

# Cyberbullying Sentiment Analysis On Tweets In X (Twitter): A Case Study Using Support Vector Machine Method

Angelin Julia<sup>a</sup>, Dewi Nasien<sup>b\*</sup>

<sup>a</sup>Department of Informatics Engineering, Pelita Indonesia Institute of Business and Technology, Pekanbaru, Indonesia

<sup>b</sup>Department of Informatics Engineering, Pelita Indonesia Institute of Business and Technology, Pekanbaru, Indonesia

## Article History

Received

3 Mei 2025

Received in revised form

28 Mei 2025

Accepted

10 June 2025

Published Online

31 June 2025

\*Corresponding author

[angelin.julia@student.pelitaindonesia.ac.id](mailto:angelin.julia@student.pelitaindonesia.ac.id)

## Abstract

The development of technology is currently experiencing very rapid progress, especially in the field of information and communication technology. The impact is so great that technology is considered an important part of everyday life. This has caused the internet to also experience major changes including in aspects of communication and information exchange. One example of technological progress is social media. The increasing number of internet users in Indonesia certainly means that there are also many social media platforms available and used by internet users to surf on the platform. Twitter (X) is a social media platform. Twitter (X) has become an effective means of spreading news and other content. Unwittingly, currently many social media users often cause various problems such as insults or defamation which is also known as cyberbullying. For this reason, this study analyzes sentiment with an approach process to identify and categorize opinions on certain topics or contexts from large data sources. By using Support Vector Machine (SVM) with the results of the Support Vector Machine method classification getting an accuracy value of 91%. The precision of the tested data results with the prediction results for the negative class is 91%, and the recall value used to measure the model's ability to predict negative data is 100% which shows that the model can detect all negative data perfectly. The results of the implementation of the sentiment prediction model using the SVM method that has been trained with a dataset that has positive and negative labeled sentences that can be effectively used to analyze sentiment in natural language texts for various purposes, such as detecting cyberbullying or analyzing public opinion.

**Keywords:** Support Vector Machine, X, Cyberbullying, Tweet

## 1.0 INTRODUCTION

The development of technology is currently experiencing very rapid progress, especially in the field of information and communication technology. The impact is so great that technology is considered an important part of everyday life. This has caused the internet to also experience major changes including in aspects of communication and information exchange.

The Indonesian Internet Service Providers Association (APJII) in February 2024 reported that the penetration rate of internet users in Indonesia in 2024 increased from 78.19% to 79.05% (apjii.or.id, 2024). In early 2024, internet users in Indonesia were reported to have reached 221.5 million people or to be precise 221,563,479 people (Fakra, 2024). Of that number, Generation Z (born 1997-2012, aged 12-27 years) is the age group that is most dominantly connected to the internet (Pebriyani, 2024). Internet users have increased from year to year, from 2018 the number of internet users as many as 106 million has increased to June 2024 at 185 million. It can be seen that the comparison of internet users with the population of Indonesia has also increased from 2018 by 40% to June 2024 by 66%. Many experts have stated that the most popular and most widely used social media platforms are Twitter (X), Facebook, Instagram, and various other internet-based applications (Harahap & Adeni, 2020).

In this context, the popularity of Twitter (X) among Generation Z is increasing 30% faster than Instagram (Gembala, 2024). Twitter (X) is a social media platform that allows users to do microblogging, which is writing short blogs in one paragraph with a maximum of 280 characters, because the number of characters in one tweet is limited (Khaira et al., 2020). With the number of users continuing to grow, Twitter (X) has become an effective means of spreading news and other content quickly (Hassolthine et al., 2023). Unwittingly, currently many social media users often cause various problems such as insults or defamation which is also known as cyberbullying.

Cyberbullying is a violent act carried out by individuals or groups through digital devices, which is carried out repeatedly against someone so that it makes it difficult for the victim to defend themselves (Dwipayana et al., 2020).

Based on the results of a survey from Ditch The Label, cyberbullying has a negative impact on mental health, including depression, eating disorders, antisocial behavior, psychosomatic symptoms and has the potential to cause victims to attempt suicide. One example of a cyberbullying case that occurred in Indonesia is the bullying experienced by a vocational high school student with the initials LNAS on TikTok social media in 2023. This case attracted public attention because it had a serious impact on the victim, such as loss of self-confidence and loss of courage to communicate with others (Ikhsanudin, 2023). Therefore, it is important to understand how cyberbullying affects a person's mental health.

Sentiment analysis is an approach process to identify and categorize opinions on a particular topic or context from large data sources (Bordoloi & Biswas, 2023). Sentiment analysis or what is often referred to as opinion mining is a branch of text classification that covers broad fields such as natural language processing, computational linguistics, and text mining related to analyzing a person's opinions, behaviors, and emotions towards objects such as events, individuals or certain topics (Betesda, 2020). Sentiment analysis can identify whether a text contains positive, negative, or neutral opinions.

Research on sentiment analysis has been widely applied and developed in several research methods such as Support Vector Machine (SVM), Naïve Bayes Classifier, Decision Tree and so on (Patel et al., 2022). The Support Vector Machine (SVM) method is a directed learning technique used to analyze data, identify patterns, and is applied in classification and regression analysis (Hakim & Kinasih, 2024).

Several studies related to cyberbullying sentiment analysis, including those implemented by Fajar Agus Maulana and lin Ernawati in 2020 entitled cyberbullying sentiment analysis on the Twitter social media network with the Naïve Bayes algorithm by conducting real-time data testing on May 12, 2020 at 01.00 WIB obtained an accuracy value of 76%. This method is effective in classifying positive and negative tweets. However, during testing, the ability to detect tweets containing cyberbullying elements was still less than optimal because there were still tweets without cyberbullying elements in the training data that were labeled negative (Maulana & Ernawati, 2020).

In 2019, Rafli Muhammad Kamal and Ednawati Rainarli also conducted research related to cyberbullying sentiment analysis entitled cyberbullying sentiment analysis on Facebook comments using the Support Vector Machine classification method. The accuracy test of this study was carried out 2 times. Where Test I uses 100 training data and 100 test data and Test II uses 100 training data and 50 test data from Facebook user comments. The results of the accuracy test carried out show that SVM can have a fairly high percentage level in the case of sentiment analysis, reaching 96% using the RBF kernel function (Kamal & Rainarli, 2019).

From various research references that have been carried out, the Support Vector Machine method is one of the method choices that the author will use in sentiment analysis. This is the author's topic to provide a solution to cyberbullying that occurs on the Twitter (X) social media platform.

## 2.0 LITERATURE REVIEW

### Sentiment Analysis

Sentiment analysis also known as opinion mining is a branch of science that analyzes opinions, sentiments, and emotions expressed in text. In the context of social media, sentiment analysis helps identify the polarity (positive/negative) of content shared by users. The process involves Text data collection, Data preprocessing, Feature extraction, Sentiment classification, and Evaluation of results.

### Social Media Twitter (X)

Twitter, which officially changed its name to X in 2023, is one of the most popular social media in Indonesia. This platform is based on a microblog with a limit of 280 characters, allowing for the rapid and interactive dissemination of information without friendship limits. X is also a space for emotional expression, including an outlet for individuals with less stable mental conditions, as well as a place to convey opinions through certain hashtags or keywords.

### Cyberbullying

Cyberbullying is a form of online bullying carried out through social media with the aim of repeatedly harming others (Kurniawan et al., 2024; Adnan et al., 2024). Even without physical contact, this behavior leaves digital

traces such as messages or recordings that can be used as evidence (Laora & Sanjaya, 2021). The impact is very serious on the victim's mental health, including depression, anxiety, and the risk of suicide.

According to Willard (2007), forms of cyberbullying include: flaming (anger), harassment (repeated harassment), denigration (defamation), cyberstalking (digital stalking), impersonation (imitation), trickery (fraud to obtain secrets), outing (spreading secrets), and exclusion (exclusion).

Legally, cyberbullying is regulated in Law No. 11 of 2008 concerning ITE which has been amended twice, most recently by Law No. 1 of 2024. Article 29 states that victims include those who suffer physical, mental, and economic losses due to criminal acts, including digital bullying. However, this regulation is still limited because it only accommodates threats of violence or intimidation. Several other articles such as Articles 27, 27A, 28(2), and Criminal Code Articles 310 and 311 are also relevant in dealing with cyberbullying. Although the Criminal Code and the ITE Law are the legal basis, victim protection is still weak. The Witness and Victim Protection Law (PSK Law) does not explicitly guarantee the rights of cyberbullying victims to medical assistance and psychological rehabilitation (Hikmawati, 2024).

### Text Preprocessing

Text preprocessing plays an important role in sentiment analysis because it aims to transform unstructured text data into a structured format using natural language processing (NLP) (Karo et al., 2022). This process is important for cleaning and preparing data so that it can be analyzed by machine learning models. Text data such as emails, newspapers, and web pages are generally unstructured, so they need to be processed into semi-structured data to detect patterns, trends, and relationships. Preprocessing in text classification generally consists of four main stages: (1) Cleaning/filtering, which removes irrelevant symbols, numbers, URLs, and duplicate words (Sugiarta et al., 2023); (2) Tokenization, which breaks text into basic word pieces or tokens (Surbakti et al., 2021); (3) Stopword removal, which removes common words that do not have significant meaning such as "yang" and "dari" (Jollyta et al., 2023); and (4) Casefolding, which is changing all letters to lower case to match the text format (Jollyta et al., 2023).

### Support Vector Machine (SVM) Method

Support Vector Machine (SVM) is a machine learning algorithm used in classification and regression, with good generalization performance, even on small datasets (Nasien et al., 2023). Support Vector Machine (SVM) works by constructing a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification (Afrianty et al., 2022). The data closest to the hyperplane is called the support vector, which determines the optimal margin between the classes (Muslim et al., 2019). The main idea of SVM is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized (Nasien et al., 2010). SVM can handle both linear and non-linear data by using kernel functions, which map the data to a higher-dimensional space so that it can be separated. Some commonly used kernels include:

1. Linear Kernel

$$K(x_1, x_2) = x_1^T x_2$$

2. Polynomial Kernel

$$K(x_1, x_2) = (\gamma x_1^T x_2 + r)^\gamma, \gamma > 0$$

3. Radial Basis Function (RBF)

$$K(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2), \gamma > 0$$

4. Sigmoid Kernel

$$K(x_1, x_2) = \tanh(\gamma x_1^T x_2 + r), \gamma > 0$$

The SVM model is built based on the principle of structural risk minimization and can be formulated mathematically with the approach Lagrangian, which produces a decision boundary. The basic equation of the hyperplane is:

$$w \cdot x + b = 0$$

with  $w$  as the weight vector and  $b$  as the bias.

With the flexibility in choosing the kernel, SVM is very effective for text classification, including in the context of sentiment analysis and cyberbullying detection.

Term Frequency-Inverse Document Frequency (TF-IDF) weighting

Term Frequency–Inverse Document Frequency (TF-IDF) is a word weighting method used in text analysis to assess the importance of a word (term) in a document to a collection of documents (corpus). This method combines two main components: Term Frequency (TF) and Inverse Document Frequency (IDF).

1. Term Frequency (TF)

TF measures the frequency of occurrence of a term  $t$  in document  $d$ . The more often the term appears, the greater its weight.

2. Inverse Document Frequency (IDF)

IDF is used to reduce the weight of terms that are too general. Terms that appear in many documents are considered less informative.

### 3. TF-IDF Calculation

The final TF-IDF value is the result of multiplying the TF and IDF weights.

The TF-IDF method is effective in highlighting important words in documents and reducing the influence of common words, making it very relevant for text classification applications such as sentiment analysis.

### Matrix Evaluation

Confusion Matrix is an evaluation method used to assess the performance of a classification model by presenting the prediction results in the form of a matrix table. This matrix compares the actual labels and predicted labels, so that it can provide a comprehensive picture of the model's performance (Maulina Putri et al., 2023).

Relying too much on a single metric such as accuracy can be misleading, especially on imbalanced data. Therefore, it is important to use various other evaluation metrics derived from the confusion matrix.

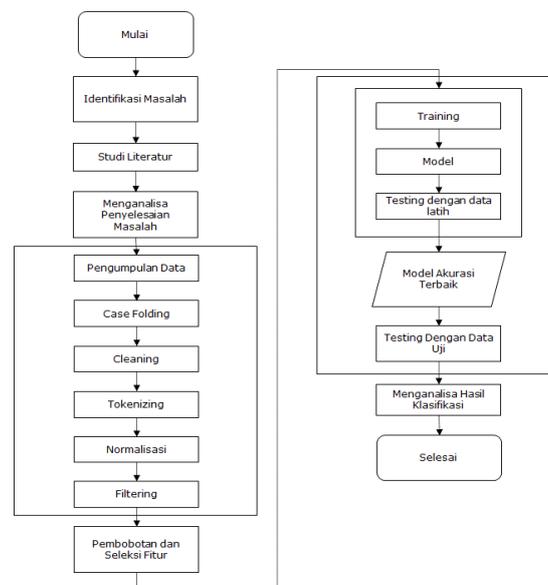
**Table 1 Classification in Confusion Matrix**

		Actual Value	
		Positive	Negative
Predicted Value	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

#### Evaluation Metrics of Confusion Matrix

1. Accuracy  
Proportion of correct predictions compared to the total number of predictions
2. Precision  
Proportion of correct positive predictions. Precision value is also known as confidence, which is the proportion of the number of cases predicted positive that are also true positives in the actual data.
3. Recall (Sensitivity)  
Proportion that measures positive instances that are actually classified as positive.
4. Specificity  
Proportion that measures negative instances that are actually classified as negative.
5. F1-Score  
Proportion that measures the combination of precision and recall. This metric is useful for considering both metrics simultaneously and F1-Score provides a single measure that reflects the balance between the two.

## 3.0 METHODOLOGY



**Figure 1. Research Framework**

### Research Stages

Research Stages are a method approach used in solving the problems studied. In this study, the author uses the SVM (Support Vector Machine) method. Data in the  $d$ -dimensional input space is denoted by  $x_i \in \mathbb{R}^d$ , while the

class label is denoted by  $y_i \in \{-1, +1\}$  for  $i = 1, 2, \dots, n$ . Where  $n$  is the amount of data. To find the values of  $x_i$  and  $y_i$  can be done when the value of each word (term) from the tf-idf weighting and the tweet label from the annotator has been obtained. The results of the tf-idf weighting are converted into the svm light data format, while the labeling results from the annotator become SVM data labels. the first stage is to input the dataset in csv format where the data from this study is in the form of cyberbullying data totaling 1000 data. The data comes from the crawling process on twitter and the selection of training data and test data is done randomly. The second stage is the data normalization process which aims to ensure that the data is in a consistent and relevant format for use in sentiment analysis. After the data is normalized, the next stage is Support Vector Machine (SVM) training, where the model is trained using the processed training data to recognize patterns related to cyberbullying sentiment. Furthermore, the testing stage is carried out on the test data to evaluate the performance of the model. At this stage, the model is tested to measure its accuracy in classifying data based on positive and negative sentiment. Finally, the classification results from the SVM model are displayed, providing output in the form of sentiment predictions on the tweet data used in this study.

### Data Collection

The data collected for this study is used as information that supports the research process. Matters related to the data collection process are as follows:

#### 1. Literature Study

Literature study is carried out by collecting and studying books, articles, and scientific journals of previous research related to the writing of this research, namely Support Vector Machine (SVM), the application of the Python programming language in sentiment analysis. Journals, e-books, and official websites as references in writing this research.

#### 2. Load Twitter Data

This stage is the process of collecting Twitter text data related to cyberbullying. Twitter provides an Application Programming Interface Key (API) that allows programmatic access to the Twitter platform, but with limitations on the amount of data that can be retrieved. To collect data, the author uses Tweet Harvest, which allows users to collect more data and can be run in the form of a Command Line Interface (CLI) and only requires an auth\_token. The author also uses one of the popular Python libraries and is used to access the Twitter API, namely Tweepy. Tweepy allows the author to automatically retrieve tweet data based on certain keywords, including tweets containing keywords related to cyberbullying. In data collection, a maximum of 500 relevant tweets were collected in Comma Separated Values (CSV) format, with tweets using Indonesian taken within the last 7 days since data collection. The Tweepy Library also allows authors to filter tweets based on language, time, and specific keywords, so that the data collected is more in line with research needs.

### Preprocessing Data

This stage is the preprocessing process, the stage of processing raw data (dirty data) so that it is ready to be processed in the testing stage. This stage is carried out to eliminate noise and speed up analysis time, so that the analysis results obtained are more accurate.

The following is an explanation of the stages in carrying out preprocessing:

1. Cleaning, is the stage of cleaning text from components that are considered meaningless for sentiment analysis such as punctuation, characters or symbols, numbers, emojis and URL links.
2. Tokenizing, is the stage of separating text sentences into word pieces according to the spaces in the text.
3. Stopword Removal, is the process of selecting or filtering words in text that have no meaning such as conjunctions, adverbs and so on, which are not needed in data modeling

After preprocessing the data, the next step is data visualization to understand the word patterns that frequently appear in the dataset.

### Data Testing

After the data is processed and sentiment analysis is performed, the next stage is testing. This testing is an important step to ensure that the sentiment analysis model built functions properly and is in accordance with the research objectives. In testing the Twitter sentiment analysis system, there are several accuracy testing models used to calculate the value, namely:

1. Precision is a metric that measures how accurate the model is in predicting the positive class and calculates the percentage of all positive predictions that actually contain cyberbullying.

$$Precision = \frac{TP}{TP + FP}$$

TP: True Positive (correct prediction for the positive class).

FP: False Positive (incorrect prediction for the positive class).

- Recall is a metric that measures how well the model finds all the positive cases that are actually in the dataset and calculates the percentage of all positive data that is successfully classified correctly.

$$Recall = \frac{TP}{TP + FN}$$

TP: True Positive (correct prediction for the positive class).

FN: False Negative (incorrect prediction for the negative class).

- F-Measure (F1-Score) is a metric that combines precision and recall to provide a more balanced picture of model performance. F1-Score is calculated as the harmonic mean of precision and recall, thus providing a single number that reflects the balance between the two.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

In addition, testing is also carried out using cross-validation, which is an additional method in data mining techniques to obtain optimal accuracy of results. This method is often known as k-fold cross validation, where testing is carried out k times on one model with the same parameters. After the validation process, the next step is to analyze the model's performance using a confusion matrix.

### Analysis of Results

The final stage of this research involves analyzing the results of the sentiment analysis that has been carried out using the Support Vector Machine (SVM) method. The evaluation was carried out using a confusion matrix, which provides a detailed picture of the model's performance in classification. The confusion matrix allows the measurement of various evaluation metrics such as accuracy, precision, recall, and F1-Score, and helps in understanding the strengths and weaknesses of the SVM model in identifying and classifying tweets accurately. By analyzing these results, the study can assess the effectiveness of the SVM method in sentiment analysis and ensure that the model can detect tweets with high accuracy. The author uses the Support Vector Machine (SVM) algorithm for classification by collecting data through the Crawling technique using the keywords "Bodoh", "Cupu" and "Jelek". This study analyzed 500 data using Python through Google Collaboratory. Based on the results of the analysis, it can be concluded that sentiment analysis of tweets related to cyberbullying can be carried out effectively using the Support Vector Machine (SVM) method.

- TF-IDF calculation

Table 2 Manual calculation of TF-IDF

TERM	TF					df	D/df	IDF(log D/df)
	D1	D2	D3	D4	D5			
main	1	0	0	1	0	2	2,5	0,39794
tangan	1	0	0	0	0	1	5	0,69897
jelek	1	0	1	0	0	2	2,5	0,39794
cupu	1	0	0	1	0	2	2,5	0,39794
bodoh	0	1	0	0	1	2	2,5	0,39794
lupa	0	1	0	0	0	1	5	0,69897
simpan	0	1	0	0	0	1	5	0,69897
kerja	0	1	0	0	0	1	5	0,69897
ketawa	0	1	0	0	0	1	5	0,69897
nyetrika	0	0	1	0	0	1	5	0,69897
hasil	0	0	1	0	0	1	5	0,69897
usaha	0	0	1	0	0	1	5	0,69897
sendiri	0	0	1	0	0	1	5	0,69897
parah	0	0	0	1	0	1	5	0,69897
mending	0	0	0	1	1	2	2,5	0,39794
hapus	0	0	0	1	0	1	5	0,69897
mati	0	0	0	0	1	1	5	0,69897

- Confusion Matrix

Labeling the sentiment as positive or negative, this dataset is all negative sentiment.

Table 3 Sentiment labeling

Teks	Label Asli	Prediksi
waduhh main tangann jelek banget pasti orangnya nih cupuuu	Negatif	Negatif
@tejuholy lagi jelek semua klo film bioskop mah	Negatif	Negatif
@93bubbletea Di mata mama dan di mata orang2 yg mengenal kamu dengan baik. Cantik tidak hanya berasal dari wajah...	Positif	Positif
@depotcepot nyari celengan gini di mana ya? terakhir punya celengan kucing tapi jelek bikinannya. ini mah cakep.	Positif	Positif
Bodoh banget sih mending mati aja luuu	Negatif	Negatif

Since all predictions match the labels, the True Positive (TP) value is 2, and True Negative (TN) is 3, there are no False Positive (FP) or False Negative (FN).

### 3. Calculating Evaluation Metrics

Accuracy:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) = (2 + 3) / (2 + 3 + 0 + 0) = 5 / 5 = 1$$

So, Accuracy = 1 (100%)

Precision:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) = 2 / (2 + 0) = 2 / 2 = 1$$

Because there is no False Positive (FP), then Precision = 1.

Recall:

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) = 2 / (2 + 0) = 2 / 2 = 1$$

Because there is no False Negative (FN), then Recall = 1.

F-1 Score:

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) = 0$$

$$\text{F1-Score} = 2 \times (1 \times 1) / (1 + 1) = 2 / 2 = 1$$

So, F1-Score = 1.

Because there is data with positive labels that are successfully predicted correctly, then the precision, recall, and F1-Score are 1, not 0. These results indicate that the model is able to perfectly classify data on the dataset used.

## 4.0 RESULTS AND DISCUSSION

### Crawling Data

In this study, the data used came from Twitter (x) with the specified keywords, namely "Bodoh", "Cupu" and "Jelek". The total data taken during the specified period was 500 tweets. In crawling this data, the researcher used the Python programming language supported by the pandas library to store and process the collected data in the form of csv format files.

### Preprocessing Data

At this stage, the dataset is processed in Google Collab using the Python programming language. This process aims to help machine learning in identifying patterns, thereby increasing the speed of the machine in learning the data set. Here are the stages of Data Preprocessing:

#### 1. Cleaning Data

This process aims to clean text data from inaccurate, damaged, or inappropriate format elements such as mentions, symbols, emoticons, and eliminate duplicate data.

#### 2. Case Folding

This stage aims to change all text to lowercase to equalize the characters in the dataset.

#### 3. Word Normalization



```
[71] # Import library yang dibutuhkan
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.svm import SVC
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Fungsi untuk memprediksi sentimen komentar
def predict_comment_sentiment(comment, svm_model, vectorizer):
    comment_tfidf = vectorizer.transform([comment]) # Transformasi ke TF-IDF
    sentiment = svm_model.predict(comment_tfidf) # Prediksi sentimen
    return sentiment[0]

# Contoh implementasi prediksi
new_comment = "keren banget"
predicted_sentiment = predict_comment_sentiment(new_comment, svm_model, vectorizer)
print(f"Hasil Prediksi dari komentar: {predicted_sentiment}")

Hasil Prediksi dari komentar: positif
```

Figure 3 Implementation Results

It can be seen that when the sentence "really cool" is entered into the model, the prediction result is positive. This shows that the model has successfully identified sentences that contain positive meanings based on words such as "really cool". Conversely, if the sentence entered has a negative meaning, such as "really bad", the model will produce negative sentiment. This process explains how the model can analyze and classify sentiments based on keywords and patterns that have been learned, whether positive or negative. The results of this implementation show that the model can effectively perform sentiment analysis on natural language texts, and can be used for various purposes, such as detecting cyberbullying or analyzing public opinion.

## 5.0 CONCLUSION

Based on the results of the research and discussion that has been carried out, it can be concluded that the Support Vector Machine (SVM) model used for cyberbullying sentiment analysis shows very good performance. This model managed to achieve an accuracy of 91% in classifying tweets containing cyberbullying. In addition, the precision value for the negative class reached 91%, indicating the model's ability to identify cyberbullying with a high level of accuracy. The recall value of 100% for negative data also indicates that the model is able to detect all cases of cyberbullying in the dataset without missing any. The consistency of model performance is also seen in k-fold cross-validation, with an average accuracy of 90.60%.

The effectiveness of data preprocessing also plays an important role in achieving these results. The process of normalization and data cleaning, such as removing stopwords and stemming, helps reduce noise in the dataset, resulting in cleaner and more relevant data for analysis. The use of the TF-IDF method in word weighting has also proven successful in optimizing the feature extraction process, which contributes to improving model performance.

In terms of practical implementation, this model has great potential to be used in detecting cyberbullying in real-time. The system built is able to identify various forms of cyberbullying in Indonesian, so it can be an effective tool for monitoring activities on social media. With promising performance, this model can be integrated into a social media monitoring system to support efforts to prevent and handle cyberbullying cases more proactively.

## References

- Adnan, A. J., Putriyani, D., Wibowo, H. A., & Ramadan, S. (2024). Perlindungan Hukum terhadap Anak Sebagai Korban Tindak Pidana Cyberbullying. *Indonesian Journal of Criminal Law and Criminology (IJCLC)*, 5(1), 25–33. <https://doi.org/https://doi.org/10.18196/ijclc.v5i1.20935>
- Afrianty, I., Nasien, D., & Haron, H. (2022). Performance Analysis of Support Vector Machine in Sex Classification of The Sacrum Bone in Forensic Anthropology. *JURNAL TEKNIK INFORMATIKA*, 15(1), 63–72. <https://doi.org/https://doi.org/10.15408/jti.v15i1.25254>
- apjii.or.id. (2024). *Pengguna Internet di Indonesia Meningkat di 2024*. Asosiasi Penyelenggara Jasa Internet Indonesia (APJII).
- Betesda. (2020). Peningkatan Optimasi Sentimen Dalam Pelaksanaan Proses Pemilihan Presiden Berdasarkan Opini Publik Dengan Menggunakan Algoritma Naïve Bayes Dan Paricle Swarm Optimization. *JSI (Jurnal Sistem Informasi) Universitas Suryadarma*, 7(2), 101–114. <https://doi.org/10.35968/jsi.v7i2.452>
- Bordoloi, M., & Biswas, S. K. (2023). Sentiment analysis: A survey on design framework, applications and future scopes. In *Artificial Intelligence Review* (Vol. 56, Issue 11). Springer Netherlands. <https://doi.org/10.1007/s10462-023-10442-2>
- Dwipayana, N. L. A. M., Setiyono, S., & Pakpahan, H. (2020). Cyberbullying Di Media Sosial. *Bhirawa Law Journal*, 1(2), 63–70. <https://doi.org/10.26905/blj.v1i2.5483>
- Fakra. (2024). *Pengguna Internet Indonesia Tembus 221 Juta, Didominasi Gen Z*. Lintasterkini.Com.

- Gembala, J. (2024). *23 Statistik Penting Twitter (X) yang Perlu Anda Ketahui di Tahun 2024*. TheSocialShepherd-Com.
- Hakim, B., & Kinasih, P. R. (2024). Sentiment Analysis of Indonesian Citizen Tweets Using Support Vector Machine on the Rebranding of Twitter to X. *Jurnal Tekno Kompak*, *18*(2), 468. <https://doi.org/10.33365/jtk.v18i2.4293>
- Harahap, M. A., & Adeni, S. (2020). Tren Penggunaan Media Sosial Selama Pandemi Di Indonesia. *Jurnal Professional FIS UNIVED*, *7*(2), 13–23.
- Hassolthine, Ramadhona, C., Riad Sahara, Haq, F. A. S. N., & Abdullah, S. (2023). Penerapan Algoritma K-Nearest Neighbour (K-Nn) Sebagai Klasifikasi Tweets Pada Twitter. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, *7*(2), 358–363. <https://doi.org/10.35145/joisiej.v7i2.3975>
- Hikmawati, P. (2024). Regulasi Cyberbullying dan Perlindungan Hukum terhadap Korbannya. *Negara Hukum: Membangun Hukum Untuk Keadilan Dan Kesejahteraan*, *15*(1), 17–35. <https://doi.org/10.22212/jnh.v15i1.4383>
- Ikhsanudin, A. (2023). *KPAI: Luluk Nuril Lakukan Cyberbullying, Korban Hilang Percaya Diri*. News.Detik.Com. <https://news.detik.com/berita/d-6918297/kpai-luluk-nuril-lakukan-cyberbullying-korban-hilang-percaya-diri>
- Kamal, R. M., & Rainarli, E. (2019). Analisis Sentimen Cyberbullying Pada Komentar Facebook Dengan Metode Klasifikasi Support Vector Machine. *Universitas Komputer Indonesia*. <https://elibrary.unikom.ac.id/id/eprint/1228/>
- Karo, I. M. K., Fudzee, M. F. M., Kasim, S., & Ramli, A. A. (2022). Karonese Sentiment Analysis: A New Dataset and Preliminary Result. *International Journal on Informatics Visualization*, *6*(2–2), 523–530. <https://doi.org/10.30630/joiv.6.2-2.1119>
- Khaira, U., Johanda, R., Utomo, P. E. P., & Suratno, T. (2020). Sentiment Analysis of Cyberbullying on Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*, *3*(1), 21–26. <https://doi.org/10.24014/ijaidm.v3i1.9145>
- Kurniawan, K., Yosep, I., Khoirunnisa, K., Mardhiyah, A., & Fitriani, N. (2024). Dampak Cyberbullying terhadap Risiko Bunuh Diri pada Remaja: Narrative Review. *JPP (Jurnal Kesehatan Poltekkes Palembang)*, *19*(1), 8–16. <https://doi.org/10.36086/jpp.v19i1.2012>
- Laora, A. Y., & Sanjaya, F. (2021). Fenomena Cyberbullying di Media Sosial Instagram (Studi Deskriptif Tentang Kesehatan Mental Pada Generasi Z Usia 20-25 Tahun di Jakarta). *Oratio Directa (Prodi Ilmu Komunikasi)*, *3*(1), 346–368. <https://www.ejurnal.ubk.ac.id/index.php/oratio/article/view/160/113>
- Maulana, F. A., & Ernawati, I. (2020). Analisa sentimen cyberbullying di jejaring sosial twitter dengan algoritma naïve bayes. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 529–538. <https://conference.upnvj.ac.id/index.php/senamika/article/view/619>
- Maulina Putri, S. S., Arhami, M., & Hendrawaty, H. (2023). Penerapan Metode SVM pada Klasifikasi Kualitas Air. *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, *3*(2), 93. <https://doi.org/10.30811/jaise.v3i2.4630>
- Muslim, M. A., Prasetyo, B., Mawarni, E. L. H., Herowati, A. J., Mirqotussa'adah, Rukmana, S. H., & Nurzahputra, A. (2019). *Data Mining Algoritma C4.5 Disertai Contoh Kasus dan Penerapannya dengan Program Computer* (E. Listiana & N. Cahyani (eds.)).
- Nasien, D., Yuhaniz, S. S., & Haron, H. (2010). Statistical Learning Theory and Support Vector Machines. *2010 Second International Conference on Computer Research and Development*, 760–764. <https://doi.org/10.1109/ICCRD.2010.183>
- Patel, A., Oza, P., & Agrawal, S. (2022). Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model. *Procedia Computer Science*, *218*, 2459–2467. <https://doi.org/10.1016/j.procs.2023.01.221>
- Pebriyani, H. (2024). *Hasil Survei APJII : Pengguna Internet di Indonesia Tembus 221 Juta, Mendominasi Gen Z*. Komite.Id.
- Sugiarta, I. K. A., Dewi, P. A. C., & Utami, N. W. (2023). ANALISA SENTIMEN MAHASISWA TERHADAP LAYANAN STMIK PRIMAKARA MENGGUNAKAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, *5*(3). <https://doi.org/https://doi.org/10.51401/jinteks.v5i3.3159>
- Surbakti, A. Q., Hayami, R., & Amien, J. Al. (2021). Analisa Tanggapan Terhadap Psbb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter. *Jurnal CoSciTech (Computer Science and Information Technology)*, *2*(2), 91–97. <https://doi.org/https://doi.org/10.37859/coscitech.v2i2.2851>
- Willard, N. E. (2007). *Cyberbullying and Cyberthreats: Responding to the Challenge of Online Social Aggression, Threats, and Distress*.