JAISEN: Journal of Advanced Information Systems and Engineering
Vol. 1, 1 (August, 2025), 39-47
ISSN: - EISSN: - , DOI: -

SENTIMENT ANALYSIS APLICATION RUANG GURU USING NAIVE BAYES AND SUPPORT VECTOR MACHINE

Amal Dluha[1*], Rizkiandi Farma Saputra[2], Santoso[3], Tony Prayitno[4]

[1] Magister Ilmu Komputer, Universitas Lancang Kuning; amaldluha10@gmail.com
[2] Magister Ilmu Komputer, Universitas Lancang Kuning; rizkiandi@gmail.com
[3] Magister Ilmu Komputer, Universitas Lancang Kuning; tosan6000@gmail.com
[4] Magister Ilmu Komputer, Universitas Lancang Kuning; tonypryitno@gmail.com

ARTICLE INFO

Keywords:

Sentiment Analysis;
Ruang Guru;
Naive Bayes;
Support Vector ;
TF-IDF.

Article history:

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

ABSTRACT

The rapid development of information technology has encouraged various innovations in the field of education, one of which is online learning applications such as Ruang Guru. As one of the largest learning platforms in Indonesia, Ruang Guru receives numerous user reviews that can be utilized to evaluate service quality. This study aims to perform sentiment analysis on user comments of the Ruang Guru application by comparing the performance of two popular classification algorithms: Naïve Bayes and Support Vector Machine (SVM). User comments were collected through web scraping from the Google Play Store and then went through a text pre-processing stage which included cleaning, case folding, tokenizing, stopword removal, and stemming. The data was then transformed into numerical representations using TF-IDF and Bag of Words feature extraction methods. The classification models were built using Multinomial Naïve Bayes and SVM with a linear kernel. The evaluation results show that the combination of SVM with TF-IDF produces higher accuracy compared to Naïve Bayes, achieving an accuracy level of more than 90% in cross- validation. These findings reinforce the evidence that SVM performs better for high-dimensional text classification tasks such as sentiment analysis in the Indonesian language. This study is expected to provide valuable insights for Ruang Guru developers to improve service quality based on user opinions and serve as a reference for further research in the field of Natural Language Processing.

Corresponding Author:
Amal Dluha
Magister Ilmu Komputer, Universitas Lancang Kuning; amaldluha10@gmail.com

1. INTRODUCTION

The massive development of digital technology has revolutionized various aspects of life, including education. The transformation in education has given rise to online learning platforms that enable flexible teaching and learning processes regardless of time and place. Ruang Guru is one of the most popular online learning applications in Indonesia, providing interactive videos, practice questions, private tutoring, and online try-outs for students from elementary to high school levels.

User reviews and feedback posted on platforms like the Google Play Store or social media are essential resources for evaluating the performance of the application. However, the sheer volume of reviews makes manual analysis inefficient and time-consuming.

Natural Language Processing (NLP) offers a solution by enabling computers to understand, process, and analyze large volumes of text automatically. Sentiment analysis is one of the widely used NLP applications, which classifies user opinions as positive, negative, or neutral. Naïve Bayes and Support Vector Machine (SVM) are popular classification algorithms commonly applied in text classification tasks due to their effectiveness and efficiency.

According to Indah and Mardiyanto (2020), the implementation of tokenizing and stemming has been shown to improve model accuracy in sentiment analysis. This is supported by Putri et al. (2021), who emphasized that the quality of classification outcomes heavily depends on the preprocessing stage. Therefore, careful execution of tokenizing and stemming is crucial for obtaining accurate and meaningful data representations (Zhafira, Rahayudi, & Indriati, 2021). Although many studies have explored these NLP techniques in the context of e-commerce and social media, their application in the education sector—particularly on user comments from online learning platforms like Ruang Guru—remains limited. In fact, sentiment analysis of LMS user feedback holds significant potential for uncovering users' perceptions and experiences of digital learning systems (Nurkumalawati & Rofli, 2023).

Based on these considerations, this study aims to analyze the sentiment of Ruang Guru user comments by implementing Support Vector Machine (SVM) and Naive Bayes algorithms, and to evaluate the performance of both models (Susandri et al., 2021). The study begins with a data preprocessing stage involving tokenizing and stemming to construct textual features, which are subsequently used in classifying sentiment into three main categories: positive, negative, and neutral (Yunitasari & Putera, 2021).

2. METHODS

This study employs a quantitative descriptive approach that integrates Natural Language Processing (NLP) techniques with machine learning-based sentiment classification (Santoso & Wibowo, 2022). The primary objective is to develop a system capable of automatically classifying Ruang Guru app reviews into positive and negative sentiment categories based on the textual content provided by users (Hendriyanto, Ridha, & Enri, 2022) To achieve this, a structured methodological framework was implemented, comprising several key stages: data collection, preprocessing, feature extraction, model training, and performance evaluation. User comments were scraped from the Google Play Store using the google_play_scraper library, resulting in a dataset of 6,000 Indonesian-language comments.

During the preprocessing phase, standard NLP techniques were applied—these included cleaning (removal of punctuation, numbers, and URLs), lowercasing, tokenizing, stopword removal, and stemming using the Sastrawi library. The clean and standardized text was then transformed into numerical representations using two widely adopted feature extraction methods: Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW). The classification task was performed using two machine learning algorithms: Support Vector Machine (SVM) and Naive Bayes. Sentiment labels (positive or negative) were assigned using a lexicon-based approach, supported by manual validation to ensure labeling accuracy. The dataset was split into training and testing subsets using an 80:20 ratio, and performance was evaluated using 10-fold cross-validation. The evaluation metrics used in this study included accuracy, precision, recall, and F1-score, allowing for a comprehensive assessment of model effectiveness. Figure 1 illustrates the overall workflow applied in this study, from raw data acquisition to final sentiment classification.

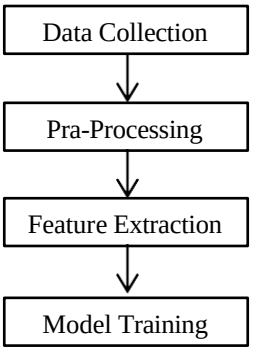


Figure 1. Structured Strages

2.1 Dataset

The dataset used in this study was obtained through an Application Programming Interface (API) call to the Ruang Guru platform, a Online Learning utilized by various educational institutions across Indonesia. The collected data consisted of original user comments—primarily from students and lecturers—providing feedback related to their experience using the application for online learning activities.

Each entry in the dataset includes comment text along with supporting metadata such as date, user role, and comment ID. Since there were no numeric scores or ratings available (as found in e- commerce platforms), sentiment labeling was conducted using a semi-automated approach:

a) Lexicon-Based Labeling, by utilizing an Indonesian sentiment lexicon containing positive and negative words.

b) Manual validation on a portion of the data to minimize misclassification.

The comments were classified into two main categories: positive sentiment and negative sentiment. Comments with neutral or ambiguous tones were excluded from the dataset to focus the analysis on clearly polarized sentiments. The outcome of this process was a clean dataset, ready for NLP processing and classification tasks.

2.2 Text Preprocessing

Text preprocessing is a critical first step in sentiment analysis using Natural Language Processing (NLP), especially when working with unstructured text data, such as user comments on the Ruang Guru platform. This stage aims to transform raw textual data into a structured, clean format suitable for feature extraction and model training. In this study, preprocessing was carried out in several systematic steps.

First, data cleaning was performed to remove irrelevant elements such as punctuation, numbers, URLs, emoticons, hashtags, mentions, and other special characters that hold no linguistic meaning. This step ensures that only meaningful words are retained.

Next, case folding was applied by converting all text into lowercase to standardize text representation and avoid duplication caused by capitalization differences (e.g., “Ruang Guru” vs. “ruang guru”). Afterward, the text was segmented into individual tokens using the tokenizing process. For example, the sentence “Ruang Guru sangat membantu pembelajaran daring” would be tokenized into [“ruang guru”, “sangat”, “membantu”, “pembelajaran”, “daring”]. This process forms the basis for semantic analysis, allowing each word to be analyzed independently.

Stopword removal followed, where common, non-informative words such as “yang”, “dan”, “di”, and “ke” were removed using a predefined stopword list from NLTK and the Sastrawi library. This was done to focus the model on meaningful words.

The next step was stemming, which converts words to their root forms using the Sastrawi library. Words like “mengakses”, “diakses”, and “pengaksesan” would all be reduced to “akses”. Stemming helps minimize unnecessary word variations and improves consistency in text modeling.

The final stage was label encoding, where comments were labeled based on sentiment polarity. Comments with positive sentiment were assigned the label 1, while those with negative sentiment

were labeled 0. Since the data consisted of free text without numeric ratings, sentiment labeling was based on a lexicon approach combined with manual validation to enhance accuracy.

As a result of this preprocessing pipeline, the corpus was cleaned, standardized, and numerically encoded—ready for the feature extraction stage using methods such as TF-IDF and Bag of Words. The success of this process greatly influences the performance of the classification algorithms used in this study, namely Support Vector Machine (SVM) and Naive Bayes.

2.3 Feature Extraction

After preprocessing, user comment texts needed to be converted into numerical representations to be processed by machine learning algorithms. This process, known as feature extraction, employed two popular text representation techniques: Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW).

TF-IDF is a word-weighting technique that evaluates the importance of a word in a document relative to the entire corpus. The Term Frequency (TF) component measures how often a word appears in a comment, while the Inverse Document Frequency (IDF) penalizes words that appear frequently across many documents, as such words tend to carry less unique information. Therefore, TF-IDF is effective in emphasizing truly relevant words in sentiment analysis and reducing the influence of common, non-discriminative words.

In contrast, Bag of Words (BoW) is a simpler method that calculates the frequency of word occurrences without considering their order or context. BoW generates high-dimensional vectors representing the presence of words in a document. Although BoW does not capture word semantics or relationships, it is widely used as a baseline in many text classification tasks due to its ease of implementation and competitive results, especially for small to medium-sized datasets.

In this study, both TF-IDF and BoW were applied in parallel and used as input to two different classifiers—SVM and Naive Bayes—to compare their performance. This comparison aimed to identify which combination of feature extraction technique and classification model was most effective in classifying the sentiment of Edlink user comments.

2.4 Classification Algorith

This study implemented two machine learning algorithms for sentiment classification of Ruang Guru user comments: Support Vector Machine (SVM) and Naive Bayes. SVM works by constructing an optimal hyperplane that separates two classes of data—positive and negative sentiment—with maximum margin. Its capability to handle high-dimensional data, such as text, makes SVM a popular choice for classification tasks, including sentiment analysis.

Naive Bayes, on the other hand, is an ensemble method based on a collection of decision trees built randomly. By aggregating predictions from multiple trees, Naive Bayes can provide stable and overfitting-resistant predictions. It is also effective in dealing with nonlinear and noisy data. In this study, both algorithms were trained using preprocessed and feature-extracted comment data through two approaches: TF-IDF and BoW. A train-test split method with a ratio of 80:20 was applied—80% for training and 20% for testing. This approach allows for an objective evaluation of each model's performance under different feature combinations.

2.5 Evaluation and Analysis

The classification models in this study were evaluated using four commonly used performance metrics: accuracy, precision, recall, and F1-score.

a) Accuracy measures the proportion of correct predictions among all test data.
b) Precision indicates the proportion of true positives among all positive predictions made by the model.
c) Recall assesses how many actual positive comments were correctly identified by the model.

d) F1-score is the harmonic mean of precision and recall and is particularly important when dealing with imbalanced data.

To assess the performance of each feature-model combination, four test scenarios were conducted:

(1) SVM with TF-IDF,

(2) SVM with BoW,

(3) Naive Bayes with TF-IDF, and

(4) Naive Bayes with BoW.

Each combination was tested to evaluate the models' accuracy and precision in classifying user comments into positive or negative sentiment categories. The evaluation results were presented in tables and graphs to facilitate comparison across models.

In addition to quantitative analysis, manual review of selected predictions was conducted to assess the models' ability to understand informal Indonesian language commonly used by Edlink users and to evaluate the consistency and relevance of their predictions. This comprehensive evaluation offers a well-rounded perspective on each model's effectiveness in performing sentiment analysis on user-generated text data.

his study implemented two machine learning algorithms for sentiment classification of Edlink user comments: Support Vector Machine (SVM) and Naive Bayes.

SVM works by constructing an optimal hyperplane that separates two classes of data—positive and negative sentiment—with maximum margin. Its capability to handle high-dimensional data, such as text, makes SVM a popular choice for classification tasks, including sentiment analysis.

Naive Bayes, on the other hand, is an ensemble method based on a collection of decision trees built randomly. By aggregating predictions from multiple trees, Naive Bayes can provide stable and overfitting-resistant predictions. It is also effective in dealing with nonlinear and noisy data.

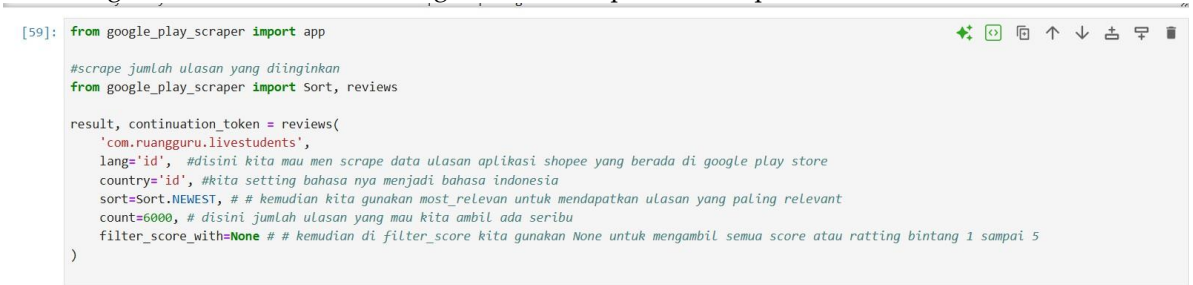
In this study, both algorithms were trained using preprocessed and feature-extracted comment data through two approaches: TF-IDF and BoW. A train-test split method with a ratio of 80:20 was applied—80% for training and 20% for testing. This approach allows for an objective evaluation of each model's performance under different feature combinations.

3. FINDINGS AND DISCUSSION

This study aims to detect sentiment in Ruang Guru app user reviews using a Natural Language Processing (NLP) approach with Support Vector Machine (SVM) and Naive Bayes algorithms. In this study, data is classified into two main labels: positive and negative. Positive labels represent reviews indicating user satisfaction, while negative labels represent complaints or dissatisfaction.

3.1 Dataset (Aplication Review Collection for Ruang Guru)

This study aims to detect sentiment in Ruang Guru app user reviews using a Natural Language Processing (NLP) approach with Support Vector Machine (SVM) and Naive Bayes algorithms. In this study, data is classified into two main labels: positive and negative. Positive labels represent reviews indicating user satisfaction, while negative labels represent complaints or dissatisfaction.



```
[59]: from google_play_scraper import app

#scrape jumlah ulasan yang diinginkan
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.ruangguru.livestudents',
    lang='id',  #disini kita mau men scrape data ulasan aplikasi shopee yang berada di google play store
    country='id', #kita setting bahasa nya menjadi bahasa indonesia
    sort=Sort.NEWEST, # # kemudian kita gunakan most_relevan untuk mendapatkan ulasan yang paling relevant
    count=6000, # disini jumlah ulasan yang mau kita ambil ada seribu
    filter_score_with=None # # kemudian di filter_score kita gunakan None untuk mengambil semua score atau ratting bintang 1 sampai 5
)
```

Figure 2. Ruang Guru Aplication Data Retrieval

3.2 Early Text Preprocessing

The first step in sentiment analysis involved data cleaning and text normalization, which aimed to remove unnecessary elements from the text such as URLs, numbers, punctuation marks, foreign characters, and excessive whitespace. Additionally, all letters were converted to lowercase to avoid differences in meaning between words like “Ruang Guru” and “ruang guru”.

[60]:

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt
0	0bfa1359-a217-4b9c-91be-55febf909841	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	seru sekali	5	0	6.97.0	2025-07-11 12:53:53	Thank you ya Reni buat feedback yang diberikan...	2025-07-11 14:29:22
1	5e372056-1673-4442-8d89-1004805ff8d0	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	timnya sering SPAM TELFON. tolong ini diperbai...	2	0	6.96.0	2025-07-11 12:39:01	Maaf ya bikin Heni jadi nggak nyaman. Jika kam...	2025-07-11 14:21:59
2	608e7816-8b07-4d5c-bb3b-87ca4333d6af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	bikin semangat belajar	4	0	6.97.0	2025-07-11 12:36:38	Keren nih, semangat belajarnya Hasya! Yuk, aja...	2025-07-11 14:15:32
3	5fb0d71f-7e0a-41b7-b1ed-5c588bb6a4af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	Sangat bagus, aplikatif dan sangat menunjang p...	5	0	6.97.0	2025-07-11 11:10:07	Terima kasih ya Riza buat apresiasi bintang 5-...	2025-07-11 14:13:15
4	88e3b8ad-3bfa-4d8d-ae59-3c10f0d9c160	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	nice app super	5	0	6.97.0	2025-07-11 10:28:33	Thanks Marfan buat bintang 5-nya. Seneng deh A...	2025-07-11 14:11:29
...	...	...	...	...	...	...	...	...	...	...
5995	b8ad5640-be3b-4f04-a537-c740f9f6978a	Cahyo Susilo	https://play-lh.googleusercontent.com/a/ACg8oc...	Plz ni y -_- jangan tanya buat no WA -_- saya ...	1	1	6.72.0	2023-11-28 19:52:28	Sedih banget dikasih bintang 1 sama Cahyo. Jel...	2023-11-29 10:15:45
5996	6ce5b7b9-1401-4f12-8da1-349c6474bc60	Latishacomel Alena	https://play-lh.googleusercontent.com/a/ACg8oc...	Bisa belajar sambil bermain	5	0	6.72.0	2023-11-28 19:08:08	Thank you ya Latisha buat feedback yang diberi...	2023-11-29 10:02:51
5997	39795238-c0b1-48e3-9644-c8a7551563e7	Henry Sinaga	https://play-lh.googleusercontent.com/a/ACg8oc...	Apk nya kadang lemot	4	0	None	2023-11-28 16:14:35	Thank you ya Henry Sinaga buat feedback yang di...	2023-11-28 16:39:48
5998	7928edc9-8304-4833-b8ce-	eko Robianto	https://play-lh.googleusercontent.com/a-/ALV-U...	Mantap	5	0	6.72.0	2023-11-27 23:11:05	Makasih eko Robianto udah kasih bintang 5.	2023-11-28 10:33:59

Figure 3. Ruang Guru Application Review Exktraction Results Data (Post-Scraping & Preprocessing)

3.3 Advanced Text Filtering

Following the initial cleaning and normalization, a more advanced filtering process was carried out to eliminate remaining irrelevant data. This step included removing empty rows or rows containing only spaces in the content_token column and eliminating very short words (fewer than three characters) that typically have little semantic value in sentiment analysis. However, exceptions were made for short but relevant terms such as “ktp”, “kk”, “apk”, “app”, and “bug,” as these frequently appear in app reviews and hold specific meanings.

[62]:

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	a
0	0bfa1359-a217-4b9c-91be-55febf909841	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	seru sekali	5	0	6.97.0	2025-07-11 12:53:53	Thank you ya Reni buat feedback yang diberikan...	2025-07-11 14:29:22	
1	5e372056-1673-4442-8d89-1004805ff8d0	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	timnya sering SPAM TELFON. tolong ini diperbai...	2	0	6.96.0	2025-07-11 12:39:01	Maaf ya bikin Heni jadi nggak nyaman. Jika kam...	2025-07-11 14:21:59	
2	608e7816-8b07-4d5c-bb3b-87ca4333d6af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	bikin semangat belajar	4	0	6.97.0	2025-07-11 12:36:38	Keren nih, semangat belajarnya Hasya! Yuk, aja...	2025-07-11 14:15:32	
3	5fb0d71f-7e0a-41b7-b1ed-5c588bb6a4af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	Sangat bagus, aplikatif dan sangat menunjang p...	5	0	6.97.0	2025-07-11 11:10:07	Terima kasih ya Riza buat apresiasi bintang 5-...	2025-07-11 14:13:15	
4	88e3b8ad-3bfa-4d8d-ae59-3c10f0d9c160	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	nice app super	5	0	6.97.0	2025-07-11 10:28:33	Thanks Marfan buat bintang 5-nya. Seneng deh A...	2025-07-11 14:11:29	
...	...	...	...	...	...	...	...	...	...	...	

Figure 4. Advanced Screening Data

3.4 Tokenizing

After the comments were cleaned and filtered, the next step was tokenizing, which is the process of splitting review texts into individual word units (tokens) using specific patterns. In this study, RegexpTokenizer from the NLTK library was used with the \w+ pattern, meaning only word characters (letters and numbers) were extracted while ignoring symbols and punctuation.

[65]:

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	a
0	0bfa1359-a217-4b9c-91be-55febf909841	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	seru sekali	5	0	6.97.0	2025-07-11 12:53:53	Thank you ya Reni buat feedback yang diberikan...	2025-07-11 14:29:22	
1	5e372056-1673-4442-8d89-1004805ff8d0	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	timnya sering SPAM TELFON. tolong ini diperbai...	2	0	6.96.0	2025-07-11 12:39:01	Maaf ya bikin Heni jadi nggak nyaman. Jika kam...	2025-07-11 14:21:59	
2	608e7816-8b07-4d5c-bb3b-87ca4333d6af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	bikin semangat belajar	4	0	6.97.0	2025-07-11 12:36:38	Keren nih, semangat belajarnya Hasya! Yuk, aja...	2025-07-11 14:15:32	
3	5fb0d71f-7e0a-41b7-b1ed-5c588bb6a4af	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	Sangat bagus, aplikatif dan sangat menunjang p...	5	0	6.97.0	2025-07-11 11:10:07	Terima kasih ya Riza buat apresiasi bintang 5-...	2025-07-11 14:13:15	
4	88e3b8ad-3bfa-4d8d-ae59-3c10f0d9c160	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	nice app super	5	0	6.97.0	2025-07-11 10:28:33	Thanks Marfan buat bintang 5-nya. Seneng deh A...	2025-07-11 14:11:29	

Figure 5. Word-Level Representation-Tokenizing

3.5 Stemming

Stemming is the process of reducing words to their base forms so that machine learning models can more easily recognize word patterns. In this study, stemming was performed using the Sastrawi library, which is specifically designed for the Indonesian language. For example, the words "mengakses," "diakses," and "pengaksesan" were all reduced to their root form "akses."

[222]:

	content	content_tokenized	content_stop	stemmed	label	stemmed_str
0	Belajar di Ruangguru membantu siswa Indonesia	[Belajar, di, Ruangguru, membantu, siswa, Indo...	[Belajar, Ruangguru, membantu, siswa, Indonesia]	[ajar, ruangguru, bantu, siswa, indonesia]	negative	ajar ruangguru bantu siswa indonesia
1	Ruangguru aplikasi edukasi online terbaik	[Ruangguru, aplikasi, edukasi, online, terbaik]	[Ruangguru, aplikasi, edukasi, online, terbaik]	[ruangguru, aplikasi, edukasi, online, baik]	negative	ruangguru aplikasi edukasi online baik
2	Sangat bermanfaat untuk semua kalangan	[Sangat, bermanfaat, untuk, semua, kalangan]	[Sangat, bermanfaat, kalangan]	[sangat, manfaat, kalang]	negative	sangat manfaat kalang
3	Gratis untuk beberapa materi dan kelas	[Gratis, untuk, beberapa, materi, dan, kelas]	[Gratis, materi, kelas]	[gratis, materi, kelas]	negative	gratis materi kelas
4	Akses mudah, fitur lengkap	[Akses, mudah, fitur, lengkap]	[Akses, mudah,, fitur, lengkap]	[akses, mudah, fitur, lengkap]	negative	akses mudah fitur lengkap

Figure 6. Stemming Results

3.6 Text Data Visualization (Preprocessing Visualization)

This stage aimed to understand the dominant characteristics of words found in Edlink user reviews. Using WordCloud techniques, a visual representation of the most frequently occurring words in positive and negative sentiment labels was displayed.

Figure 7. Visualization Results

3.7 Word Frequency Analysis

This phase aimed to identify the most frequently occurring words in positive and negative sentiment comments after stemming. The Counter function from the collections library was used to count the frequency of each token (word).

```
Top 10 Kata Paling Banyak Muncul - Positif:

Top 10 Kata Paling Banyak Muncul - Negatif:
ruangguru: 2
ajar: 1
bantu: 1
siswa: 1
indonesia: 1
aplikasi: 1
edukasi: 1
online: 1
baik: 1
sangat: 1
```

Figure 8. Top 10 Words Most Frequently Used by Users

3.8 Feature Extraction Using TF-IDF

After cleaning and stemming the user review texts, the next step was to convert the text into a numerical representation that could be processed by machine learning models. This study employed Term Frequency-Inverse Document Frequency (TF-IDF) as the feature extraction method. TF-IDF assigns a weight to each word based on how frequently it appears in a specific document compared to the entire corpus.

[108]:

	aamiin	abis	abjad	absen	absensi	ada	adaa	adakan	adaptasi	admim	...	wiffi	wifi	woiii	word	wuiz	ya
0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
657	0.0	0.0	0.0	0.725154	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
658	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
659	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
660	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	
661	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	

662 rows × 1113 columns

Figure 9. Feature Extraction Output Using TF-IDF

3.9 SVM Model Evaluation Using Cross-Validation and Confusion Matrix

Once feature representations were created using TF-IDF, the Support Vector Machine (SVM) algorithm was applied to classify user sentiment in Ruang Guru reviews. Model performance was evaluated using 10-fold cross-validation to obtain a comprehensive view of the model's accuracy.

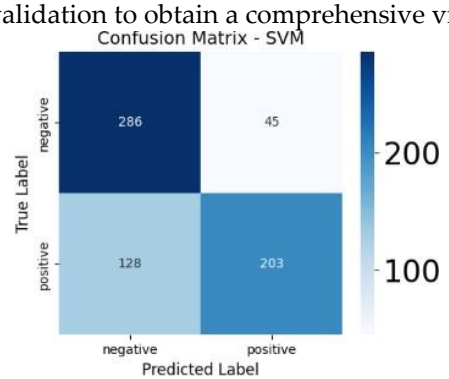


Figure 10. SVM Model Evaluation

3.10 Naive Bayes Model Evaluation Using Cross-Validation and Confusion Matrix

To compare the performance of classification models, this study also implemented the Multinomial Naive Bayes algorithm for sentiment classification of Ruang Guru user comments.

Evaluation was conducted using 10-fold cross-validation to determine the overall model accuracy. In addition, a confusion matrix was used to assess the classification precision between positive and negative sentiment labels.

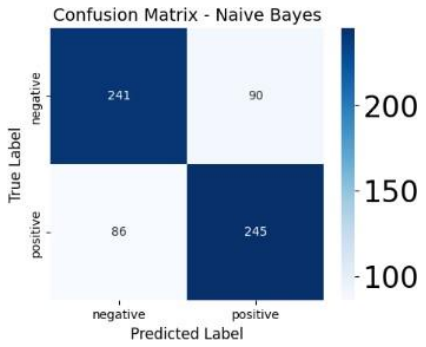


Figure 11. Naive Bayes Model Evaluation

4. CONCLUSION

This study demonstrates that the Natural Language Processing (NLP) approach combined with the Support Vector Machine (SVM) algorithm and TF-IDF features is highly effective for analyzing the sentiment of Ruang Guru app user comments. The evaluation results show that the SVM model with TF-IDF produces higher accuracy than Naive Bayes, especially in accurately distinguishing positive and negative comments. The use of comprehensive text preprocessing techniques, such as cleaning, tokenizing, stopword removal, and stemming, plays a crucial role in improving data quality and model performance. These findings strengthen evidence from previous studies that SVM is a reliable algorithm for classifying Indonesian-language text. Acknowledgments: The authors thank Universitas Lancang Kuning for support and the lecturers of the NLP course for their guidance throughout this project.

REFERENCES

Fitri, E. (2020). Sentiment analysis of the Ruangguru application using Naive Bayes, Random Forest and Support Vector Machine algorithms. *Jurnal Transformasi*, 18(1), 71.

Hendriyanto, M. D., Ridha, A. A., & Enri, U. (2022). Analisis sentimen ulasan aplikasi Mola pada Google Play Store menggunakan algoritma Support Vector Machine. *Journal of Information Technology and Computer Science*, 5(1), 1–7.

Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis sentimen ulasan aplikasi Dana dengan metode Random Forest. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 6(9), 4305–4313. Retrieved from http://j-ptiik.ub.ac.id

Nurkumalawati, I., & Rofii, M. S. (2023). Public review of M-Paspor application in Indonesia: Mobile government, digital resilience, cyber security. In *Atlantis Press SARL*.

Santoso, D. P., & Wibowo, W. (2022). Analisis sentimen ulasan aplikasi Buzzbreak menggunakan metode Naïve Bayes classifier pada situs Google Play Store. *Jurnal Sains dan Seni ITS*, 11(2).

Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis sentimen pengguna Twitter terhadap polemik persepakbolaan Indonesia menggunakan pembobotan TF-IDF dan K-Nearest Neighbor. *Jurnal Intelligent Systems and Computing*, 1(1), 43–49.

Susandri, S., Defit, S., Riandari, F., & Sinaga, B. (2021). Ekplorasi timeline: Waktu respon pesan terbaik WhatsApp group "Gurauan kita STMIK Amik". *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, 20(2), 317–324.

Suswadi, S., & Erkamim, M. (2023). Sentiment analysis of Shopee app reviews using Random Forest and Support Vector Machine. *Ilkom: Jurnal Ilmiah*, 15(3), 427–435.

Yunitasari, Y., & Putera, A. R. (2021). Analisis sentimen masyarakat di Twitter terkait pandemi Covid-19. *Smatika Jurnal*, 11(01), 22–26.

Zhafira, D. F., Rahayudi, B., & Indriati, I. (2021). Zhafira, Dhaifa Farah Rahayudi, Bayu Indriati, Indriati. *Jurnal Sistem Informasi, Teknologi Informasi, dan Edukasi Sistem Informasi*, 2(1), 55–63.