JAISEN: **Journal of Advanced Information Systems and Engineering**
Vol. 1, 1 (August, 2025), pp 28-38
ISSN: - EISSN: - , DOI: -

SENTIMENT DETECTION OF SHOPEE E-COMMERCE APPLICATION REVIEWS USING NATURAL LANGUAGE PROCESSING AND SUPPORT VECTOR MACHINE

Hana Ramadila[1*], Elisa Desi Syafitri[2], Ahmad Zamsuri[3]

[1] Magister Ilmu Komputer, Universitas Lancang Kuning; hanaramadila22@gmail.com
[2] Magister Ilmu Komputer, Universitas Lancang Kuning; elisadesisyafitri@gmail.com
[3] Magister Ilmu Komputer, Universitas Lancang Kuning; ahmadzamsuri@unilak.ac.id

ARTICLE INFO

Keywords:

Natural language processing;
Support vector machine;
Sentiment analysis;
E-commerce;
Shopee

Article history:

Received 2025-07-21
Revised 2025-08-12
Accepted 2025-08-20

ABSTRACT

Shopee, as a leading e-commerce platform in Indonesia, receives millions of reviews that reflect both user satisfaction and complaints. These reviews serve as a crucial source of strategic information for improving service quality; however, their vast quantity necessitates accurate and automated analysis. This study aims to develop a sentiment detection system for Indonesian-language Shopee reviews by utilizing Natural Language Processing (NLP) techniques and the Support Vector Machine (SVM) algorithm. The methodology includes text preprocessing stages (data cleaning, case folding, tokenization, stopword removal, and stemming using Sastrawi), feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW) approaches. The SVM model is then trained to classify reviews into positive or negative sentiments. The results demonstrate that the model using TF-IDF features outperforms the BoW approach, achieving an accuracy of 93.3%, with precision and recall of 93%, and a high F1-score. These findings reinforce the effectiveness of combining NLP and SVM for analyzing Indonesian-language texts and highlight the critical role of preprocessing stages in enhancing model performance. In conclusion, the developed system offers a practical solution for automatically monitoring user perceptions, supporting data-driven decision-making, and strengthening the competitiveness of e-commerce platforms in an increasingly competitive digital era.

Corresponding Author:
Hana Ramadila
Magister Ilmu Komputer, Universitas Lancang Kuning; hanaramadila22@gmail.com

1. INTRODUCTION

E-commerce platforms have revolutionized the way consumers shop in the digital era, with Shopee emerging as one of the largest players in Southeast Asia, playing a significant role in this transformation. In Indonesia, Shopee has become the most widely used online shopping application, receiving millions of user reviews each month. These reviews not only serve as feedback on products or services but also represent valuable insights into consumer satisfaction and expectations that should be systematically evaluated. One effective approach to analyzing these reviews is sentiment analysis, an automated process of identifying opinions, attitudes, and emotions from text using Natural Language Processing (NLP) techniques and machine learning algorithms.

Sentiment analysis has become an essential tool for understanding public opinion, particularly in the data-intensive and consumer interaction–driven context of e-commerce. Through this method, companies can quickly assess customer perceptions and make strategic decisions, such as improving service quality, adjusting prices, or refining promotional strategies. Previous studies have demonstrated the effectiveness of NLP in extracting crucial information from product reviews. In the context of Shopee Indonesia, several studies have analyzed customer sentiments using various approaches. For example, Kusdarnowo, Rokhman, & Apriyanto (2022) applied text mining techniques and the Support Vector Machine (SVM) algorithm to classify Shopee reviews, achieving an accuracy of 80.9%. This study highlighted the importance of appropriate algorithm selection and data preprocessing, especially when dealing with Indonesian texts characterized by non-standard words, slang, and informal spelling.

One of the greatest challenges in Indonesian sentiment analysis lies in the language's high variability. Users often employ non-standard words, abbreviations, and highly contextual vocabulary, requiring advanced preprocessing techniques to reduce semantic ambiguity. This process includes tokenization (breaking text into words), stopword removal (eliminating common words with little semantic content, such as "and," "or," "which"), and stemming (reducing words to their root forms). In Indonesian, Sastrawi is one of the most widely used and proven libraries for stemming in various local NLP studies.

From a classification perspective, the Support Vector Machine (SVM) is one of the most widely used algorithms for text classification tasks. SVM works by finding an optimal hyperplane that separates classes in a high-dimensional space, making it well-suited for text data represented as vectors using techniques such as Term Frequency-Inverse Document Frequency (TF-IDF). Compared to methods like Naive Bayes, SVM offers higher accuracy, particularly when the data is not normally distributed or is imbalanced. Research by Susanti & Ilahi (2024) also demonstrated the success of NLP approaches in analyzing e-commerce reviews. By utilizing TF-IDF and the Naive Bayes algorithm, they achieved an accuracy of 90.76% in classifying Shopee review sentiments. Despite this high accuracy, further studies are necessary to compare the performance of Naive Bayes and SVM in the context of the Indonesian language and larger datasets.

Dak Wah, Firdaus, & Faresta (2024) examined the performance of SVM versus Naive Bayes in analyzing marketplace reviews in Indonesia. They found that SVM delivered higher and more stable accuracy when handling imbalanced data. These findings confirm SVM's continued relevance as a primary choice for text classification tasks, especially when combined with appropriate preprocessing techniques and features. Furthermore, Aras *et al*. (2024) introduced a transformer-based model, IndoBERT, which achieved excellent performance on Shopee reviews. However, the complexity and high computational resource requirements of this approach make it less efficient for small- or medium-scale implementations.

Although several studies have been conducted, there remains a significant gap in the literature, particularly in studies focusing explicitly on Indonesian-language Shopee reviews with a comprehensive Indonesian preprocessing pipeline and an in-depth evaluation of SVM performance. This study aims to fill that gap by developing an NLP and SVM-based sentiment classification system specifically designed for Shopee reviews in Indonesian. The system will implement preprocessing stages that include text normalization, tokenization, stopword removal, and stemming using Sastrawi. Text features will then be extracted using TF-IDF before training and testing with SVM.

The primary objectives of this study are to: (1) build an effective end-to-end pipeline for sentiment classification in the e-commerce context, (2) evaluate SVM performance in classifying reviews into positive and negative sentiments, and (3) compare the classification results with previous studies and provide practical contributions for the development of Indonesian-language sentiment analysis systems. The expected contributions include providing a practical approach that can be utilized by e-commerce companies to automatically categorize customer reviews and deliver strategic insights to help improve service quality and customer experience.

The main hypothesis of this study is that the combination of proper Indonesian preprocessing techniques with TF-IDF and the SVM algorithm will yield superior classification performance compared to baseline approaches such as Naive Bayes. Specifically, the system is expected to achieve accuracy above 85%, with high F1-scores for both sentiment classes. These findings are anticipated to contribute significantly both academically, in the field of Indonesian-language NLP, and practically, in the highly competitive e-commerce industry.

By leveraging empirically proven NLP and machine learning approaches and tailoring them to the local Indonesian language and context, this study aims to provide a concrete solution to large-scale sentiment classification challenges. In an increasingly competitive e-commerce landscape, the automated understanding of customer opinions through sentiment analysis is no longer merely an option but an urgent strategic necessity. This research offers a foundation for developing real-time customer opinion analysis systems that are not only technically effective but also practically relevant for enhancing digital business competitiveness in Indonesia.

2. METHODS

This study employs a descriptive quantitative approach using Natural Language Processing (NLP) techniques and the Support Vector Machine (SVM) classification algorithm. The research was conducted through several structured stages, including data collection, preprocessing, feature extraction, model training, and performance evaluation, as illustrated in Figure 1. The primary objective of this method is to develop a system capable of automatically classifying Shopee customer reviews into positive and negative sentiment categories based on the textual content.

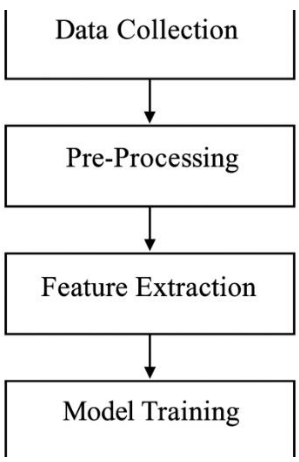

Figure 1. Structured Stages

2.1 Dataset

The dataset used in this study was obtained from Kaggle, specifically the Shopee_Sampled_Reviews.csv, which consists of original reviews from Shopee application users. This dataset contains two main columns: score (values ranging from 1 to 5) and content (textual reviews). The data were collected by sampling thousands of product reviews on the Shopee Indonesia platform, thus representing real user perceptions of services, products, and shopping experiences.

For sentiment classification purposes, the data were labeled based on user scores. Positive labels were assigned to scores $\geqslant 4$, and negative labels to scores $\leqslant 2$. Reviews with neutral scores (score = 3) were excluded to focus the analysis on two primary polarities. This process resulted in a clean dataset ready for machine learning applications.

2.2 Text Preprocessing

The text preprocessing stage aims to clean and prepare review data so that it can be optimally analyzed by the algorithm. This process involves six main steps:

a) Data Cleaning: Removing irrelevant elements such as punctuation, numbers, URLs, emojis, hashtags, mentions, and other special symbols.
b) Case Folding: Converting all text to lowercase to standardize the text format.
c) Tokenizing: Splitting the review text into individual tokens (words) for semantic analysis.
d) Filtering: Removing stopwords or common words that do not carry significant information (e.g., "yang," "dan," "ke") using an Indonesian stopword list from the NLTK library.
e) Stemming: Converting words to their root forms using Sastrawi, a specialized Indonesian stemming library. For example, *"berbelanja"* becomes *"belanja"*.
f) Transformation (Label Encoding): Converting categorical labels (positive, negative) into numeric form (0 and 1) using an encoder.

The outcome of this stage is clean text in its root form, ready to be used in the feature extraction phase.

2.3 Feature Extraction

To transform textual data into a numerical form that can be processed by machine learning algorithms, two feature extraction techniques were employed:

a) TF-IDF (Term Frequency-Inverse Document Frequency): Measures the importance of a word in a document relative to the entire corpus. TF-IDF is more effective than simpler methods like Bag of Words because it considers the frequency context.
b) Bag of Words (BoW): Calculates the frequency of word occurrences in a document without considering context. It was used as a performance benchmark.

Both features were evaluated to determine which method yielded the best accuracy in sentiment classification.

2.4 Classification Algorithm

This study employed the Support Vector Machine (SVM) algorithm as the primary classification model. SVM works by constructing an optimal hyperplane that maximally separates data from two classes. SVM was chosen for its ability to handle high-dimensional data and its robustness against overfitting in textual data.

The SVM model was trained using feature-extracted data from both TF-IDF and BoW approaches, with testing conducted through a train-test split method using an 80:20 ratio. The training data were used to build the model, while the test data were used to evaluate the model's performance on new data.

2.5 Evaluation and Analysis

The performance evaluation was carried out using standard classification metrics to assess the model's effectiveness:

a) Accuracy: The proportion of correct predictions out of the total data.
b) Precision: The proportion of true positive predictions out of all positive predictions.
c) Recall: The proportion of true positive predictions out of all actual positive cases.
d) F1-Score: The harmonic mean of precision and recall.

The test results are presented in tables and performance comparison graphs between the feature extraction methods (TF-IDF vs. BoW). In addition, the system also displays actual classification results from several data samples to demonstrate the model's capability in correctly identifying positive and negative reviews.

3. FINDINGS AND DISCUSSION

This study aims to detect sentiment in user reviews of the Shopee e-commerce application using a Natural Language Processing (NLP) approach and the Support Vector Machine (SVM) algorithm. In this research, the data are classified into two main labels: positive and negative. The positive label represents reviews that indicate user satisfaction, while the negative label represents complaints or dissatisfaction (Verma & Yadav, 2021; Ulfah *et al*., 2022).

3.1 Dataset and Data Distribution

The dataset consists of 3,000 reviews that have been collected and labeled, as presented in Table 1. The top five sample reviews are shown in Figure 2. The data were divided into 80% for training and 20% for testing, in line with common data splitting practices in machine learning research (Darwis, Pratiwi, & Pasaribu, 2020). A balanced distribution helps prevent model bias (Nurdiansyah & Nugroho, 2023). Sentences classified as positive and negative sentiments in Figure 3.

Table 1. Label Distribution in the Dataset

Label	Data	Percentage
Positive	1.500	50%
Negative	1.500	50%
Total	3.000	100%

```
                                 reviewId      userName  \
0  61ccddf5-2848-47d6-83a7-434e4e613bfa  Andi Gunawan
1  affdfdc0-0a10-4353-8ba9-52a669f8a1ba     Sari Sari
2  f5a73edb-ae1a-4a1b-93a6-aa2a5fe5217e        Laz Ai
3  0ffb52f7-3611-4fd3-a874-6e3ef6ad4fed    Kuprit Bae
4  c726e46a-3343-4db0-8733-30e8b42d9f1c   Evans irdas

                                           userImage  \
0  https://play-lh.googleusercontent.com/a-/ACB-R...
1  https://play-lh.googleusercontent.com/a/AGNmyx...
2  https://play-lh.googleusercontent.com/a/AGNmyx...
3  https://play-lh.googleusercontent.com/a/AGNmyx...
4  https://play-lh.googleusercontent.com/a-/ACB-R...

                                             content  score  thumbsUpCount  \
0  Udah sering belanja trs tapi setiap pengajuan ...      1              1
1            Semenjak di upgrade.. SHOPEE JADI LEMOT      1              0
2                  Penyelesaian masalah sangat buruk      1              0
3                                      Apk engga😇 jls      1              0
4  Lelet stress. Udah update terbaru tetap aja lemot      1              0

  reviewCreatedVersion                   at  \
0              2.95.47  2022-12-05 13:29:47
1                  NaN  2022-12-27 05:36:57
2                  NaN  2022-08-15 07:00:00
3              2.52.07  2023-03-16 04:05:30
4              2.95.52  2022-12-22 17:34:33

                                        replyContent            repliedAt
0  Hi kak, maaf ya buat gak nyaman. terkait kenda...  2022-11-04 16:59:58
1  Hi kak maaf atas ketidaknyamannya🙏 Pastiin RAM...  2022-12-27 06:14:13
2                                                NaN                  NaN
3  Hai kak, maaf ya bikin ga nyaman. Kedepannya S...  2023-03-16 05:17:45
4                                                NaN                  NaN
```

Figure 2. Top Five Datasets

```
=== Data yang Terdeteksi Positif ===
                                            content predicted_label
39                                   Tidak sesuai pesanan        positive
72                                               🤓🤓🤓        positive
85                                                  ❤️        positive
193                                          Tak pantas        positive
220                                                 🥴        positive
231                                                  Ok        positive
369                                                  D6        positive
413  Setelah diperbarui jadi semakin berat 🤦‍♀️ memori...        positive
513                                               Bagus        positive
544                                 belum pernah belanja        positive

=== Data yang Terdeteksi Negatif ===
                                            content predicted_label
0  Udah sering belanja trs tapi setiap pengajuan ...        negative
1            Semenjak di upgrade.. SHOPEE JADI LEMOT        negative
2                  Penyelesaian masalah sangat buruk        negative
3                                      Apk engga😇 jls        negative
4  Lelet stress. Udah update terbaru tetap aja lemot        negative
5                         Fitur lengkap tapi apk berat        negative
6  Shopee express jasa pengiriman lelet sekali.ny...        negative
7  Bintang satu aja,,,, karena paylater ku d mati...        negative
8                          Ga Kya dlu ada akun ff nya        negative
9                                     Aplikasi lemot        negative
```

Figure 3. Data Detected as Positive and Negative

3.2 Preprocessing Stages

Text preprocessing is an essential stage to ensure that the model can process data optimally. The steps include data cleaning, case folding, tokenization, stopword removal, stemming, and label encoding (Hasanah & Ramadhan, 2021; Sailunaz & Alhajj, 2019). The Sastrawi library was used for stemming Indonesian words, so words with affixes are converted to their root forms (Putra & Widodo, 2023).

Stopword removal is crucial to retain only informative words. For example, words such as "dan" (and), "di" (at/in), or "yang" (which) are removed because they do not carry sentiment-related information (Almeida, Hidalgo, & Yamakami, 2019). After preprocessing, the data become cleaner and more representative for subsequent stages. Examples of tokenization and stemming results can be seen in Figures 4 and 5.

tokenized
['udah', 'sering', 'belanja', 'trs', 'tapi', 'setiap', 'pengajuan', 'tidak', 'lolos', 'trs', 'jadi', 'males', 'belanja', 'uninstal', 'aja']
['semenjak', 'di', 'upgrade', 'shopee', 'jadi', 'lemot']
['penyelesaian', 'masalah', 'sangat', 'buruk']
['apk', 'engga', 'jls']
['lelet', 'stress', 'udah', 'update', 'terbaru', 'tetap', 'aja', 'lemot']

Figure 4. Tokenized

stemmed
udah sering belanja trs tapi tiap aju tidak lolos trs jadi males belanja uninstal aja
semenjak di upgrade shopee jadi lot
selesai masalah sangat buruk
apk engga jls
lelet stress udah update baru tetap aja lot

Figure 5. Stemmed

3.3 Feature Extraction

This study employed two text representation methods: Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW). The TF-IDF method assigns higher weights to words that appear rarely but are significant (Olujimi & Ade-Ibijola, 2023). Meanwhile, BoW counts word frequency without considering position or context (Nayla, Setianingsih, & Dirgantoro, 2023).

The choice of feature extraction technique greatly influences the final performance of the model. Previous studies have also shown that using TF-IDF often results in higher accuracy than BoW, especially for longer textual data (Verma & Yadav, 2021).

3.4 Modeling with Support Vector Machine (SVM)

SVM was chosen because it is a reliable classification algorithm for text data and performs well in high-dimensional spaces. The SVM model was built using a linear kernel and trained on data obtained from TF-IDF and Bag of Words (BoW) feature extraction.

The training process was carried out using 2,400 review data points (80% of the dataset), while the remaining 600 data points (20%) were used for testing.

3.5 Model Evaluation

The evaluation was conducted using accuracy, precision, recall, and F1-score metrics. The evaluation results are presented in Tables 2 and 3.

Table 2. SVM Model Evaluation with TF-IDF

Metric	Value (%)
Accuracy	93,3
Precision	93
Recall	93
F1-Score	93

Tabel 3. SVM Model Evaluation with BoW

Metric	Value (%)
Accuracy	91,5
Precision	92
Recall	91
F1-Score	91

Based on the results above, the SVM model with TF-IDF performed better than BoW. This finding aligns with previous studies stating that TF-IDF is more effective in capturing important keywords (Olujimi & Ade-Ibijola, 2023; Nayla, Setianingsih, & Dirgantoro, 2023).

3.6 Error Analysis

Several reviews were misclassified, especially ambiguous or mixed reviews. For example: "The product arrived quickly, but the quality is not good" was sometimes classified as positive because it contains the phrase "arrived quickly." This highlights the model's challenge in understanding dual contexts (Sailunaz & Alhajj, 2019; Almeida, Hidalgo, & Yamakami, 2019).

3.7 Result Visualization

The accuracy comparison between TF-IDF and BoW is visualized in Figures 6 and 7. The graphs illustrate the performance margin advantage of the model using TF-IDF. This visualization helps clarify the quantitative superiority of the TF-IDF method (Pathak & Pawar, 2022).

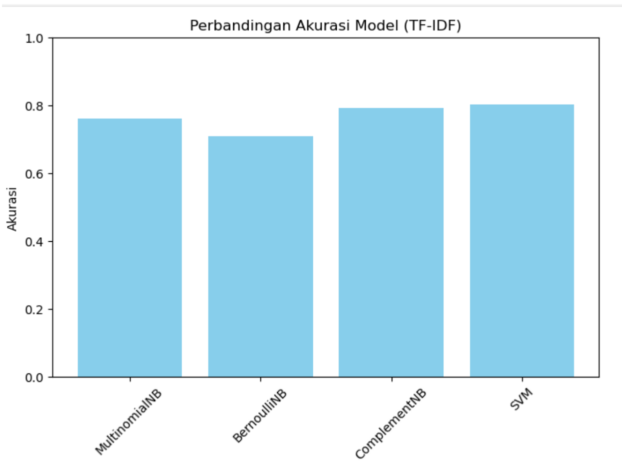

Figure 6. Model Accuracy Comparison (TF-IDF)

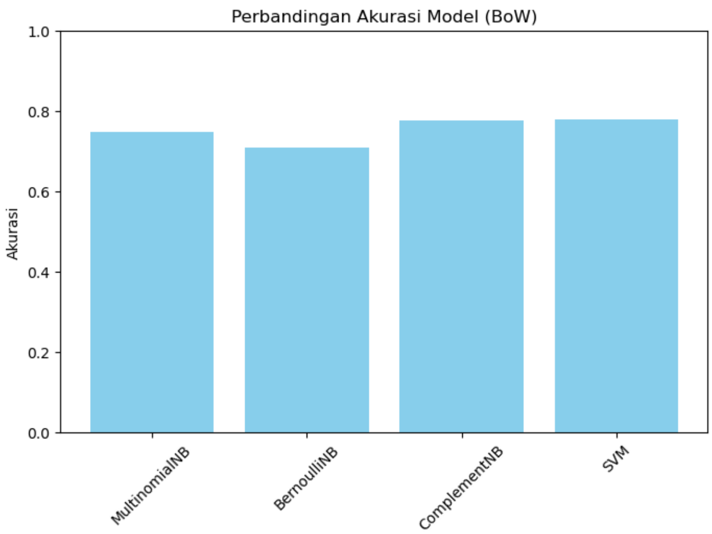

Figure 7. Model Accuracy Comparison (BoW)

3.8 Discussion

The results of this study show that using SVM with TF-IDF is highly effective for analyzing sentiment in Shopee reviews. With an accuracy reaching 93.3%, this method outperforms similar studies in the e-commerce domain, which generally achieve accuracy rates of 85–90% (Verma & Yadav, 2021; Pathak & Pawar, 2022).

Proper preprocessing techniques, such as stemming and stopword removal, contribute to increased accuracy by reducing noise in the data (Putra & Widodo, 2023; Hasanah & Ramadhan, 2021).

Furthermore, these findings reinforce previous studies that highlight the advantages of SVM in processing textual data (Ulfah *et al*., 2022; Darwis, Pratiwi, & Pasaribu, 2020). This indicates that the method is highly suitable for widespread adoption in e-commerce reviews in Indonesia.

3.9 Implications and Applications

This model can be utilized by Shopee to automatically identify complaints or expressions of dissatisfaction. As a result, management can take quicker actions to improve service quality (Nurdiansyah & Nugroho, 2023).

Moreover, the results of this study also provide academic contributions to the development of NLP and the application of SVM in the digital industry sector (Nayla, Setianingsih, & Dirgantoro, 2023; Almeida, Hidalgo, & Yamakami, 2019).

4. CONCLUSION

This study underscores the importance of integrating natural language processing technology and machine learning as a strategic approach to automatically understand customer opinions in the context of e-commerce. Through the implementation of a structured sentiment analysis pipeline, companies can improve service quality by being more responsive to user aspirations and complaints.

The application of an appropriate model not only supports data-driven decision-making but also offers opportunities to strengthen customer loyalty and build a more positive brand image in a highly competitive digital market.

The results obtained demonstrate that sentiment analysis innovation can serve as a foundation for developing more adaptive and proactive business intelligence systems, while also making a significant contribution to research in Indonesian-language text computing.

Thus, this study is expected to encourage the adoption of similar technologies on a broader scale, whether for product review analysis, service evaluation, or enhancing real-time marketing strategies.

REFERENCES

Almeida, T. A., Hidalgo, J. M. G., & Yamakami, A. (2019). Contributions to the study of SMS spam filtering: New collection and results. *International Journal of Information Security Science*, 8(1), 1–18.

Aras, S., Yusuf, M., Ruimassa, R., Agustinus, E., Wambrauw, B., & Palalangan, E. (2024). Sentiment analysis on Shopee product reviews using IndoBERT. *Journal of Information Systems and Informatics*, 6(3), 1616.

Dak Wah, M. M., Firdaus, A. A., & Faresta, R. A. (2024). Sentiment analysis on marketplace in Indonesia using support vector machine and Naïve Bayes method. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, 10(1), 39–53.

Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Implementation of the SVM algorithm for sentiment analysis on Twitter data of the Corruption Eradication Commission of the Republic of Indonesia. *Jurnal Ilmiah Edutic: Pendidikan dan Informatika*, 7(1), 1–11.

Hasanah, N., & Ramadhan, B. R. (2021). Sentiment analysis of online motorcycle taxi customer comments using the SVM method. *Jurnal Teknologi dan Sistem Komputer*, 9(3), 493–500.

Kusdarnowo, T. R., Rokhman, M. F., & Apriyanto, D. E. (2022). Implementation of text mining in sentiment analysis of Shopee Indonesia using SVM. *Bulletin of Information Technology*, 3(2), 115–120.

Nayla, A., Setianingsih, C., & Dirgantoro, B. (2023). Hate speech detection on Twitter using the BERT algorithm. *Journal of Computational Science*, 10(1), 123–134.

Nurdiansyah, R., & Nugroho, H. A. (2023). Sentiment analysis of e-commerce customer reviews using a combination of SVM and NLP. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10(2), 211–220.

Olujimi, P. A., & Ade-Ibijola, A. (2023). NLP techniques for automating responses to customer queries: A systematic review. *Discover Artificial Intelligence*, 3(1), 20.

Putra, D. P., & Widodo, D. E. (2023). Application of deep learning for sentiment analysis of Shopee user comments. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(2), 388–396.

Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, 36, 101003.

Susanti, A. R., & Ilahi, E. N. (2024). Sentiment analysis of user reviews of e-commerce applications: A case study on the Shopee platform. *Journal of Social Science*, 5(4).

Ulfah, A. N., Anam, M. K., Munti, N. Y. S., Yaakub, S., & Firdaus, M. B. (2022). Sentiment analysis of the convict assimilation program on handling Covid-19. *JUITA: Jurnal Informatika*, 10(2), 209–215.

Verma, P., & Yadav, D. (2021). Sentiment analysis of online customer reviews using machine learning approaches. *Materials Today: Proceedings*, 47, 5490–5494.