term Frequency) and how rarely the word appears throughout the document (Inverse Document Frequency). With this approach, words that are more informative in determining sentiment will have a higher weight than common words that appear frequently in all reviews.

In this study, TF-IDF was applied to 2,970 processed reviews, resulting in a feature matrix with dimensions corresponding to the number of unique words in the dataset. TF-IDF was applied using the scikit-learn library, with parameters adjusted to ensure that only words with high significance are retained in the modelling. This process produces a numerical vector that will be used as input for the Naïve Bayes and SVM algorithms in the sentiment classification stage.

**Modelling**

After the features are extracted using TF-IDF, the next step is to build a classification model using Naïve Bayes and Support Vector Machine (SVM). These two algorithms were chosen because they have different approaches in handling text data, thus allowing performance comparison in sentiment analysis of TikTokShop user reviews.

1. Support Vector Machine (SVM)

   SVM is a margin-based algorithm that aims to find the best separator between sentiment categories. In this research, Linear SVM is used, as it is able to handle high-dimensional text data generated from TF-IDF representation. The model works by mapping the data into a feature space and determining the optimal dividing line that maximises the margin between sentiment classes.

   The model training is done by dividing the data into 80% for training and 20% for testing. Preprocessed data that has been converted using TF-IDF is used as input to build the classification model. SVM has the advantage of handling complex data and tends to produce more accurate classifications than probabilistic-based algorithms. However, SVM requires higher computation time, especially when handling large datasets.

2. Naïve Bayes

   Naïve Bayes is a probability-based algorithm that classifies text based on the distribution of words in the dataset. In this research, Multinomial Naïve Bayes is used, which is commonly applied in text classification with frequency-based features such as TF-IDF. This model assumes that each word in the text is independent, thus allowing the calculation of sentiment class probabilities based on the occurrence of certain words.

   As in SVM, training is done with 80% of the data for training and 20% for testing. This model has advantages in computational efficiency and good performance on large datasets, especially for texts with a balanced distribution of words. However, the assumption of independence between words can be a limitation, especially when there is a strong correlation between words in a review.

**Evaluation**

After the model was developed using SVM and Naïve Bayes, an evaluation was conducted to measure its performance in classifying the sentiment of TikTokShop user reviews. This evaluation aims to understand the effectiveness of each algorithm in recognising sentiment patterns as well as identifying potential model weaknesses. The two main methods used in this research are Confusion Matrix and K-Fold Cross Validation, which are complemented by several additional analyses to provide deeper insights into the model's prediction results.

1. Confussion Matrix

   Confusion Matrix is used to compare the model predictions with the original labels in the test dataset. This matrix provides an overview of the number of correct and incorrect predictions in the positive, neutral, and negative categories. From these results, evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to assess the effectiveness of the model in distinguishing review sentiment. To clarify the analysis, the confusion matrix results are visualised in the form of heatmaps, as well as bar charts comparing the distribution of actual data with the model's predicted results.

2. K-Fold Cross Validation

   Further evaluation is done with K-Fold Cross Validation to reduce bias in model testing. In this study, 10-Fold Cross Validation is used, where the dataset is divided into 10 parts that are used alternately as training and test data. This technique ensures that the model is tested thoroughly and does not rely solely on one particular subset of data. The results of the cross validation are visualised using a line chart that shows the variation in accuracy at each fold, as well as a boxplot that illustrates the distribution of the model's performance throughout the validation process.

In addition, some additional analyses were conducted to understand the characteristics of the model more deeply. The Confidence Score distribution was visualised in the form of a histogram to see the confidence level of the model in its predictions. Error Case Analysis was used to identify patterns of errors in classification. ROC Analysis displays the ROC curve and AUC value as an indicator of the balance between True Positive Rate and False Positive Rate. Comparison of actual and predicted data distributions is visualised with bar charts or pie charts, while Word Cloud is used to show the most dominant words in each sentiment category. This evaluation provides a comprehensive overview of the effectiveness of SVM and Naïve Bayes in sentiment analysis of TikTokShop reviews.

## Results and Discussion

This chapter presents the evaluation results of Support Vector Machine (SVM) and Naïve Bayes models in classifying TikTokShop review sentiments. The analysis was conducted using Confusion Matrix as well as accuracy, precision, recall, and F1-score metrics to assess the effectiveness of both algorithms. In addition, 10-Fold Cross Validation was conducted to measure the stability of the model on various subsets of data.

To further understand the model performance, additional analyses included Confidence Score Distribution, Error Case Analysis, and ROC Analysis to evaluate the balance between True Positive Rate and False Positive Rate. In addition, data distribution analysis compares the classification results with the original labels, and Word Cloud Analysis is used to identify the dominant words in each sentiment category.

The final section of this chapter discusses the performance comparison of the two algorithms as well as the implications of the research results in the context of sentiment analysis on video-based e-commerce platforms. The discussion begins with an evaluation of the classification performance of each model.

### Classification Model Performance

**Table 4**. Performance of the Algorithm SVM & Naive Bayes

| Metrics | Support Vector Machine | Naive Bayes |
|---|---|---|
| Accuracy | 68.86% | 64.48% |
| Precision | 68.43% | 64.19% |
| Recall | 68.86% | 64.48% |
| F1-Score | 68.58% | 62.46% |
| Training Time | 156.83 sec | 0.70 sec |
| Prediction Time | 6.72 sec | 0.07 sec |

SVM achieved 68.86% accuracy in classifying TikTokShop review sentiments, with balanced precision (68.43%) and recall (68.86%). The model performed exceptionally well for negative sentiment detection (84% F1-score) but showed moderate performance for neutral and positive sentiments (62% and 59% F1-scores respectively). The confusion matrix reveals that while SVM excelled at identifying negative reviews, it had more difficulty distinguishing between neutral and positive categories. Despite its superior performance, SVM required significant computational resources with a training time of 156.83 seconds.

Naïve Bayes achieved 64.48% accuracy, with similar precision and recall values. The model demonstrated remarkable performance in detecting negative sentiments (94% recall) but struggled significantly with neutral reviews (only 38% recall). The confusion matrix shows that neutral reviews were frequently misclassified as either negative or positive. However, Naïve Bayes proved extremely efficient computationally, with a training time of just 0.70 seconds—224 times faster than SVM.

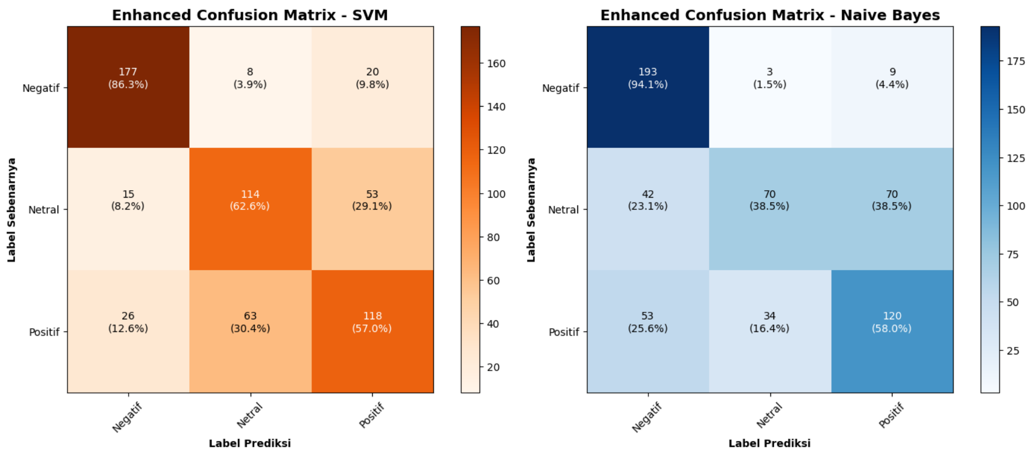


**Figure 2.** Confussion Matrix

The confusion matrix visualisation shows different classification patterns between the two algorithms. SVM shows a more balanced prediction distribution with good classification ability in all categories, especially in negative sentiment (86.3%) and moderately good in positive (57.0%) and neutral sentiment (62.6%). Meanwhile, Naïve Bayes showed a strong tendency to identify negative sentiments with a very high success rate (94.1%), but had significant difficulty in recognising neutral sentiments (only 38.5% correctly classified). An interesting pattern of misclassification was seen in the Naïve Bayes model, where 70 neutral reviews (38.5%) were misclassified as positive, demonstrating the challenge in distinguishing the nuances between these two categories. The SVM also experienced similar but more moderate difficulties, with 53 neutral reviews (29.1%) misclassified as positive. This comparison indicates that while Naïve Bayes excelled at identifying negative sentiment, SVM offered a more balanced and accurate classification overall, especially for neutral sentiment which proved the most challenging to classify correctly.

**K-Fold Cross Validation Analysis**

To test the consistency and reliability of the two models in sentiment classification, a 10-fold cross validation method was used. This method divides the dataset into 10 parts that are used alternately as testing data, thus providing a more comprehensive picture of performance.

**Table 5**. Performance Results on Each Fold

| Fold | Support Vector Machine (SVM) | | | | Naive Bayes | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy | Precision | Recall | F1-Score | Accuracy | Precision | Recall | F1-Score |
| 1 | 0.673 | 0.668 | 0.673 | 0.670 | 0.633 | 0.645 | 0.633 | 0.602 |
| 2 | 0.714 | 0.718 | 0.714 | 0.710 | 0.670 | 0.668 | 0.670 | 0.655 |
| 3 | 0.694 | 0.688 | 0.694 | 0.690 | 0.653 | 0.673 | 0.653 | 0.629 |
| 4 | 0.714 | 0.709 | 0.714 | 0.710 | 0.677 | 0.697 | 0.677 | 0.651 |
| 5 | 0.697 | 0.693 | 0.697 | 0.694 | 0.660 | 0.672 | 0.660 | 0.640 |
| 6 | 0.657 | 0.654 | 0.657 | 0.655 | 0.633 | 0.638 | 0.633 | 0.614 |
| 7 | 0.653 | 0.653 | 0.653 | 0.652 | 0.660 | 0.655 | 0.660 | 0.644 |
| 8 | 0.741 | 0.742 | 0.741 | 0.739 | 0.714 | 0.715 | 0.714 | 0.701 |
| 9 | 0.727 | 0.728 | 0.727 | 0.724 | 0.650 | 0.638 | 0.650 | 0.632 |
| 10 | 0.667 | 0.664 | 0.667 | 0.664 | 0.589 | 0.637 | 0.589 | 0.568 |
| **Avrg** | **0.654** | **0.664** | **0.654** | **0.634** | **0.694** | **0.692** | **0.694** | **0.691** |

The 10-fold cross validation evaluation results show that SVM consistently outperforms Naïve Bayes on all performance metrics. SVM achieved an average accuracy of 69.4%, while Naïve Bayes achieved 65.4%. The performance difference was also significant in F1-score, where SVM achieved an average of 69.1% compared to 63.4% for Naïve Bayes. In addition, both models showed variation in performance between folds, with SVM achieving the highest performance on the 8th fold (accuracy 74.1%, F1-score 73.9%) and Naïve Bayes also showing the best performance on the same fold (accuracy 71.4%, F1-score 70.1%). However, Naïve Bayes showed a significant drop in performance at the 10th fold with an accuracy of only 58.9%. The stability of SVM was better with smaller performance deviations between the highest and lowest folds, indicating better generalisation ability to variations in the data. These results reinforce the conclusion that SVM is more reliable for sentiment classification of TikTokShop reviews despite requiring longer computation time.

**Confidence Score Distribution Analysis**

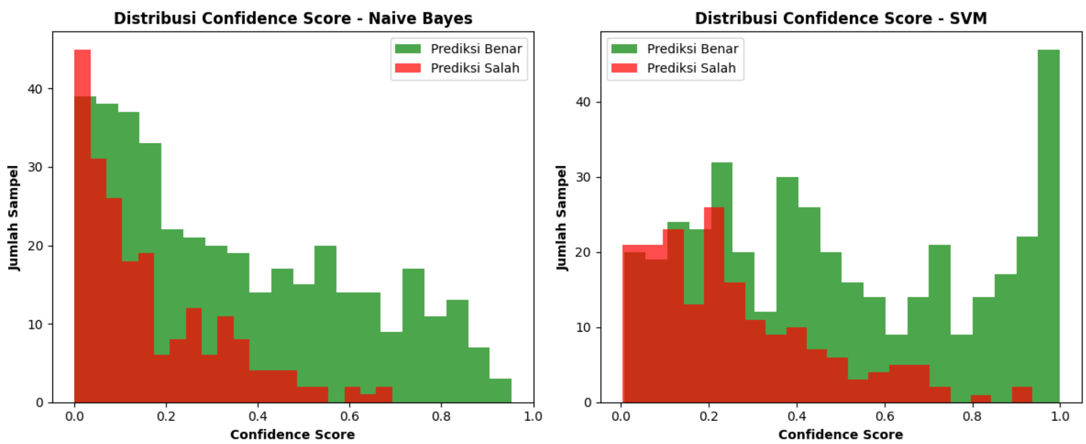


**Figure 3.** Confidence Score Distribution

The confidence score analysis reveals that SVM exhibits higher confidence in its predictions compared to Naïve Bayes. SVM achieves an average confidence of 0.5012 for correct predictions and 0.2578 for incorrect predictions, while Naïve Bayes records 0.3470 for correct predictions and 0.1642 for incorrect predictions. Visualization of the distribution shows that SVM's correct predictions are concentrated at high confidence scores (close to 1.0), whereas Naïve Bayes's correct predictions are spread across low to high scores, with a higher density in the lower range (0.0-0.3).

This pattern indicates that SVM is more "confident" in correct predictions, while Naïve Bayes shows lower confidence even for correct predictions. These findings align with SVM's higher accuracy and suggest that an optimal confidence score threshold could enhance model precision, particularly for SVM.

**Error Case Analysis**

The error case analysis highlights distinct patterns in the misclassification behavior of the Support Vector Machine (SVM) and Naïve Bayes (NB) models. Specifically, there were 37 cases where Naïve Bayes correctly classified the sentiment, but SVM made errors, and 63 cases where SVM was correct, but Naïve Bayes failed. This indicates that while

SVM generally performs better overall, Naïve Bayes occasionally identifies certain cases more accurately, particularly in specific sentiment categories.

**Table 5**. Error Case Analysis

| SVM Error Pattern when Naive Bayes is Correct | | Naïve Bayes Error Pattern when SVM is Correct | |
|---|---|---|---|
| **Error** | **Total Cases** | **Error** | **Total Cases** |
| Positive → Neutral | 15 cases | Neutral → Positive | 27 cases |
| Negative → Positive | 10 cases | Neutral → Negative | 20 cases |
| Negative → Neutral | 6 cases | Positive → Negative | 13 cases |
| Neutral → Positive | 3 cases | Positive → Neutral | 3 cases |
| Positive → Negative | 3 cases | Negative → Neutral | 0 cases |
| Neutral → Negative | 0 cases | Negative → Postive | 0 cases |

SVM's misclassifications primarily occurred when Naïve Bayes was correct. For instance, SVM struggled to distinguish between positive and neutral sentiments, mislabeling 15 positive reviews as neutral. Additionally, SVM incorrectly classified 10 negative reviews as positive, suggesting difficulty in detecting strongly negative language. There were also 6 cases where negative reviews were misclassified as neutral, indicating challenges in identifying subtle negative cues. In a smaller number of cases, SVM misclassified neutral reviews as positive (3 cases) and positive reviews as negative (3 cases), likely due to ambiguous or mixed language. Notably, SVM did not misclassify any neutral reviews as negative, demonstrating relative strength in this area.

On the other hand, Naïve Bayes exhibited significant challenges when SVM was correct. The most frequent errors occurred with neutral reviews, where 27 neutral reviews were misclassified as positive and 20 as negative. This suggests that Naïve Bayes struggles to capture the nuances of neutral sentiment, often misinterpreting it as either positive or negative. Additionally, 13 positive reviews were misclassified as negative, further highlighting the model's difficulty in handling ambiguous or context-dependent language. Interestingly, Naïve Bayes did not misclassify any negative reviews as positive or neutral, indicating a stronger ability to identify negative sentiments compared to neutral or positive ones.

These findings underscore the strengths and weaknesses of both models. SVM demonstrates a more balanced performance but still faces challenges in distinguishing between positive and neutral sentiments, as well as identifying subtle negative cues. Naïve Bayes, while efficient and occasionally more accurate in specific cases, struggles significantly with neutral and ambiguous reviews. This analysis suggests that improving the models' ability to handle neutral and mixed sentiments, as well as incorporating contextual understanding, could significantly enhance their overall performance.

**ROC Analysis**

The ROC (Receiver Operating Characteristic) analysis provides a detailed evaluation of the performance of both the Support Vector Machine (SVM) and Naïve Bayes (NB) models in classifying TikTokShop review sentiments. The ROC curves and their corresponding AUC (Area Under the Curve) values are visualized in the accompanying graph, offering insights into the models' ability to distinguish between negative, neutral, and positive sentiments.

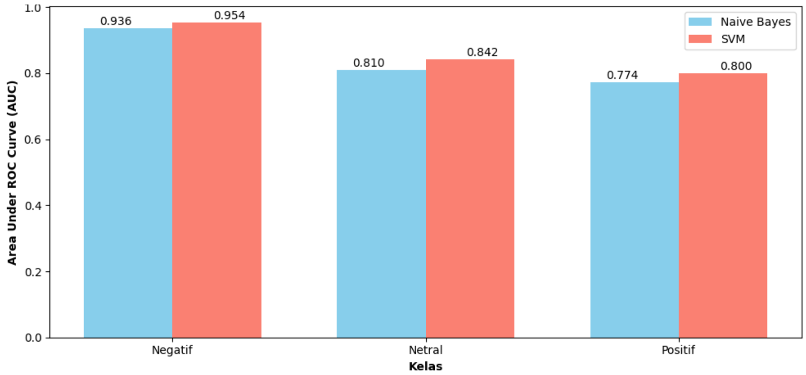


**Figure 4.** ROC Distribution

From the graph, it is evident that SVM consistently outperforms Naïve Bayes across all sentiment categories, as indicated by its higher AUC values. For negative sentiment, SVM achieves an AUC of 0.9644 compared to Naïve Bayes' 0.9369, demonstrating its superior ability to differentiate negative reviews from others. In the case of neutral sentiment, SVM maintains a stronger performance with an AUC of 0.8421, while Naïve Bayes lags slightly behind at 0.8104. Similarly, for positive sentiment, SVM's AUC of 0.7995 surpasses Naïve Bayes' 0.7736, though the margin is narrower. These results align with the overall trend observed in previous analyses, where SVM exhibits greater robustness and accuracy in sentiment classification. The ROC curves further highlight the challenges both models face in classifying neutral and positive sentiments, as reflected in their relatively lower AUC values compared to negative sentiment. This analysis underscores the importance of refining the models' ability to handle nuanced and ambiguous sentiments, particularly in the neutral and positive categories.

**Analysis Distribution of Actual and Predicted Sentiments**

The distributional analysis compares the actual sentiment distribution in the testing data with the predicted distributions generated by the Naïve Bayes (NB) and Support Vector Machine (SVM) models. This analysis helps evaluate how well the models align with the true sentiment distribution and identifies potential biases or discrepancies in their predictions.

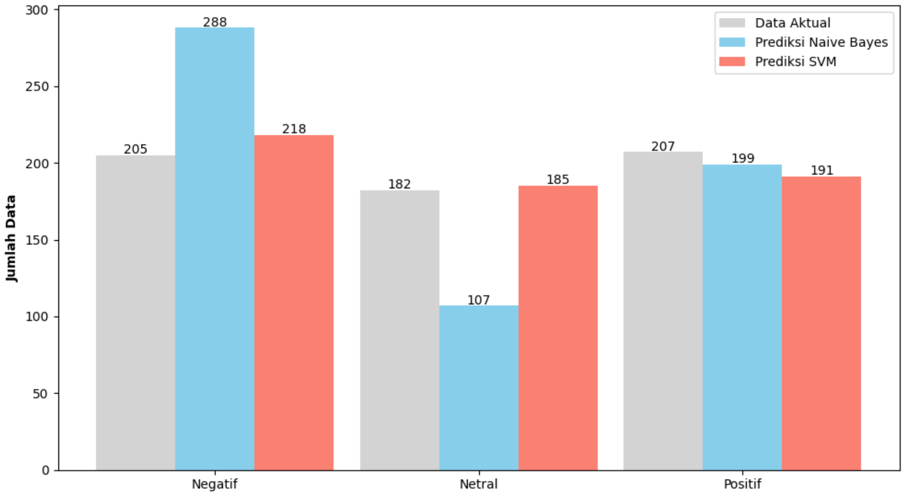

**Figure 5.** Distribution of Actual and Predicted Label

The actual distribution of the testing data shows a relatively balanced spread across the three sentiment categories, with negative sentiments comprising 34.5% (205 data points), neutral sentiments at 30.6% (182 data points), and positive sentiments at 34.8% (207 data points). This balance provides a fair basis for evaluating the models' performance.

In contrast, the predicted distribution by Naïve Bayes reveals a significant bias toward negative sentiments, with 48.5% (288 predictions) of its outputs classified as negative. This over-prediction of negative sentiment comes at the expense of neutral sentiment, which is underrepresented at only 18.0% (107 predictions). This aligns with earlier findings that Naïve Bayes struggles with neutral reviews, often misclassifying them as negative or positive. Positive sentiments are predicted at 33.5% (199 predictions), which is close to the actual distribution but still reflects some imbalance.

On the other hand, SVM's predictions are more balanced and closely aligned with the actual distribution. Negative sentiments are predicted at 36.7% (218 predictions), neutral sentiments at 31.1% (185 predictions), and positive sentiments at 32.2% (191 predictions). While SVM also slightly over-predicts negative sentiment, it does so to a much lesser extent than Naïve Bayes and maintains a better representation of neutral and positive sentiments. This suggests that SVM is more effective at capturing the true distribution of sentiments in the data.

Overall, the distributional analysis highlights the strengths and weaknesses of both models. Naïve Bayes tends to over-predict negative sentiments and under-predict neutral ones, indicating a potential bias in its classification approach. SVM, while not perfect, demonstrates a more balanced and accurate alignment with the actual sentiment distribution, further reinforcing its superior performance in this sentiment analysis task.

## Conclusions

1. This study conducted a comprehensive comparison between Support Vector Machine (SVM) and Naïve Bayes algorithms for sentiment analysis of TikTokShop user reviews. The evaluation using multiple performance metrics reveals several significant findings that contribute to the understanding of sentiment analysis in video-based e-commerce platforms.
2. SVM consistently outperformed Naïve Bayes across all evaluation metrics, achieving higher accuracy (68.86% vs. 64.48%), precision (68.43% vs. 64.19%), recall (68.86% vs. 64.48%), and F1-score (68.58% vs. 62.46%). The 10-fold cross-validation further confirmed SVM's superior performance, demonstrating its robustness and reliability in handling the nuanced language patterns found in TikTokShop reviews.
3. One of the most striking differences between the two algorithms lies in their classification patterns. SVM exhibited a more balanced performance across all sentiment categories, while Naïve Bayes showed exceptional strength in identifying negative sentiments (94.1% accuracy) but struggled significantly with neutral reviews (only 38.5% correctly classified). This imbalance is further evidenced in the distributional analysis, where Naïve Bayes demonstrated a bias toward negative sentiment classification at the expense of neutral sentiment recognition.
4. The confidence score analysis revealed that SVM not only makes more accurate predictions but also does so with higher confidence. This suggests that SVM's decision boundaries are more effectively capturing the underlying sentiment patterns in the feature space. The ROC analysis further confirmed this advantage, with SVM achieving higher AUC values across all sentiment categories, particularly for negative sentiment (0.9644 vs. 0.9369).
5. Despite SVM's superior performance, it required significantly more computational resources, with a training time 224 times longer than Naïve Bayes (156.83 seconds vs. 0.70 seconds). This computational efficiency gives Naïve Bayes a practical advantage in scenarios where processing speed is prioritized over classification accuracy,

particularly for real-time applications or platforms with limited computational resources.

6. Error case analysis revealed complementary strengths between the two algorithms. SVM struggled primarily with distinguishing between positive and neutral sentiments, while Naïve Bayes showed remarkable accuracy in identifying negative sentiments but frequently misclassified neutral reviews. This suggests potential value in ensemble approaches that leverage the strengths of both algorithms.
7. For e-commerce platforms like TikTokShop that rely heavily on user-generated content to drive purchasing decisions, these findings have significant implications. The ability to accurately classify sentiment, particularly neutral sentiment which proved challenging for both algorithms, is crucial for businesses to understand consumer perceptions and refine their offerings accordingly. The balanced performance of SVM makes it more suitable for comprehensive sentiment analysis, while Naïve Bayes might be preferable for specific applications focused on negative sentiment detection where computational efficiency is essential.
8. Future research should focus on enhancing the models' ability to recognize neutral sentiment, which proved to be the most challenging category for both algorithms. Additionally, exploring ensemble methods that combine the strengths of SVM and Naïve Bayes could potentially improve overall sentiment classification performance. Incorporating domain-specific features and context-aware preprocessing techniques could also address the challenges posed by the informal and dynamic nature of user reviews on video-based e-commerce platforms like TikTokShop.

In conclusion, while SVM demonstrates superior overall performance for sentiment analysis of TikTokShop user reviews, the choice between SVM and Naïve Bayes should be guided by specific application requirements, balancing classification accuracy against computational efficiency.

## Acknowledgments

This research journey was made possible through the collective support of numerous individuals who contributed their expertise, time, and encouragement. I extend my deepest appreciation to my thesis advisors who provided exceptional mentorship, critical insights, and steadfast encouragement throughout this investigative process. Their expertise in machine learning and sentiment analysis fundamentally elevated this work and deepened my understanding of computational linguistics.

My sincere thanks go to my academic colleagues in the informatics engineering department at Universitas Malikussaleh who offered valuable technical discussions, collaborative troubleshooting sessions, and consistent moral support. The intellectual exchanges we shared proved instrumental in refining the methodological approach and enhancing the interpretation of experimental results.

I wish to acknowledge the contributions of friends who generously devoted time to review drafts, suggest improvements, and provide fresh analytical perspectives. Their enthusiastic interest in my research and continued encouragement maintained my momentum through challenging phases of this study.

I reserve my most profound gratitude for my family members, whose patience, understanding, and emotional backing formed the foundation of my perseverance. Their unshakeable confidence in my abilities and constant reinforcement served as a powerful motivating force, particularly during periods of difficulty. This academic accomplishment stands as a testament to their unconditional support and belief in my potential.

## References

Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification a review. *Qubahan Academic Journal*, *1*(2), 81–90.

Ahmed, A., Raj, A., Goel, P. K., & Mahur, L. S. (2024). Social Media Integration with E-Commerce Using AI & Augmented Reality. *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0*, 1–5.

Bharadwaj, L. (2023). Sentiment analysis in online product reviews: mining customer opinions for sentiment classification. *Int J Multidiscip Res*, *5*(5).

Boyko, N., & Boksho, K. (2020). Application of the Naive Bayesian Classifier in Work on Sentimental Analysis of Medical Data. *IDDM*, *2020*, 3rd.

Bray, G. A. (2024). Capturing Consumer Attention: An In-Depth Analysis Of Tiktok Live Shopping. *Law and Economy*, *3*(1), 32–46.

Bustami, B., & Aryani, C. F. (2023). The Nutritional Classification of Pregnant Women Using Support Vector Machine (SVM). *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, *7*(1), 174–182.

Chodak, G. (2024). Social Commerce. In *The Future of E-commerce: Innovations and Developments* (pp. 169–185). Springer.

Elkari, B., Chaibi, Y., & Kousksou, T. (2024). Random forest with feature selection and K-fold cross validation for predicting the electrical and thermal efficiencies of air based photovoltaic-thermal systems. *Energy Reports*, *12*, 988–999.

Gupta, V., & Rattan, P. (2023). Improving twitter sentiment analysis efficiency with SVM-PSO classification and EFWS heuristic. *Procedia Computer Science*, *230*, 698–715.

Hasibuan, S. S. (2024). Sentimen Analisis Terhadap Fitur Tiktok Shop Menggunakan Naive Bayes dan K-Nearest Neighbor. *Jurnal Media Informatika Budidarma*, *8*(1), 303–311.

Huang, H., Zavareh, A. A., & Mustafa, M. B. (2023). Sentiment analysis in e-commerce platforms: A review of current techniques and future directions. *IEEE Access*, *11*, 90367–90382.

Kabra, B., & Nagar, C. (2023). Convolutional neural network based sentiment analysis with tf-idf based vectorization. *Journal of Integrated Science and Technology*, *11*(3), 503.

Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, *10*(1), 5.

Khurana, A., & Verma, O. P. (2022). Optimal feature selection for imbalanced text classification. *IEEE Transactions on Artificial Intelligence*, *4*(1), 135–147.

Korkmaz, S. (2020). Deep learning-based imbalanced data classification for drug discovery. *Journal of Chemical Information and Modeling*, *60*(9), 4180–4190.

Kunekar, P., Nimbolkar, A., Patil, A., Lakde, V., Wadile, P., & Rathod, K. (2024). Product Review Sentiment Analysis using Web Crawler and Machine Learning. *Grenze International Journal of Engineering & Technology (GIJET)*, *10*.

Mehta, P., & Pandya, S. (2020). A review on sentiment analysis methodologies, practices and applications. *International Journal of Scientific and Technology Research*, *9*(2), 601–609.

Muttaqin, M. N., & Kharisudin, I. (2021). Analisis sentimen aplikasi gojek menggunakan support vector machine dan k nearest neighbor. *UNNES Journal of Mathematics*, 22–27.

Praneswara, A. O., & Cahyono, N. (2023). Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes. *The Indonesian Journal of Computer Science*, *12*(6).

Premasudha, B. G., & Rampalli, V. (2024). A Comparative Study of Logistic Regression, Support Vector Machines, and LSTM Networks for Sentiment Classification in Academic Reviews. *2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, 1–11.

Rahmawati, L., & Santoso, D. B. (2023). Implementasi Metode Naive Bayes Untuk Klasifikasi Ulasan Aplikasi E-Commerce Tokopedia. *Intecoms: Journal Of Information Technology And Computer Science*, *6*(1), 116–124.

Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., & Choi, G. S. (2021). A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. *Plos One*, *16*(2), e0245909.

Sabila, F. L., & Andni, R. (2023). Live Streaming Strategy on TikTok Shop Media in Increasing Sales of Areta Wholesale Negligee. *Journal of Social Science Utilizing Technology*, *1*(2), 86–97.

Sarina, S., & Tanniewa, A. M. (2023). Implementasi algoritma support vector learning Terhadap Analisis Sentimen Penggunaan Aplikasi Tiktok shop seller center. *Prosiding SISFOTEK*, *7*(1), 165–170.

Septiani, A., Voutama, A., Siska, S., Hendriadi, A. A., & Heryana, N. (2024). Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Regulasi Tiktok Shop Pada Media Sosial X (Twitter). *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(4), 5729–5735.

Silva, D. (2025). E-Commerce and the Evolution of Consumer Behavior. In *Evolving Strategies for Organizational Management and Performance Evaluation* (pp. 133–154). IGI Global Scientific Publishing.

Suwanda, R., Fikry, M., & Anshari, S. F. (2024). Implementation of the Naïve Bayes Method in a Web-Based Fish Species Classification System. *Proceedings of Malikussaleh International Conference on Multidisciplinary Studies (MICoMS)*, *4*, 46.

Syahputra, R., Yanris, G. J., & Irmayani, D. (2022). SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, *6*(2), 671–678.

Wahyuni, I., Arifiansyah, R., Meidasari, E., & Siskawati, I. (2024). The role of brand equity, digital marketing, customer experience, and social media engagement on consumer purchase intention on TikTok Shop application. *International Journal of Business, Law, and Education*, *5*(2), 2806–2815.

Yacob, A., Ghazali, N. E., & Hassan, F. M. (2024). Sentiment Analysis of ChatGPT Using the KNN Algorithm and K-Fold Cross-Validation Optimization of the K Value. *JICO: International Journal of Informatics and Computing*, *1*(2), 48–55.

Zhang, C. (2024). Marketing in the tourism industry: Legal insights from. *Adapting to Evolving Consumer Experiences in Hospitality and Tourism*, 87.