Indonesian Journal of Data and Science

Volume 7 Issue 2
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v7i2.459

Research Article

The Performance of Support Vector Machine in Classifying Public Sentiment toward Student Suicide Cases

I Gusi Gede Bagus Ngurah Sarjana[1]; Made Leo Radhitya[2]; Ni Wayan Suardiati Putri[3]
[1] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, ngurahsarjana2002@gmail.com
[2] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, leo.radhitya@instiki.ac.id
[3] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, suardiatiputri@instiki.ac.id

Correspondence should be addressed to Author; leo.radhitya@instiki.ac.id

Received 10 March 2026; Accepted 15 June 2026; Published 31 July 2026

Abstract:
Introduction: The rapid growth of social media has generated large volumes of user-generated content that can be analyzed to understand public responses to sensitive social issues. This study evaluates the performance of Support Vector Machine (SVM) in classifying public sentiment toward a widely discussed student suicide case based on YouTube comments. **Method**: A total of 5,000 comments were collected from a video on the Denny Sumargo YouTube channel using the YouTube Data API and categorized into positive and negative sentiments. Text preprocessing included cleaning, normalization, tokenization, stop-word removal, and stemming. Term Frequency-Inverse Document Frequency (TF-IDF) was used for feature extraction, while Synthetic Minority Over-sampling Technique (SMOTE) addressed class imbalance. The dataset was divided into 80% training and 20% testing data, and SVM was applied for binary sentiment classification. **Results and Discussion**: The SVM model achieved 99.96% training accuracy and 89.25% test accuracy, with precision, recall, and F1-score consistently around 89%. These results indicate that the TF-IDF, SMOTE, and SVM pipeline effectively classified Indonesian social media comments despite the linguistic complexity of discussions surrounding sensitive issues. **Conclusion**: SVM demonstrates effective and robust performance for classifying public sentiment in Indonesian YouTube comments and provides a useful approach for analyzing public responses to sensitive social phenomena.

Keywords: Sentiment Analysis; Support Vector Machine; YouTube Comments; Text Mining

1. Introduction

The rapid development of information and communication technology has changed the way society acquires, conveys, and responds to information. Social media now serves as a digital public space where users can openly and quickly express opinions, comments, and responses to various social phenomena. One of the social media platforms that significantly shapes public opinion is YouTube. In addition to serving as an entertainment medium, YouTube also becomes a forum for discussion and exchange of views on social, legal, and political issues, as well as various viral cases that attract widespread public attention. The high level of user participation in providing comments makes YouTube a potential data source for observing and analyzing public opinion trends in the digital space [1]. YouTube is a video-based platform that provides various types of content created by creators. In Indonesia, YouTube has become one of the most popular video streaming platforms and is used by various groups, ranging from individuals and communities to institutions. In addition to providing entertainment and educational content, YouTube also features discussions on various current social phenomena. The presence of the comment feature enables two-way interaction between creators and viewers, so YouTube serves not only as a medium for information but also as a public discussion space that reflects public opinion and responses [1]. One of the issues that has garnered significant attention is the case of a student who committed suicide and became a topic of discussion on various social media platforms, including YouTube. The case was discussed on Denny Sumargo's YouTube channel, uploaded on October 25, 2025. The video elicited a range of responses from the public, from expressions of empathy and support to negative comments and

harsh criticism. The numerous comments that emerged indicate a high level of public engagement in responding to sensitive social issues. Analysis of social media user comments can provide an objective picture of public perception regarding an event. However, the large volume of data and the variety of language styles used make manual sentiment classification difficult. Therefore, a sentiment analysis method capable of automatically categorizing public opinion into specific sentiment categories is needed. Sentiment analysis is a branch of text mining and Natural Language Processing (NLP) used to identify, extract, and classify a person's opinion or emotion toward an object from text data [2]. One of the machine learning algorithms widely used in sentiment analysis is Support Vector Machine (SVM). SVM is a supervised machine learning algorithm that finds the optimal hyperplane to separate data into multiple classes with maximum margin. This algorithm is known to have good generalization capabilities, especially in handling high-dimensional data such as text [3]. In the context of sentiment analysis, SVM can recognize patterns and relationships among word features derived from feature extraction, enabling effective and accurate sentiment classification. Several previous studies have shown that the SVM algorithm performs well in sentiment classification on social media. The research conducted by Desti Mualfah using the SVM algorithm to classify YouTube comments related to the deportation case of Ustad Abdul Somad achieved an accuracy of 95.02% [4]. Another study by Dedy Atmajaya on sentiment analysis of Twitter users' opinions on ChatGPT use shows that the SVM method can achieve good classification performance when supported by appropriate sentiment labeling methods [5]. Additionally, research by Marganingsih on sentiment analysis of YouTube comments for MasterChef Indonesia shows that the SVM algorithm achieves better classification performance than Gaussian Naïve Bayes [2]. Although various previous studies have shown that the SVM algorithm is effective in sentiment analysis on social media, most research still focuses on topics such as entertainment, politics, and marketing. Research on sentiment analysis regarding sensitive social issues, particularly student suicide cases, is still relatively limited. Additionally, the characteristics of YouTube comments, which are long, narrative, and emotionally charged, have the potential to affect the performance of sentiment classification algorithms. Therefore, research is needed to evaluate the performance of the Support Vector Machine (SVM) algorithm for classifying public sentiment in sensitive, complex social discussions. Based on this background, this study aims to analyze and classify the sentiment of YouTube user comments on student suicides into positive and negative categories using the Support Vector Machine (SVM) algorithm with default parameters. This research is expected to contribute to the development of sentiment analysis research based on machine learning, particularly for Indonesian-language text data related to sensitive social issues.

2. Method

This research was conducted online using public comment data from the YouTube platform, collected from November 1 to November 22, 2025. The system's workflow in this study is structured as a pipeline that integrates all stages of data processing.

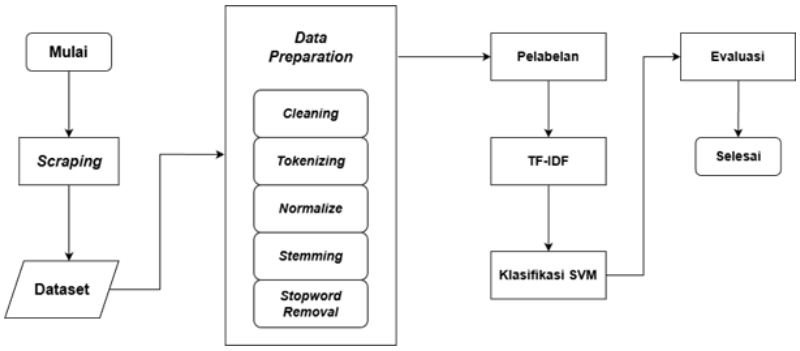

Figure 1. Research Flow

1. *Web Scraping*

Web scraping is the process of automatically extracting data from a website using a computer program or script. The data extracted usually consists of text, tables, images, or other publicly available information on the web page,

which is then stored in a structured format such as CSV, Excel, or a database for further analysis. In this research, scraping will be conducted on the YouTube platform. [6].

Data was collected using the automatic web scraping method with the help of the Python programming language thru the YouTube Data API on the Denny Sumargo channel. The amount of raw data successfully collected was 5,000 comments. The labeling process was done manually into binary categories, namely positive sentiment (support, empathy, positive judgment) and negative sentiment (criticism, disagreement, negative remarks).

2. *Suport Vector Machine*

Support Vector Machine is a learning system that uses a hypothesis space in the form of linear functions in a high-dimensional feature space and implements a learning bias derived from statistical learning theory trained with a learning algorithm. SVM is one of the machine learning algorithms used for classification and regression tasks [3]. Some advantages of SVM include using its kernel function to apply hyperplanes to high-dimensional non-linear input, making SVM superior. [7]

3. SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) is a data balancing technique used to address the issue of class imbalance in the classification process. SMOTE works by generating new synthetic data in the minority class based on the proximity of existing data, rather than directly duplicating data. By adding synthetic data to the minority class, the distribution of data between classes becomes more balanced, allowing the classification model to learn data patterns more fairly and without bias toward the majority class. SMOTE is widely used in classification research, including sentiment analysis, to improve the model's performance in predicting the minority class. [8], [9].

4. Confusion Matrix

Confusion Matrix is a commonly used evaluation method to measure the performance of a classification model in the fields of machine learning and data mining. Confusion matrix is used to illustrate the prediction results of a model by comparing the actual class and the predicted class, so that the error rate and accuracy of the built classification model can be determined. The confusion matrix is usually used in binary classification and consists of four possible prediction outcomes, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). [10]

Table 1. Confusion Matrix

Actual	Prediction	
	+	-
+	TP (True Positive)	FN (False Negative)
-	FP (False Positive)	FN (False Negative)

The final stage of the research is the evaluation of the classification model. The evaluation is conducted using a confusion matrix to determine the alignment between the model's prediction results and the actual data. [5]

Accuracy is the percentage of correct predictions compared to the total data tested. Accuracy is used to assess the general ability of the model in performing classification.

Accuracy formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision shows the level of accuracy of the model in predicting positive data. The higher the precision value, the smaller the likelihood of false positive predictions.

Precision formula:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall shows the model's ability to retrieve all truly positive data. A high recall value indicates that the model can recognize most of the positive data.

Recall formula:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

F1-Score is the harmonic mean of precision and recall. This metric is used to balance precision and recall, especially when the data is imbalanced (imbalanced dataset).

F1-Score formula:

$$\text{F1} - \text{Score} = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (4)$$

Text Processing

Text preprocessing is the initial stage in text data processing that aims to improve data quality before feature extraction and classification processes are carried out. Text data obtained from social media generally still contains various unstructured elements, such as inconsistent capitalization, punctuation, non-standard words, and words that do not have significant meaning in sentiment analysis [11], [12]. Therefore, text preprocessing is necessary to ensure that the data used is cleaner, more structured, and relevant. In general, text preprocessing in sentiment analysis consists of several main stages as follows:

1. Cleaning

Cleaning is the process of purifying text data from irrelevant elements that can disrupt the analysis process. This stage includes the removal of punctuation, numbers, symbols, emoticons, URLs, mentions, hashtags, and other special characters that do not contribute to sentiment determination. The cleaning process aims to produce cleaner text data that is ready to be processed in the next stage.

Table 2. Cleaning Example

Before Cleaning	After Cleaning
Jahat sekali iblis2 yg membully almarhum, tunggu saja kalian pembuli the power of netizen	*jahat sekali iblis yg membully almarhum tunggu saja kalian pembuli the power of netizen*
Kecerdasan Timi diturunkan dari ibu 💭💐	*kecerdasan timi diturunkan dari ibu*
aga berbelit belit bang penjelasannya 😓	*aga berbelit belit bang penjelasannya*
The best podcast ever Thankyou densu dan thankyou mama timmy sudah mendoakan untuk kita semua	*the best podcast ever thankyou densu dan thankyou mama timmy sudah mendoakan untuk kita semua*

2. Tokenizing

Tokenizing is the process of breaking down text into smaller units called tokens, usually in the form of words. This process aims to separate each word in a sentence so that it can be further processed at the feature extraction stage.

Table 2. Tokenizing

Before Tokenizing	After Tokenizing
ibunya bilang ada lihat air besar di doanya sekarang banjir di sumatra	*['ibu', 'bilang', 'lihat', 'air', 'doa', 'banjir', 'sumatra']*

Before Tokenizing	After Tokenizing
beliau semasa hidup sering belanja ke indomaret saya turut berduka cita	*['beliau', 'hidup', 'belanja', 'indomaret', 'duka', 'cita']*
terima kasih bu sudah hadir	*['terima', 'kasih', 'bu', 'hadir']*
ibunya bilang ada lihat air besar di doanya sekarang banjir di sumatra	*['ibu', 'bilang', 'lihat', 'air', 'doa', 'banjir', 'sumatra']*

3. Normalize

Text normalization is performed to standardize non-standard words (abbreviations and slang) in order to maintain data consistency.

Table 3. Normalize

Before Normalize	After Normalize
['jahat', 'sekali', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']	*['jahat', 'sekali', 'iblis', 'yang', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']*
['kecerdasan', 'timi', 'diturunkan', 'dari', 'ibu']	*['kecerdasan', 'timi', 'diturunkan', 'dari', 'ibu']*
['aga', 'berbelit', 'belit', 'bang', 'penjelasannya']	*['aga', 'berbelit', 'belit', 'bang', 'penjelasannya']*

4. Stopword Removal

Stopword removal is the process of eliminating common words that frequently appear in text but do not have a significant impact on sentiment determination. Stopword removal aims to reduce data dimensions and improve the efficiency of the classification process.

Table 4. Stopword Removal

Before *Stopword*	**After *Stopword***
['jahat', 'sekali', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']	['jahat', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'pembuli', 'the', 'power', 'of', 'netizen']
['kecerdasan', 'timi', 'diturunkan', 'dari', 'ibu']	['kecerdasan', 'timi', 'diturunkan']
['aga', 'berbelit', 'belit', 'bang', 'penjelasannya']	['aga', 'berbelit', 'belit', 'bang', 'penjelasannya']
['jahat', 'sekali', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']	['jahat', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'pembuli', 'the', 'power', 'of', 'netizen']

5. Stemming

Stemming is the process of converting inflected words into their base forms. This stage aims to unify variations of words that have the same meaning, thereby enhancing the consistency of feature representation.

Table 5. Stemming

Before Stemming	**After *Stemming***
['jahat', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'pembuli', 'the', 'power', 'of', 'netizen']	['jahat', 'sekali', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']
['kecerdasan', 'timi', 'diturunkan']	['cerdas', 'timi', 'turun', 'dari', 'ibu']
['aga', 'berbelit', 'belit', 'bang', 'penjelasannya']	['aga', 'belit', 'belit', 'bang', 'jelas']

Before Stemming	After *Stemming*
['jahat', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'pembuli', 'the', 'power', 'of', 'netizen']	['jahat', 'sekali', 'iblis', 'yg', 'membully', 'almarhum', 'tunggu', 'saja', 'kalian', 'pembuli', 'the', 'power', 'of', 'netizen']

Labeling

At the labeling stage, a list of positive words and negative words is used as a sentiment word dictionary. The system will count the number of positive and negative words in each comment, then determine the sentiment label as positive or negative based on the most dominant word count.[13].

1. Positive Words

   Positive words are a list of words that have positive meanings or sentiments. These words are used as a reference to detect whether a comment or text contains opinions that are good, supportive, satisfied, or pleasant.

```
positive_words = [
'aamiin','abadi','adem','adil','afdal','afiat','agung','ajaib','akrab','aktif',
'aktual','alqurani','aman','amazing','ambisius','anggun','anugerah','apik',
'apresiasi','arif','asik','asyik','autentik','bagus','bahagia','baik','bakat',
'balance','bangga','bantuan','baru','baru','beradab','berani','berarti',
'berbakat','bercahaya','berdaya','berdoa','berfaedah','berhasil','berharga',
'berhikmah','berilmu','berjasa','berkah','berkarisma','berkelas','bermanfaat',
'bermartabat','bernilai','berpengaruh','bersemangat','bersih','bersinar',
```

Figure 2. Positive Words

2. Negative Words

   Negatuve words adalah daftar kata-kata yang memiliki makna atau sentimen negatif. Daftar ini digunakan untuk mendeteksi apakah suatu komentar mengandung opini yang bersifat buruk, keluhan, ketidakpuasan, atau kritik.

```
negative_words = [
'abal','abis','abrutal','absurd','acuh','adab rendah','agan','agresif buruk',
'agitasi','agu','aib','ajg','ajg','akal pendek','akamsi','aklak buruk','alay',
'amburadul','ancam','ancaman','angkara','angkuh','anjing','anjg','anjrit',
'anjrot','aneh','antipati','apes','arogan','asal','asmara gelap','asam',
'asli toxic','asam','asu','asshole','bad','badut','bajingan','bakar','banci',
'bangke','bangsat','bantai','barbar','barbaric','baper buruk','bareng buruk',
'basi','bebal','becanda kasar','bekicot','belagu','belenggu','belingsatan',
'belit','bengis','benci','berbahaya','berantakan','berdosa','bergosip',
'berlebihan','bermasalah','berrasa','bersalah','biadab','biang','biang kerok',
```

Figure 3. Negative Words

SVM Modeling

At this stage, the Support Vector Machine (SVM) modeling process is carried out, which involves building a classification model using the SVM algorithm based on text data that has undergone preprocessing and sentiment labeling. This stage aims to train the system to recognize patterns from sentiment data and classify comments into positive or negative categories.

Define X and Y

The initial stage of the testing scheme is the definition of variables X and Y. Variable X represents the feature data in the form of text comment feature extraction results using the TF-IDF method, while variable Y represents the sentiment class labels, which consist of two categories: positive sentiment and negative sentiment. The definition of X and y aims to separate the input data (features) and target data (labels) so that they can be processed by the Support Vector Machine (SVM) algorithm in a structured manner.

Figure 4. Define X and Y

Spliting Data

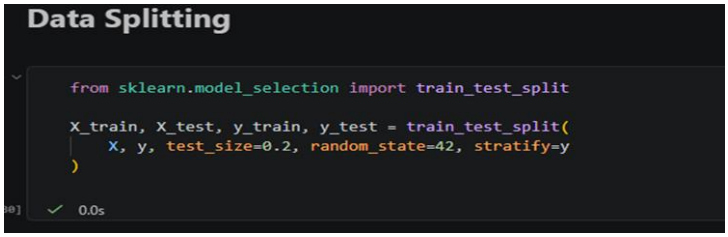

Figure 5. Data Spliting

At this stage, the data splitting process is carried out, which involves dividing the dataset into training data and testing data. The purpose of this stage is to separate the data used to train the model from the data used to test the model's performance.

The parameter test_size=0.2 indicates that 20% of the data is used as test data, while 80% of the data is used as training data. Meanwhile, the parameter random_state=42 is used so that the data splitting process always produces the same results each time the code is run, making the evaluation and reproduction of research results easier. The data splitting stage is very important in SVM modeling because it allows the model to be tested on data that has never been seen before, thus enabling the model's performance to be measured objectively.

Define Model Pipeline

At this stage, the process of defining the model pipeline is carried out, which involves structuring the model's workflow from the feature extraction process to classification using Support Vector Machine (SVM). Pipeline is used to combine several stages of data processing into a single interconnected flow. By using a pipeline, the process of training and testing the model becomes neater, more efficient, and minimizes errors in data processing.

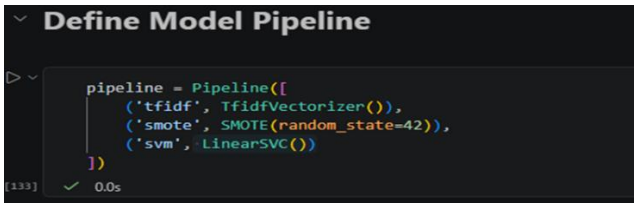

Figure 6. Define Model Pipeline

Data Train

At this stage, the data training process is carried out, which is the model training process using the training data that was divided in the previous stage. The goal of this stage is for the Support Vector Machine (SVM) model to learn the pattern of the relationship between the text data and the sentiment labels.

Figure 7. Data Train

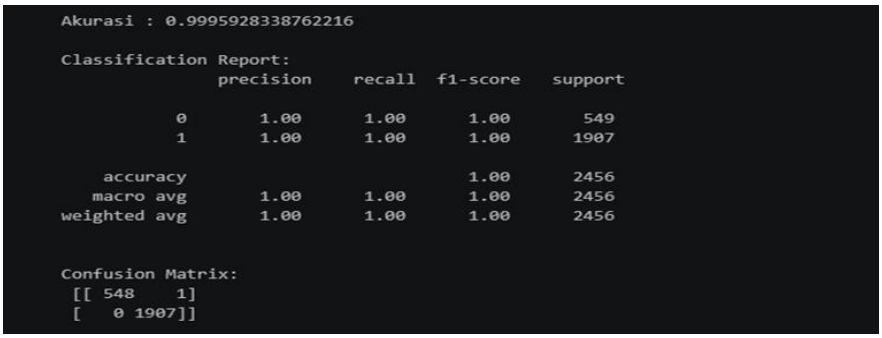

Figure 8. Data Train Result

Table 7. Data Train Confusion Matrix

Actual	Prediction	
	Positive	***Negative***
Positive	548	1
Negative	0	1907

Predict Data test

At this stage, the process of predicting test data is carried out, which is the process of testing the model that has been trained using the testing data. The purpose of this stage is to determine the ability of the Support Vector Machine (SVM) model to predict sentiment on data that has never been seen before.

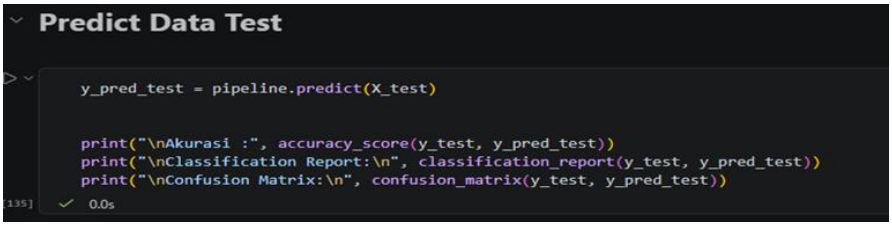

Figure 9. Predict Data Test Code

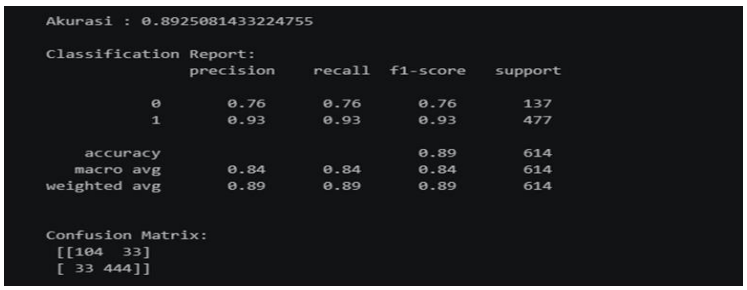

Figure 10. Predict Data Test Result

Based on the testing results on the testing data, the Support Vector Machine (SVM) model achieved an accuracy score of 0.8925 or 89.25%. This value indicates that the model has a good ability to predict sentiment on data that has not been used in the training process.

3. Result and Discussion

Confusion Matrix Result

At this stage, the Confusion matrix process is carried out, which is an evaluation method used to measure the performance of the Support Vector Machine (SVM) classification model. The Confusion matrix serves to compare the model's prediction results with the actual data labels..

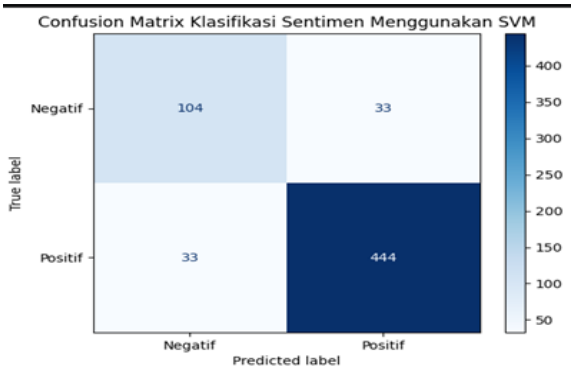

Figure 11. Confusion Matrix Result

Based on the testing results on the testing data, the Support Vector Machine (SVM) model achieved an accuracy score of 0.8925 or 89.25%. This value indicates that the model has a good ability to predict sentiment on data that has not been used in the training process.

Confusion Matrix Analysist

Based on the evaluation matrix visualization of a total of 614 independent data samples tested, the model's prediction performance is detailed as follows:

- True Positive (TP): The model successfully predicted 444 positive sentiment data points accurately.
- True Negative (TN): The model successfully predicted 104 negative sentiment data points accurately.
- False Predictions: A total of 66 prediction errors were found (a combination of false positives and false negatives).

Table 8. Evaluation Result

Model	Accuracy	Precision	Recall	F1-score
Support Vector Machine	0.89	0.89	0.89	0.89

These results confirm that the integration of SMOTE and TF-IDF in a single pipeline effectively mitigates the adverse effects of text class imbalance. However, the model appears to be slightly more sensitive and optimal in recognizing word characteristics in the positive sentiment group compared to the negative sentiment group. Overall, a performance above 89% proves that the SVM algorithm has a very high reliability in extracting and recognizing public emotion patterns on social media, even without advanced parameter optimization.

Evaluation of Classification Model Performance

Testing the SVM algorithm with default parameters shows very high performance during the training process. The model is able to learn the training data patterns optimally, achieving an accuracy of 99.96% on the training data.

Subsequently, testing the model on the test data (predict data test), which involves 20% of the data (614 test data), demonstrates the model's strong generalization capability. The SVM classification model achieved a final accuracy rate of 89.25%. The stability of the model's performance is supported by other balanced evaluation metrics, where the precision, recall, and f1-score scores consistently remain at 89%.

Discussion

Based on the research results, the Support Vector Machine (SVM) algorithm shows strong capability for sentiment classification of public comments on the YouTube platform related to the student suicide case. The use of the SVM method in this study is considered effective because it can handle high-dimensional text data and produce stable classification processes on social media comment data.

The preprocessing stages significantly impact the classification model's performance. The processes of cleaning, case folding, tokenizing, stop-word removal, and stemming successfully reduced noise in the comment data, making the text more structured and easier for the system to process. YouTube comments generally contain non-standard language, abbreviations, emoticons, and repeated words, so preprocessing is necessary to improve data quality before classification.

At the feature extraction stage, the TF-IDF method assigns weights to each word based on its frequency in the document. Using TF-IDF helps the system identify words that significantly affect comment sentiment. Words that frequently appear but lack significant meaning will have a low weight, while words relevant to the user's opinion will have a higher weight. This helps the SVM algorithm determine the boundary between sentiment classes more optimally.

The classification results show that negative sentiment is more dominant compared to positive sentiment. The dominance of negative sentiment is caused by the numerous comments expressing sadness, social criticism, empathy, and public concern regarding the student suicide case discussed in the YouTube video. Meanwhile, positive sentiment typically includes moral support, prayers, and calls to be more attentive to students' mental health.

The SVM algorithm classifies comments quite well because it finds the best hyperplane that separates data between sentiment classes. In addition, SVMs have strong generalization capabilities, making them suitable for text-based sentiment analysis research. The results of this study align with several previous studies that report that SVM achieves high performance in text classification and social media sentiment analysis.

Nevertheless, this research still has several limitations. One limitation is the use of data from a single YouTube video source, which limits the variation in comments. In addition, some ambiguous comments or those containing elements of sarcasm are still difficult for the classification system to accurately recognize. Therefore, future research can use a larger dataset, incorporate word embeddings, or compare the performance of SVM with other machine learning algorithms, such as Naïve Bayes, Random Forests, or Deep Learning, to achieve better results.

4. Conclusion

This research successfully implemented a classification model based on the Support Vector Machine (SVM) algorithm to map public sentiment toward student suicide cases based on 5,000 YouTube comments. Through structured data processing that combines TF-IDF weighting and SMOTE class balancing, the SVM model with default parameters proved to be highly reliable. This model recorded outstanding performance with an accuracy of 99.96% on the training data and successfully maintained a generic accuracy of 89.25% on the test data. With all evaluation metrics (precision, recall, and F1-score) stable at around 89%, it can be concluded that the default SVM approach is very effective and robust for recognizing and classifying unstructured Indonesian opinion texts on social media platforms.

References:

[1] A. M. Putra, Candra Saputra, Rahmaddeni, Safril Irsandi, and Vawana Muzaki, "Analisis Sentimen Masyarakat Terhadap Kasus Gas LPG 3 Kg Pada Youtube Kompas Menggunakan Metode Support Vector Machine," *Explore*, vol. 15, no. 2, pp. 163–171, Jul. 2025, doi: https://doi.org/10.35200/ex.v15i2.159.

[2] D. Marganingsih, H. Oktavianto, and G. Abdurrahman, "Analisis Sentimen Komentar Youtube Masterchef Indonesia Menggunakan Algoritma Support Vector Machine dan Gaussian Naïve Bayes," *Jurnal Informatika dan Teknologi Pendidikan*, vol. 5, no. 1, May 2025, doi: https://doi.org/10.59395/jitp.v5i1.117.

[3] M. R. Kana, N. Rahmi, and M. Mulkal, "Penggunaan Machine Learning Algoritma Support Vector Machine (SVM) Untuk Mengidentifikasi Kadar Pasir Besi di Kabupaten Aceh Besar," *Jurnal Pertambangan dan Lingkungan*, vol. 5, no. 1, p. 36, Jul. 2024, doi: https://doi.org/10.31764/jpl.v5i1.23216.

[4] R. R. G. D. M. S. Desti Mualfah1), "Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM," Apr. 2023.

[5] D. Atmajaya, A. Febrianti, H. Darwis, I. Artikel Abstrak, and K. Kunci, "Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 4, p. 2173, Aug. 2023.

[6] H. Herlawati, R. T. Handayanto, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, and D. Y. Septia, "Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 2, pp. 153–163, Nov. 2021, doi: https://doi.org/10.31603/komtika.v5i2.6280.

[7] V. Fitriyana *et al.*, "Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine," Apr. 2023.

[8] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," vol. 6, no. 3, 2021.

[9] Andika Mahesa putra, Candra Saputra, Rahmaddeni, Safril Irsandi, and Vawana Muzaki, "Analisis Sentimen Masyarakat Terhadap Kasus Gas LPG 3 Kg Pada Youtube Kompas Menggunakan Metode Support Vector Machine," *Explore*, vol. 15, no. 2, pp. 163–171, Jul. 2025, doi: https://doi.org/10.35200/ex.v15i2.159.

[10] B. Savita and D. D. Gore, "Sentiment Analysis on Twitter Data Using Support Vector Machine," *International Journal of Computer Science Trends and Technology*, vol. 4.

[11] F. Putrawansyah, "Penerapan Metode Support Vector Machine Terhadap Klasifikasi Jenis Jambu Biji," *JIKO (Jurnal Informatika dan Komputer)*, vol. 8, no. 1, p. 193, Feb. 2024, doi: https://doi.org/10.26798/jiko.v8i1.988.

[12] Ade Dwi Dayani, Yuhandri, and G. Widi Nurcahyo, "Analisis Sentimen Terhadap Opini Publik pada Sosial Media Twitter Menggunakan Metode Support Vector Machine," *Jurnal KomtekInfo*, pp. 1–10, Mar. 2024, doi: https://doi.org/10.35134/komtekinfo.v11i1.439.

[13] Chely Aulia Misrun, E. Haerani, M. Fikry, and E. Budianita, "Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 1, pp. 207–215, Apr. 2023, doi: https://doi.org/10.37859/coscitech.v4i1.4790.