Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

Sentiment Analysis of Starlink on Twitter Using Support Vector Machine Algorithm

Sardin[1)*], Agung Nugroho[2)], Nanang Tedi Kurniadi[3)]
[1)*2)3)]Informatics Engineering, Universitas Pelita Bangsa, Bekasi Regency, Indonesia
[1)*]sardin.s31@mhs.pelitabangsa.ac.id, [2)]agung@pelitabangsa.ac.id, [3)]nanang@pelitabangsa.ac.id

ABSTRACT

Indonesia faces unique challenges in the provision of internet services. Cable and fiber optic infrastructure is often difficult and expensive to implement in many areas. Based on data from the Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), by 2024 internet users will reach 221.5 million. Starlink, Elon Musk's satellite-based internet service through his SpaceX company, offers an innovative solution to fill the void in Indonesia's telecommunications infrastructure. However, Starlink's presence has raised concerns among local service providers, particularly regarding potential market disruption and existing regulations. Starlink could also pose a potential threat to Indonesia's security and sovereignty. Starlink has been a hot topic on various social media platforms, including Twitter. Twitter is a very popular social media platform with millions of active users who often share their opinions in real-time. The number of public responses in assessing the presence of Starlink in Indonesia, became a reference for a sentiment analysis. The Support Vector Machine algorithm is used to classify opinions into positive and negative categories. Based on testing that has been done using the Cross Validation technique with a K-Fold value with a total of 1976 tweets data. The results show that 1112 tweets contain positive sentiment and 864 tweets contain negative sentiment. This shows that 56.3% of people agree with the presence of starlink in Indonesia. While from the use of the Support Vertor Machine algorithm, the Accruracy value is 76.22%, Precision is 77.48%, and Recall is 81.38%.

Keywords: Sentiment Analysis; Internet; Starlink; Classification; Support Vector Machine

1. INTRODUCTION

The development of internet technology has brought profound and extensive changes in various aspects of human life. The internet originated from the ARPANET (Advanced Research Projects Agency Network) project developed by the United States Department of Defense in the late 1960s (Andy Nugroho, 2020). The Internet has become one of the most significant innovations in human history, changing the way we communicate, work and live on a daily basis. Since it was first introduced in the late 20th century, internet technology has experienced rapid development and continues to transform, creating new opportunities and challenges in various aspects of life.

At present, the use of the internet in Indonesia is growing extraordinarily fast and the internet plays an important role in all aspects of human life (Ananda & Pristyanto, 2021). Based on data from the Indonesian Internet Service Providers Association (APJII), by 2024 internet users will reach 221.5 million out of a total population of 278.6 million Indonesians in 2023 (Muhammad Arif, 2024). Fast and stable internet access plays an important role in driving economic growth, improving education, and facilitating communication and information.

Indonesia faces unique challenges in the provision of internet services. Cable and fiber optic infrastructure is often difficult and expensive to implement in many areas. Starlink, with its satellite technology, offers a solution that can reach remote and infrastructure-poor areas more efficiently. Starlink is a satellite internet service developed by SpaceX, an Aerospace manufacturing and private space transportation services company founded by Elon Musk (Muhammad Fuad, 2024).

Recently, the public has been shocked by the presence of Starlink in the Indonesian market. The presence of Starlink, a satellite-based internet service owned by Elon Musk through the SpaceX company, offers an innovative solution to fill the void of telecommunications infrastructure in Indonesia. Starlink uses Low Earth Orbit (LEO) satellite technology that can provide broad and consistent internet coverage at high speeds (Sekretariat KADIN

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

Indonesia, 2024). However, Starlink's presence has raised concerns among local service providers, particularly regarding potential market disruption and existing regulations. Starlink may also pose a potential threat to Indonesia's security and sovereignty (Agus Tri Haryanto, 2024).

Starlink has become a hot topic on various social media platforms, including Twitter. Twitter is a very popular social media platform with millions of active users who often share their opinions in real-time (Ismail & Raden Bagus Fajriya Hakim, 2023). The number of public responses in assessing the presence of Starlink in Indonesia is a reference for researchers to conduct sentiment analysis. Sentiment analysis is the process of extracting subjective information from text and determining the author's attitude towards a particular topic, whether positive or negative (Malik Zuhdi, Utami, & Raharjo, 2019). The basic task in sentiment analysis is to classify the polarity of the text in documents, sentences, or opinions (Amrullah, Sofyan Anas, Adrian, & Hidayat, 2020). The algorithm that will be used to perform sentiment analysis is Support Vector Machine (SVM).

There are many algorithms that can be used to analyze sentiment. SVM has several advantages, namely high accuracy, efficient in memory usage and can handle data that is not normally distributed. Based on these advantages, the SVM algorithm was chosen to be used in this study. Quoted from several sources, it was found that the Support Vector Machine algorithm is superior to other algorithms. The results on the accuracy calculation show that the Support Vector Machine method is superior with a value of 89.70% without K-Fold Cross Validation and 88.76% with K-Fold Cross Validation (Rangga, Nasution, & Hayaty, 2019).

As a result of this study, through analyzing the sentiment towards Starlink on Twitter, the researcher gained a deeper understanding of public opinion. This helps researchers explore the factors that influence people's perceptions of new technologies, as well as how those perceptions may change over time. Researchers can use these findings to strengthen sentiment analysis methodology with the Support Vector Machine (SVM) algorithm. This study can serve as a reference in developing big data analysis techniques and improving the accuracy of sentiment classification on various social media platforms. Overall, this study provides many positive implications for researchers. It opens up new opportunities for further exploration, development of better analysis tools, and contributions to the existing scientific literature.

2. LITERATURE REVIEW

Related research conducted by Styawati, Nirwana Hendrastuty, Auliya Rahman Isnain, Ari Yanti Rahmadhani, (2021), namely "Analysis of Public Sentiment Towards the Pre-Employment Card Program on Twitter with the Support Vector Machine Method" (Styawati, Hendrastuty, & Isnain, 2021). The results of the classification using the Support Vector Machine method which is divided into three neutral classes are 98.34%, negative classes are 0.99%, and positive classes are 0.66%. The results of the evaluation carried out on the linear kernel accuracy value of 98.67%, precission 98%, recall 99%, and F1-Score 98%, while at the RBF kernel accuracy value of 98.34%, precission 97%, recall 98%, F1-Score 98%, it can be concluded that public sentiment from twitter users towards the pre-employment card program during the pandemic is more inclined to neutral by 98.34%.

Other related research was conducted, Dedi Darwis, Eka Shintya Pratiwi, A. Ferico Octaviansyah Pasaribu, (2020) namely "Application of the Svm Algorithm for Sentiment Analysis on Twitter Data of the Corruption Eradication Commission of the Republic of Indonesia" (Darwis, Shintya Pratiwi, Ferico, & Pasaribu, 2020). The results of classification using the SVM method are divided into three classes, namely positive classes as much as 8%, negative classes as much as 77%, and neutral classes as much as 15%. Based on the results of testing and evaluation carried out on the accuracy, precession, recall, and F1-Score values, it can be concluded that the sentiment of the twitter user community regarding the performance of the KPK with a very unfavorable percentage is indicated by the appearance of a negative sentiment skew of 77% with the accuracy of the accuracy test results of 82% and precision testing of 90%, as well as recall of 88% and f1-score of 89%.

Another related research was conducted by Dian Siti Utami, Adhitia Erfina, (2021), namely "Sentiment Analysis of Online Loans on Twitter using the Support Vector Machine (Svm) Algorithm" (Utami & Erfina, 2021). The results of the implementation of the Support Vector Machine (SVM) algorithm, in this study the data shows that the accuracy rate for online loans is 62.00%. The accuracy results in the study are considered quite good. Sentiment analysis with the Support Vector Machine (SVM) algorithm. successfully classifies community sentiment on Twitter about online loans. The classification results show that negative sentiment dominates over positive sentiment, there are 59% Negative reviews and 41% Positive reviews. It can be concluded that there are still many people who are hesitant

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

about online loans because the risk is quite high. And the number of victims of online loans makes people afraid to try. on the other hand, some people think positively about online loans because they can help financial needs.

Other related research was conducted, Ristyani Slamet, Windu Gata, Annisa Novtariany, Khairunisa Hilyati, Febri Ainun Jariyah, (2022) namely "Twitter Sentiment Analysis of the Use of South Korean Artists as Brand ambassadors for Local Beauty products" (Slamet, Gata, Novtariany, Hilyati, & Jariyah, 2022). The best model that can be used for sentiment classification analysis is SVM with an accuracy value of 83.60% with precision 83.86% and recall 99.62%. The results of the sentiment classification analysis on the use of South Korean artists as brand ambassadors for local beauty products show the tendency of the community to be positive about the decision.

Another related research was conducted by Aldiansyah Putra, Dede Haeirudin, Hasna Khairunnisa, Retnani Latifah, (2021), namely "Analysis of Public Sentiment Towards PPKM Policy on Twitter Social Media Using Svm Algorithm" (Putra, Haeirudin, Khairunnisa, & Latifah, 2021). The results of Sentiment analysis regarding PPKM from Twitter data using SVM show an accuracy of 64% with a Linear kernel and regularization penalty C = 1. Where from the prediction results on non-label tweet data, it is known that the distribution is 57%-60% labeled positive and the rest labeled negative. However, from the word cloud analysis, it is known that the SVM algorithm is able to recognize PPKM rejection tweets as tweets with negative tendencies and can identify important words related to these negative tendencies such as 'restaurants' and 'slowly'.

3. METHOD

Framework

This framework is made to make it easier to understand the direction of research from the process to the results that will be obtained. The following proposed framework can be seen in the figure below:

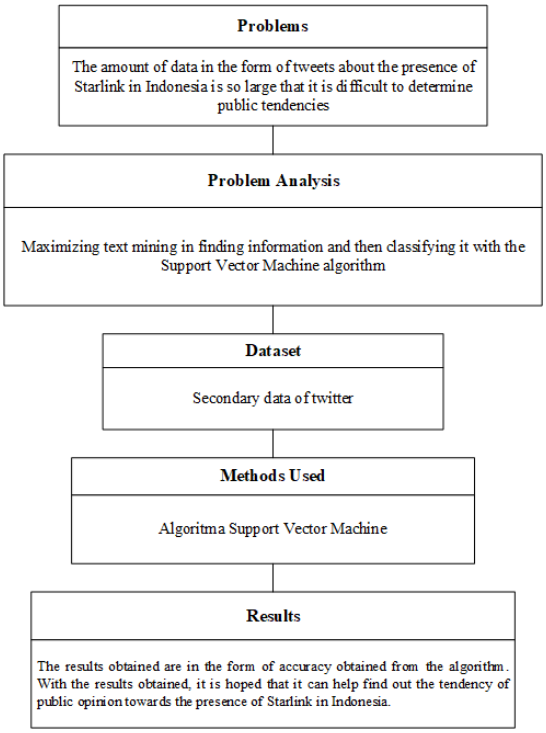


Fig. 1 Framework

At this stage, the existing problems will be explained as the background of the research and then make the scope of the problem and describe the problem to make it simpler. The data used is secondary data, the dataset is taken from the twitter website with the help of the twitter API and the algorithm used is Support Vector Machine. The results

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

obtained from this research are in the form of algorithm accuracy results and public opinion regarding the presence of Starlink in Indonesia.

Research Method

This research focuses on analyzing the sentiment towards the presence of starlink in Indonesia. The research will go through several processes, the following is the modeling of the proposed Support Vector Machine algorithm.

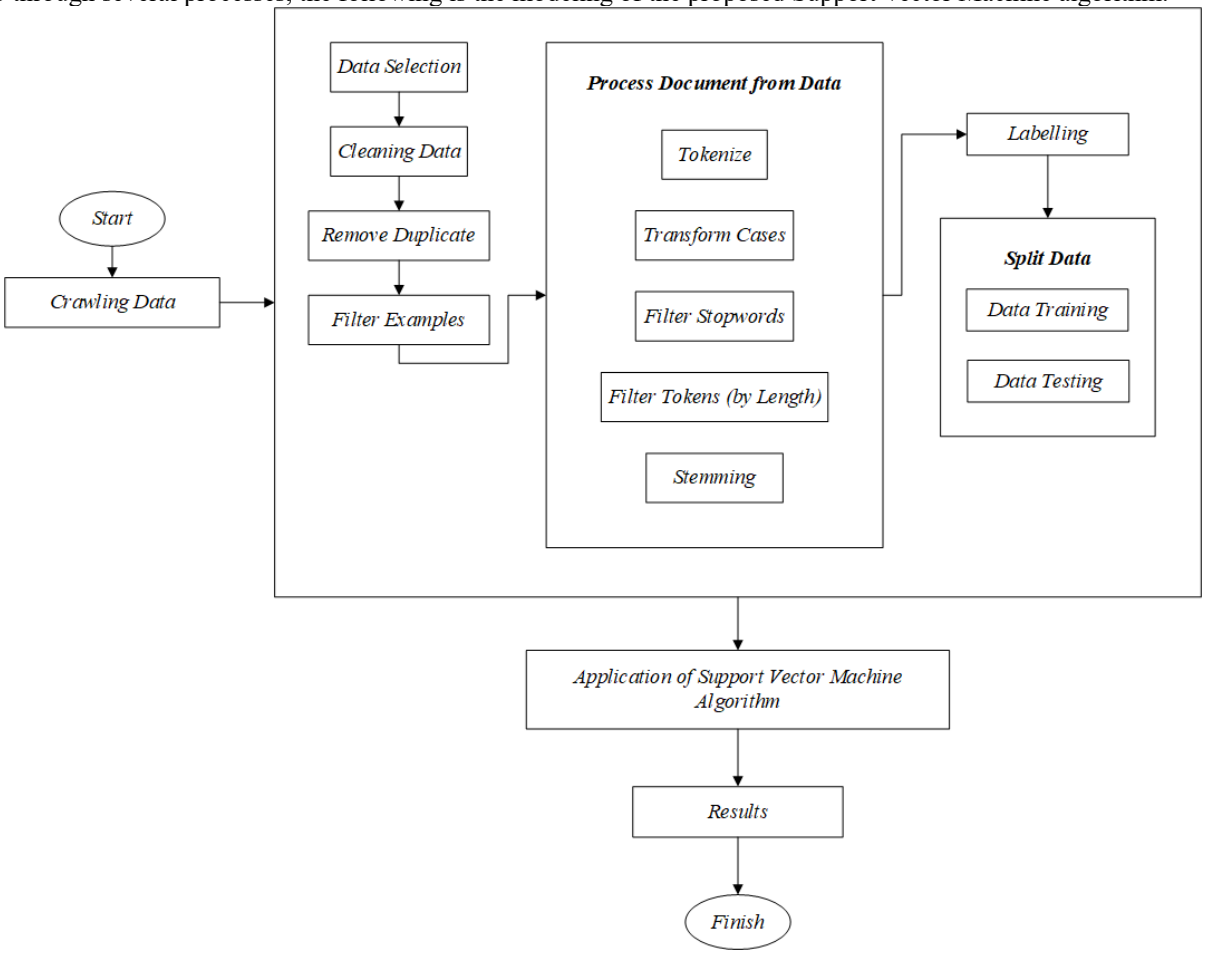


Fig. 2 Research Method

Data collection is done by crawling on twitter social media. To facilitate the sentiment analysis process, the data needs to be normalized first. The purpose of preprocessing is to remove noise or minimize the presence of data that is not needed in the sentiment analysis process. The first preprocessing component used is as follows: Data Selection, Cleansing, Remove Duplicates, Filter Examples. The second preprocessing component used is as follows: Tokenize, Transform Case (to lower), Filter Stopwords, Filter Tokens (by Length), Stem (Dictionary). Determination of attribute classes is based on the nature of the tweet data that has been obtained. To classify, the data that has been cleaned is first separated into training data and testing data. The algorithm used in this research is Support Vector Machine. The evaluation is carried out to determine the accuracy and performance of the applied modeling. Validation aims to see the accuracy of the model used. The validation technique used in this research is Performance Classification, the accuracy of the algorithm will be measured using Confusion Matrix.

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

4. RESULT

Dataset

In this study, the data collection process was carried out using Google Collaboratory tools with the query "starlink indonesia", a data limit of 3000 data, and the language "id". The data obtained during the crawling process is 2669 data and the total final data is 2470.

Table 1 Dataset

Query	Data Crawling	Data Akhir	Positif	Negatif
starlink indonesia	2669	2470	1405	1065

Data Selection

To perform sentiment analysis, it is necessary to sort out the attributes that will be used only. The attributes selected for this sentiment analysis process are attributes in the full_text column or public opinion.

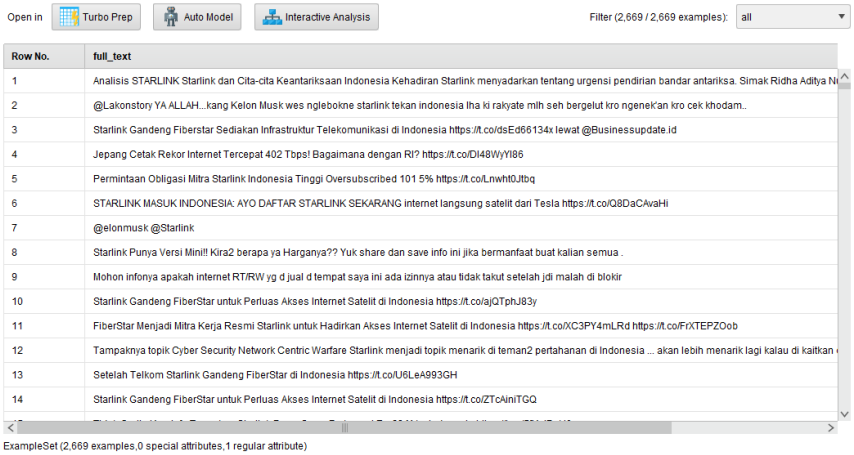


Fig. 3 Data Selection

Cleansing

This cleansing process aims to eliminate unnecessary words, characters and symbols. In this research process, it is carried out in a subprocess to manage processes such as RT (retweet), URL, Hastag, Mention, Symbol, and Trim. Everything is done in the subprocess.

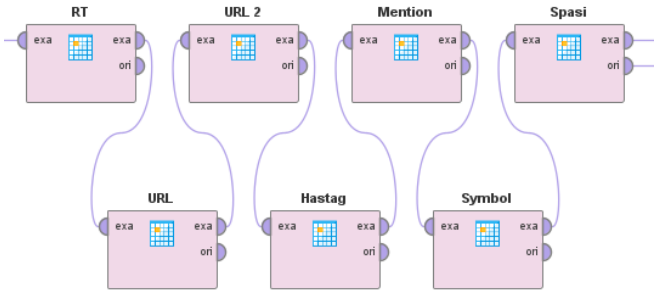


Fig. 4 Process Cleansing

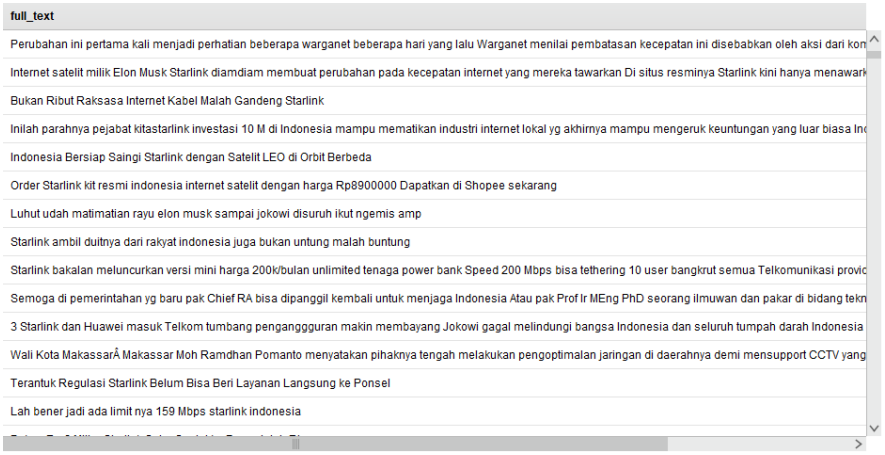
Process of cleansing using several operators including, Replace RT to remove words from retweets, Replace URL 1 and URL 2 are used to remove links. Replace Hastag is used to remove the # symbol, Replace Mention is used to remove the @ symbol. Replace Symbol is used to remove symbols that are not needed, Replace Trim is used to remove excess spaces. Here are the results:

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

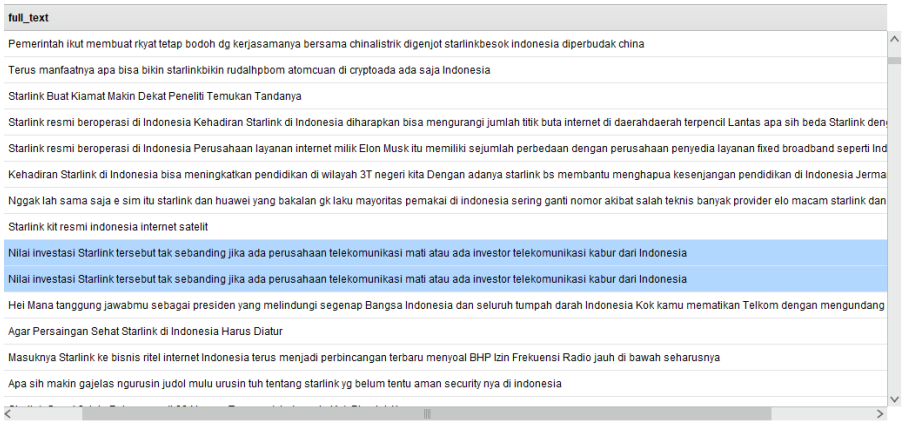


full_text
Perubahan ini pertama kali menjadi perhatian beberapa warganet beberapa hari yang lalu Warganet menilai pembatasan kecepatan ini disebabkan oleh aksi dari kon
Internet satelit milik Elon Musk Starlink diamdiam membuat perubahan pada kecepatan internet yang mereka tawarkan Di situs resminya Starlink kini hanya menawark
Bukan Ribut Raksasa Internet Kabel Malah Gandeng Starlink
Inilah parahnya pejabat kitastarlink investasi 10 M di Indonesia mampu mematikan industri internet lokal yg akhirnya mampu mengeruk keuntungan yang luar biasa In
Indonesia Bersiap Saingi Starlink dengan Satelit LEO di Orbit Berbeda
Order Starlink kit resmi indonesia internet satelit dengan harga Rp8900000 Dapatkan di Shopee sekarang
Luhut udah matimatian rayu elon musk sampai jokowi disuruh ikut ngemis amp
Starlink ambil duitnya dari rakyat indonesia juga bukan untung malah buntung
Starlink bakalan meluncurkan versi mini harga 200k/bulan unlimited tenaga power bank Speed 200 Mbps bisa tethering 10 user bangkrut semua Telkomunikasi provic
Semoga di pemerintahan yg baru pak Chief RA bisa dipanggil kembali untuk menjaga Indonesia Atau pak Prof Ir MEng PhD seorang ilmuwan dan pakar di bidang tekn
3 Starlink dan Huawei masuk Telkom tumbang pengangggguran makin membayang Jokowi gagal melindungi bangsa Indonesia dan seluruh tumpah darah Indonesia
Wali Kota MakassarÂ Makassar Moh Ramdhan Pomanto menyatakan pihaknya tengah melakukan pengoptimalan jaringan di daerahnya demi mensupport CCTV yang
Terantuk Regulasi Starlink Belum Bisa Beri Layanan Langsung ke Ponsel
Lah bener jadi ada limit nya 159 Mbps starlink indonesia

Fig. 5 Result Cleansing

Remove Duplicates

The purpose of this process is to filter out the same words, this happens because twitter has a retweet feature so it can cause many of the same tweets to be retweeted by other users and make the text repetitive or the text becomes the same.



full_text
Pemerintah ikut membuat rkyat tetap bodoh dg kerjasamanya bersama chinalistrik digenjot starlinkbesok indonesia diperbudak china
Terus manfaatnya apa bisa bikin starlinkbikin rudalhpbom atomcuan di cryptoada ada saja Indonesia
Starlink Buat Kiamat Makin Dekat Peneliti Temukan Tandanya
Starlink resmi beroperasi di Indonesia Kehadiran Starlink di Indonesia diharapkan bisa mengurangi jumlah titik buta internet di daerahdaerah terpencil Lantas apa sih beda Starlink den
Starlink resmi beroperasi di Indonesia Perusahaan layanan internet milik Elon Musk itu memiliki sejumlah perbedaan dengan perusahaan penyedia layanan fixed broadband seperti Ind
Kehadiran Starlink di Indonesia bisa meningkatkan pendidikan di wilayah 3T negeri kita Dengan adanya starlink bs membantu menghapua kesenjangan pendidikan di Indonesia Jerma
Nggak lah sama saja e sim itu starlink dan huawei yang bakalan gk laku mayoritas pemakai di indonesia sering ganti nomor akibat salah teknis banyak provider elo macam starlink dan
Starlink kit resmi indonesia internet satelit
Nilai investasi Starlink tersebut tak sebanding jika ada perusahaan telekomunikasi mati atau ada investor telekomunikasi kabur dari Indonesia
Nilai investasi Starlink tersebut tak sebanding jika ada perusahaan telekomunikasi mati atau ada investor telekomunikasi kabur dari Indonesia
Hei Mana tanggung jawabmu sebagai presiden yang melindungi segenap Bangsa Indonesia dan seluruh tumpah darah Indonesia Kok kamu mematikan Telkom dengan mengundang
Agar Persaingan Sehat Starlink di Indonesia Harus Diatur
Masuknya Starlink ke bisnis ritel internet Indonesia terus menjadi perbincangan terbaru menyoal BHP Izin Frekuensi Radio jauh di bawah seharusnya
Apa sih makin gajelas ngurusin judol mulu urusin tuh tentang starlink yg belum tentu aman security nya di indonesia

Fig. 6 Repeated Tweet Example

Filter Examples

When starting data processing, it is very common to encounter missing values. Missing Values are tweet data that is invalid or missing. To eliminate missing data, you can use the Filter Examples operator by specifying the parameter “is not missing”, meaning that in this study the data to be displayed is data that contains values only. The following is an example of missing data:

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

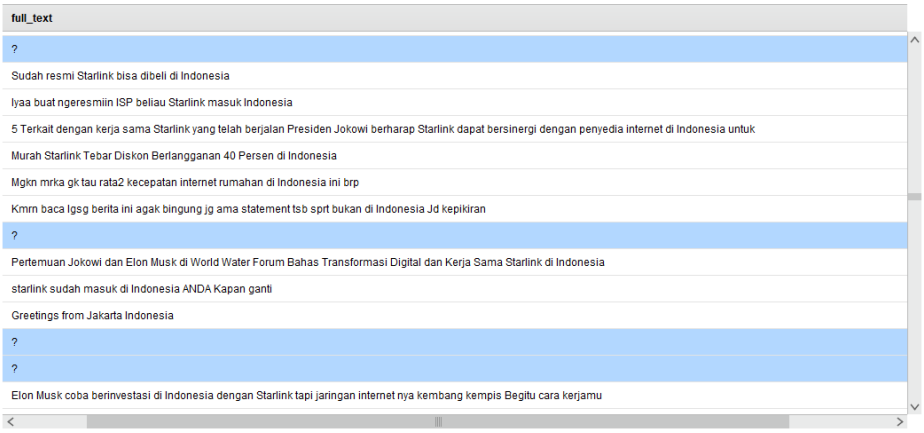


Fig. 7 Missing Values Example

Process Document from Data

This preprocessing will be applied to training data and testing data. The components used include Tokenize, Transform Cases, Filter Stopwords (Dictionary), Filter Tokens (by Length), Stem (Dictionary).

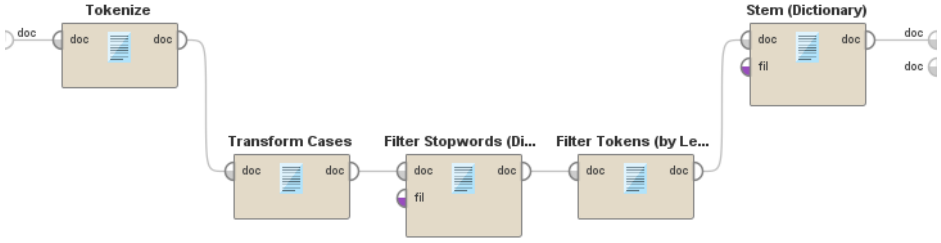


Fig. 8 Process Document from Data

Tokenize

Tokenizing is the process of cutting the input string document on each word constituent. At this stage, a sentence is broken down into word units.

Transform Cases

At this stage, all text in the processed document will be converted into lowercase letters.

Filter Stopwords (Dictionary)

At this stage, words that have no meaning or meaning when separated from other words will be removed so that sentiment analysis only focuses on words that have meaning.

Filter Tokens (by Length)

The Tokens Filter (by Length) process in this research is used to determine the short length of words that have been determined by researchers. The minimum words allowed are 4 words, while the maximum words allowed are 25 words.

Stem (Dictionary)

At this stage, a word is decomposed into basic words by removing the affixes contained in each word.

Support Vector Machine Classification Process

At this stage, the process of creating a classification model using the Support Vector Machine algorithm will be carried out which aims to predict class labels based on the training process. The training process uses training data that already has a label, then after that the results of training models that have patterns for classification can be used for the labeling process for testing data.

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

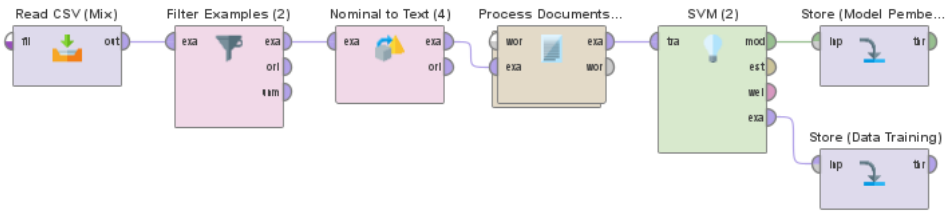


Fig. 9 Learning Model SVM Algorithm

Read CSV is used to read data from a file that we have previously saved in the form (.csv). Filter Examples is used to select data that has been labeled. Nominal to Text is used to convert data with text type. Procces Document from Data will be used for the preprocessing process. Support Vector Machine algorithm is used to learn the data pattern so that it can produce a classification model. Furthermore, the modeling is stored in the store (algorithm learning model) and the processing results are stored in the store (training data).

Support Vector Machine Classification Model

After the classification model has been formed, it can be used to predict the label of the testing data. Testing data is entered into the classification model. The testing data that has been entered will later be predicted by the classification model.

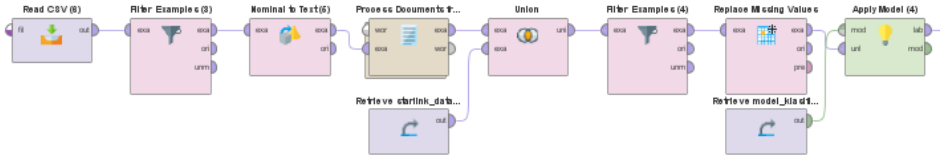


Fig. 10 SVM Algorithm Classification Model Process

Read CSV is used to read data from files that we have previously saved in the form (.csv). Filter Examples is used to select data that has not been labeled. Nominal to Text is used to convert text-type data. Procces Document from Data is used for the preprocessing process. Store (training data) the attributes in this store are combined with the attributes of the testing data. Union is used to combine the attributes of both and fill them in as Missing Values or pools that have no values. Filter Examples is used to remove sentiment data (training data) that is included. Replace Missing Values is used to replace the missing values in this data to zero. Store (algorithm learning model) the Support Vector Machine model previously created is used to analyze data that has not been labeled and Apply Model is used to apply the model that has been created previously. After all processes have been completed and successful, the results will be as shown below:

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

sentiment	prediction(s...	confidence(...	confidence(...	text	aaah	aaraf	abang	abangnya	al
?	Positif	0.634	0.366	hahahaha jok...	0	0	0	0	0
?	Positif	0.582	0.418	emangnye ka...	0	0	0	0	0
?	Positif	0.718	0.282	puskesmas ...	0	0	0	0	0
?	Positif	0.658	0.342	pakai internet...	0	0	0	0	0
?	Positif	0.666	0.334	starlink layan...	0	0	0	0	0
?	Positif	0.553	0.447	kwkkwkwkww...	0	0	0	0	0
?	Positif	0.651	0.349	dibukakan pi...	0	0	0	0	0
?	Positif	0.701	0.299	harga langga...	0	0	0	0	0
?	Negatif	0.451	0.549	maju trus ind...	0	0	0	0	0
?	Positif	0.669	0.331	sblm peluncu...	0	0	0	0	0
?	Positif	0.670	0.330	telkomsat sta...	0	0	0	0	0
?	Positif	0.605	0.395	bali elon mus...	0	0	0	0	0
?	Positif	0.658	0.342	terusyang ma...	0	0	0	0	0
?	Positif	0.563	0.437	elon musk be...	0	0	0	0	0

Fig. 11 SVM Algorithm Sentiment Prediction Results

DISCUSSIONS

Validation Using Cross Validation

The results of the process that has formed the classification model need to be validated to find out how well and also how accurate the results of the analysis have been carried out. Therefore, a validation using the 10-fold cross validation method is carried out using training data and testing of testing data where each experiment takes one different segment. The testing process uses the Support Vector Machine method and testing is done 10 times to find out the best accuracy, precision, and recall comparison results.

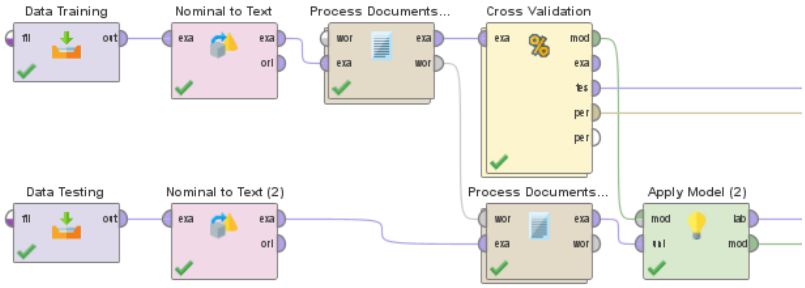


Fig. 12 Testing Process

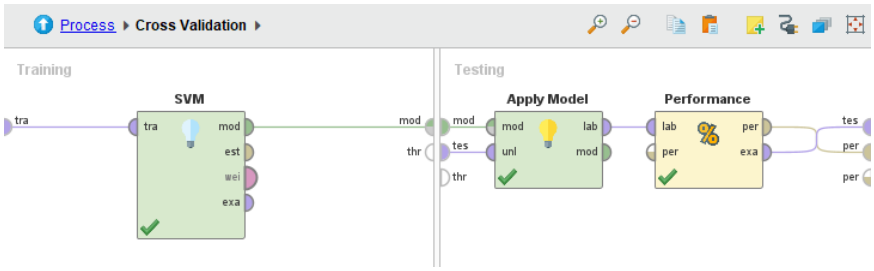


Fig. 13 Application of Support Vector Machine Algorithm

Read CSV (1) is used to read training data. Read CSV (2) is used to read testing data. Nominal to Text is used to convert text-type data. Procces Document from Data is used for preprocessing. Cross Validation has two separate sections, namely the training section used for the Support Vector Machine method and the testing section used for the Apply Model and Performance which functions to display the confusion matrix to see the results of accuracy, precision, and recall from the testing process.

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

Support Vector Machine Algorithm Testing Results

From the results of tests carried out using the Support Vector Machine algorithm, another test was carried out with the Cross Validation technique using K-Folds in order to get more accurate validation results. Then the results that have been obtained from the test results are as follows:

Table 2 Testing Results

K-Folds	*Accucary*	*Precision*	*Recall*
2	68,22%	70,07%	75,99%
3	71,92%	73,34%	78,69%
4	72,42%	74,42%	77,70%
5	74,04%	75,19%	80,40%
6	74,95%	76,03%	81,03%
7	76,22%	77,48%	81,38%
8	76,01%	77,22%	81,38%
9	75,05%	76,48%	80,40%
10	75,40%	76,57%	81,12%

Based on the results of the table above, the conclusion that can be drawn is that the highest accuracy, precision and recall values are in k-fold 7 with an accuracy value of 76.22%, precision 77.48% and recall 81.38%. From this it can be concluded that the use of Cross Validation techniques with k-folds is quite influential on the results that will be obtained. The following is a diagram of the test comparison results:

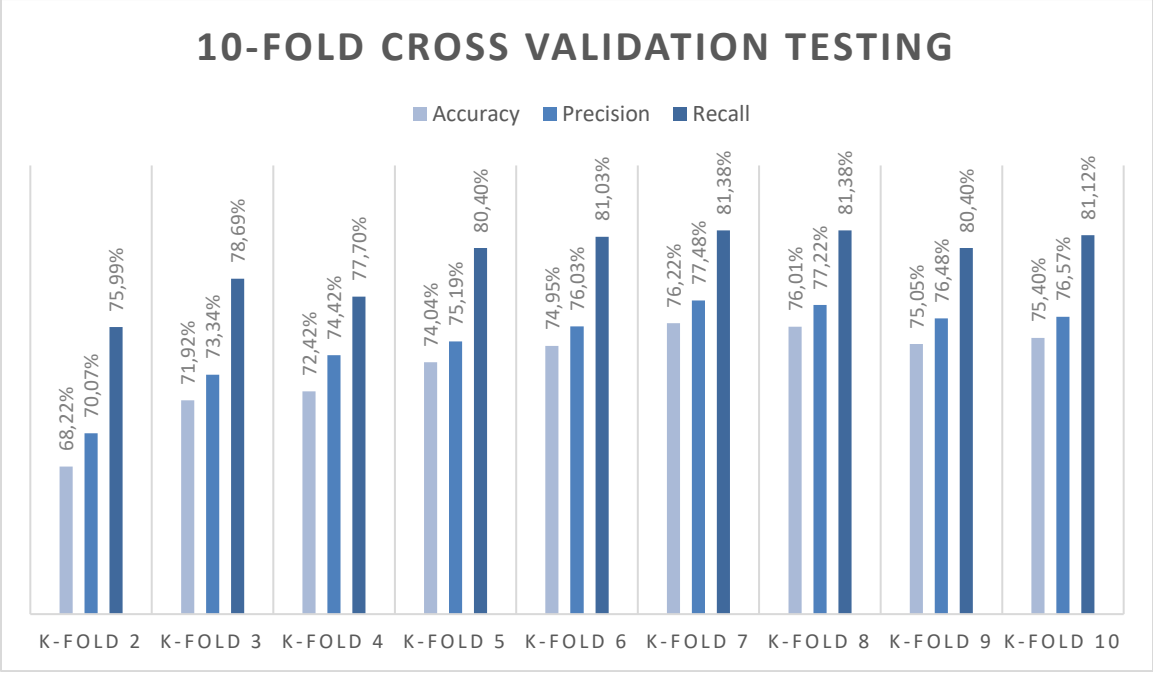


Fig. 14 Diagram of K-Fold Comparison Results

Sentiment Analysis Results

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

The results of the classification of public sentiment regarding the presence of starlink in Indonesia obtained from training data before and after testing with the Support Vector Machine algorithm. The following is a diagram of the results of the classification of public sentiment about starlink:

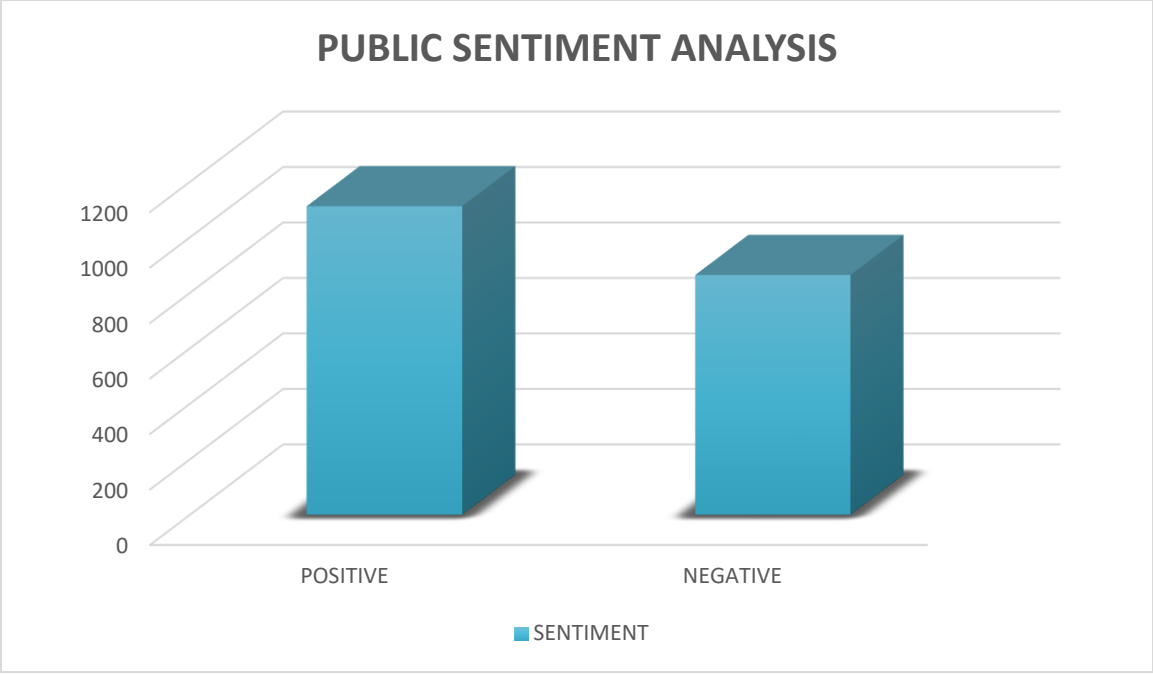


Fig. 15 Public Sentiment Analysis

From a total of 1976 tweets data tested using the Support Vector Machine method, there are 1112 tweets data containing positive sentiment and 864 tweets data containing negative sentiment. From the figure above, it can be concluded that the results of the classification of community sentiment from this test are superior to positive tweets with a percentage of 56.3% and negative tweets with a percentage of 43.7%.

Wordcloud

Next, the data visualization process is carried out using wordcloud based on the results of sentiment data classification. Wordcloud will display all the words in the dataset, so the word that has the highest frequency will have a larger font size in the wordcloud visualization. In this study, the results of wordcloud for all data regarding the presence of starlink in Indonesia can be seen in the following figure:



Fig. 16 Wordcloud

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 3, July 2024
https://doi.org/10.47709/cnapc.v6i3.4348

Submitted : 20 July 2024
Accepted : 26 July 2024
Published : 31 July 2024

5. CONCLUSION

Analysis of public sentiment towards the presence of starlink in Indonesia on twitter social media from a total of 1976 tweets data tested using the Support Vertor Machine algorithm, the results are 1112 tweets data containing positive sentiment and 864 tweets data containing negative sentiment. This shows that 56.3% of people agree with the presence of starlink in Indonesia. While from the use of the Support Vertor Machine algorithm, the resulting accuracy is quite good with the Cross Validation technique using the K-Fold value and then obtained the highest Accruracy, Precision, and Recall values in the 7th test or K-Fold value of 7, the value obtained for Accuracy is 76.22%, for Precision is 77.48% and for Recall is 81.38%.

6. REFERENCES

Agus Tri Haryanto. (2024, April 26). Dampak Baik-Buruk Starlink Elon Musk Jualan Internet di Indonesia.

Amrullah, A. Z., Sofyan Anas, A., Adrian, M., & Hidayat, J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal*, *2*(1). https://doi.org/10.30812/bite.v2i1.804

Ananda, F. D., & Pristyanto, Y. (2021). Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *20*(2), 407–416. https://doi.org/10.30812/matrik.v20i2.1130

Andy Nugroho. (2020, May 5). Sejarah ARPANET Sebagai Cikal Bakal Adanya Internet.

Darwis, D., Shintya Pratiwi, E., Ferico, A., & Pasaribu, O. (2020). PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA. In *Jurnal Ilmiah Edutic* (Vol. 7).

Ismail, A. R., & Raden Bagus Fajriya Hakim. (2023). Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Menentukan Rekomendasi Pantai Di DI Yogyakarta Berdasarkan Data Twitter. *Emerging Statistics and Data Science Journal*, *1*(1), 37–46. https://doi.org/10.20885/esds.vol1.iss.1.art5

Malik Zuhdi, A., Utami, E., & Raharjo, S. (2019). *ANALISIS SENTIMENT TWITTER TERHADAP CAPRES INDONESIA 2019 DENGAN METODE K-NN* (Vol. 5).

Muhammad Arif. (2024, February 7). APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang.

Muhammad Fuad. (2024, May 20). Starlink: Revolusi Konektivitas Internet di Indonesia.

Putra, A., Haeirudin, D., Khairunnisa, H., & Latifah, R. (2021). *Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm*.

Rangga, M., Nasution, A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *JURNAL INFORMATIKA*, *6*(2), 212–218. Retrieved from http://ejournal.bsi.ac.id/ejurnal/index.php/ji

Sekretariat KADIN Indonesia. (2024, June 7). Dampak dan Manfaat Kehadiran Starlink.

Slamet, R., Gata, W., Novtariany, A., Hilyati, K., & Jariyah, F. A. (2022). Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai Brand Ambassador Produk Kecantikan Lokal. *INTECOMS: Journal of Information Technology and Computer Science*, *5*(1), 145–153. https://doi.org/10.31539/intecoms.v5i1.3933

Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT*, *6*(3), 150–155. https://doi.org/10.30591/jpit.v6i3.2870

Utami, D. S., & Erfina, A. (2021). *ANALISIS SENTIMEN PINJAMAN ONLINE DI TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)*.

* Corresponding author

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.