Sentiment Analysis of 2024 Presidential Candidates Election Using SVM Algorithm

Michael Alfonso[1*)], Dionisia Bhisetya Rarasati, S.Kom., M.T.I[2]
[1]Program Studi Teknik Informatika, Fakultas Teknologi dan Desain, Universitas Bunda Mulia
[2] Program Studi Teknik Informatika, Fakultas Teknologi dan Desain, Universitas Bunda Mulia
Email: [1]s32190039@student.ubm.ac.id, [2]l1783@lecturer.ubm.ac.id

Abstract − Elections for presidential candidates are held every 5 years with various candidates, especially on Twitter, arguments about political matters often occur that many Twitter users participate in discussions about the election for presidential candidate. Therefore, this study focuses on sentiment analysis to infer user responses to the presidential election and validate it by looking for a correlation between electability survey results and Twitter sentiment data using Pearson Correlation. In sentiment analysis model, the 10-Fold Cross Validation method is used to find the best model from a dataset with a division of training data and test data with 90:10 split. Then the alphabetic data will be converted into numeric data using the TF-IDF weighting method. To validate the best model, Confusion Matrix is used to get the best f1-score. The model is using Support vector machine algorithm with the Gaussian RBF (Radial Basis Function) kernel. The results of the analysis are compared with the results of the news portal electability survey which contains the 3 candidates using Pearson Correlation. This study produces the best fold for each data on each presidential candidate with the f1-score to find the best model for each fold. In the Peason Correlation result, the higher positive sentiment of each presidential candidate, the higher electability survey data. For further research, research can be discuss about hyper tuning parameters and using other kernels on Support vector machine algorithm.

Keywords – NLP, Pearson Correlation, Sentiment analysis, SVM, TF-IDF

I. INTRODUCTION

Elections occur every 5 years with the number of voters increasing each period, with an increase of 28.90% from 2004 to 2019 [1]. Social media such as Facebook, Twitter, LinkedIn, YouTube have become part of human life, it is proven that most interactions can be done with social media, until it becomes self-identity to the general public [2]. On social media, especially Twitter, there is a lot of dissemination of information, entertainment, and political opinion. This is not spared from the election which is a political activity that will be held in 2024, referred to as a democratic party, this is certainly a lot of discussion on social media. One way to convey opinions and through social media is to use certain hashtags in tweets posted on Twitter. Twitter itself is used because of the ease in replying to tweets with likes, comments and even replies, so that topics that are currently being discussed can be easily spread [3]. Therefore, further analysis is needed to see how much Twitter users agree and disagree with the 2024 presidential election, then data will be obtained from Twitter user tweets related to this topic, namely the 2024 general election. Sentiment analysis or Opinion Mining is a study that learn about opinions, sentiments, evaluations, behavior and emotions towards an entity [2], [3]. Sentiment analysis is used to find out which tweets are posted containing positive, negative or neutral words [4]. In this study, the Support vector machine (SVM) method will be used, which is one of the classification methods in machine learning (supervised learning) that predicts classes based on patterns from training results. Classification is done by creating a hyperplane between positive and negative classes [2]. The results of this study are the sentiments of Twitter users towards the presidential election in 2024.

II. RESEARCH METHODOLOGY

This research consists of 6 processes in outline, namely data collection, preprocessing, Term Frequency-Inverse Document Frequency weighting, Support Vector Machine training using 10-Fold Cross Validation, Choosing the best fold to be used as a model, and testing as shown below.
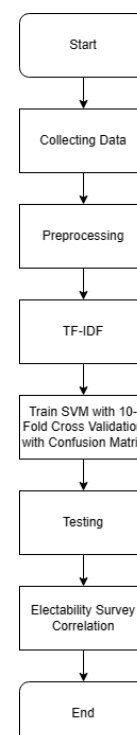

Figure 1. Research Flow

A. Collection Data

The first stage of this research is to scrape data sourced from Twitter with the keyword "(Name of Candidate) President" to narrow down the context of a Twitter user's Tweet. The data collection process is carried out with the help of the snscrape library by python. Data was collected from 1 January 2023 to 31 March 2023. This scraping process resulted in a total of 18,267 data with 7,810 data that were duplicate and irrelevant to the theme "2024 Presidential Candidate Election" resulting in a total of 10,457 clean data from 3 presidential candidates.

B. Preprocessing

The pre-processing stage functions to clean and transform data so that it can be processed by the system easily and accurately. Several stages of pre-processing in this study include:

a. Lowercasing

Lowercasing is a step to change all letters to lowercase so that the word being processed becomes the same form. Table 1 shows an example of the lowercasing process from some of Anies Baswedan's tweets:

Table 1. Lowercasing Result

No.	Before Process	After Process
1	PKS Mendukung Anies Baswedan Menjadi Calon Presiden 2024	pks mendukung anies baswedan menjadi calon presiden 2024
2	Anies Baswedan Presiden Indonesia	anies baswedan presiden indonesia
3	Walau sekarang didukung Demokrat dan PKS, ada empat alasan Anies Baswedan gak bakal bisa menang jadi presiden !!	walau sekarang didukung demokrat dan pks, ada empat alasan anies baswedan gak bakal bisa menang jadi presiden !!
4	Anies Baswedan lh yg pantas jadi presiden	anies baswedan lh yg pantas jadi presiden
5	Pak Anies Baswedan calon presiden 2024	pak anies baswedan calon presiden 2024

b. Remove Special Character and Number

This stage functions to produce sentences without punctuation, emoticons, and numbers. Table 2 shows an example of the Remove Special Character and Number process:

Table 2. Remove Special Character and Number Result

No.	Before Process	After Process
1	pks mendukung anies baswedan menjadi calon presiden 2024	pks mendukung anies baswedan menjadi calon presiden
2	anies baswedan presiden indonesia	anies baswedan presiden indonesia
3	walau sekarang didukung demokrat dan pks, ada empat alasan anies baswedan gak bakal bisa menang jadi presiden !!	walau sekarang didukung demokrat dan pks ada empat alasan anies baswedan gak bakal bisa menang jadi presiden
4	anies baswedan lh yg pantas jadi presiden	anies baswedan lh yg pantas jadi presiden
5	pak anies baswedan calon presiden 2024	pak anies baswedan calon presiden

c. Stemming

At this stage, all affixed words will be processed to become basic words, as shown in table 3 below:

Table 3. Stemming Result

No.	Before Process	After Process
1	pks mendukung anies baswedan menjadi calon presiden	pks dukung anies baswedan jadi calon presiden
2	anies baswedan presiden indonesia	anies baswedan presiden indonesia
3	walau sekarang didukung demokrat dan pks ada empat alasan anies baswedan gak bakal bisa menang jadi presiden	walau sekarang dukung demokrat dan pks ada empat alasan anies baswedan gak bakal bisa menang jadi presiden
4	anies baswedan lh yg pantas jadi presiden	anies baswedan lh yg pantas jadi presiden
5	pak anies baswedan calon presiden	pak anies baswedan calon presiden

d. Tokenize

This stage break sentences into tokens per word. This process is carried out with the aim that the next process can be applied at the word level. The process is shown in table 4 below:

Table 4. Tokenize Result

No.	Before Process	After Process
1	pks dukung anies baswedan jadi calon presiden	[[pks] [dukung] [anies] [baswedan] [jadi] [calon] [presiden]]
2	anies baswedan presiden indonesia	[[anies] [baswedan] [presiden] [Indonesia]]
3	walau sekarang dukung demokrat dan pks ada empat alasan anies baswedan gak bakal bisa menang jadi presiden	[[walau] [sekarang] [dukung] [demokrat] [dan] [pks] [ada] [empat] [alasan] [anies] [baswedan] [gak] [bakal] [bisa] [menang] [jadi] [presiden]]
4	anies baswedan lh yg pantas jadi presiden	[[anies] [baswedan] [lh] [yg] [pantas] [jadi] [presiden]]
5	pak anies baswedan calon presiden	[[pak] [anies] [baswedan] [calon] [presiden]]

e. Remove Stopwords

Stopwords are for removing words that don't have an impact on sentiment, but often appear in sentences. Table 5 shows the Remove Stopwords process as follows:

Table 5. Remove Stopwords Result

No.	Before Process	After Process
1	[[pks] [dukung] [anies] [baswedan] [jadi] [calon] [presiden]]	[[pks] [dukung] [anies] [baswedan] [calon] [presiden]]
2	[[anies] [baswedan] [presiden] [Indonesia]]	[[anies] [baswedan] [presiden] [Indonesia]]
3	[[walau] [sekarang] [dukung] [demokrat] [dan] [pks] [ada] [empat] [alasan] [anies] [baswedan] [gak] [bakal] [bisa] [menang] [jadi] [presiden]]	[[sekarang] [dukung] [demokrat] [pks] [empat] [alasan] [anies] [baswedan] [gak] [bakal] [menang] [presiden]]
4	[[anies] [baswedan] [lh] [yg] [pantas] [jadi] [presiden]]	[[anies] [baswedan] [lh] [pantas] [presiden]]
5	[[pak] [anies] [baswedan] [calon] [presiden]]	[[pak] [anies] [baswedan] [calon] [presiden]]

JISA (Jurnal Informatika dan Sains) (e-ISSN: 2614-8404) is published by Program Studi Teknik Informatika, Universitas Trilogi
under Creative Commons Attribution-ShareAlike 4.0 International License.

2

f. *Slangwords Filtering*

This stage removes slang words into standard words. The slang dictionary is taken from the github repository posted by louisowen6 which is shown in Table 6 as follows:

Table 6. Slangwords Result

No.	Before Process	After Process
1	[[pks] [dukung] [anies] [baswedan] [calon] [presiden]]	[[pks] [dukung] [anies] [baswedan] [calon] [presiden]]
2	[[anies] [baswedan] [presiden] [Indonesia]]	[[anies] [baswedan] [presiden] [Indonesia]]
3	[[sekarang] [dukung] [demokrat] [pks] [empat] [alasan] [anies] [baswedan] [gak] [bakal] [menang] [presiden]]	[[sekarang] [dukung] [demokrat] [pks] [empat] [alasan] [anies] [baswedan] [tidak] [akan] [menang] [presiden]]
4	[[anies] [baswedan] [lh] [pantas] [presiden]]	[[anies] [baswedan] [pantas] [presiden]]
5	[[pak] [anies] [baswedan] [calon] [presiden]]	[[pak] [anies] [baswedan] [calon] [presiden]]

g. *Remove Synonim*

At this stage, the writer removes synonyms in the same sentence if any. Table 7 shows the Remove Synonim process as follows:

Table 7. Remove Synonim Result

No.	Before Process	After Process
1	[[pks] [dukung] [anies] [baswedan] [calon] [presiden]]	[[pks] [dukung] [anies] [baswedan] [calon] [presiden]]
2	[[anies] [baswedan] [presiden] [Indonesia]]	[[anies] [baswedan] [presiden] [Indonesia]]
3	[[sekarang] [dukung] [demokrat] [pks] [empat] [alasan] [anies] [baswedan] [tidak] [akan] [menang] [presiden]]	[[sekarang] [dukung] [demokrat] [pks] [empat] [alasan] [anies] [baswedan] [tidak] [akan] [menang] [presiden]]
4	[[alhamdulillah] [mudah] [mudah] [anies] [baswedan] [takdir] [allah] [swt] [presiden] [aamiin] [yra]	[[alhamdulillah] [mudah] [anies] [baswedan] [takdir] [allah] [swt] [presiden] [aamiin] [yra]
5	[[pak] [anies] [baswedan] [calon] [presiden]]	[[pak] [anies] [baswedan] [calon] [presiden]]

C. *TF-IDF*

TF-IDF weighting is a method of assessing the weight of each word in a document, TF-IDF is a strong feature weighting method that measures the frequency and distribution of words to determine their weight. TF-IDF consists of TF (Term Frequency), which is the frequency of occurrence of words in documents and IDF (Inverse Document Frequency), which is the opposite of the frequency of documents, the TF-IDF value can be obtained by using the following equation:

a. The value of Term Frequency (TF) is obtained by the value of the frequency of appearance of the t feature in document d[5]

$$TF_t = (t, d) \qquad (1)$$

b. The Inverse Document Frequency (IDF) value is obtained from the logarithm of the number of documents n divided by df documents containing t features[6]

$$IDF_t = log \frac{n}{df(t)} + 1 \qquad (2)$$

c. The value of the Term Frequency Inverse Document Frequency (TF-IDF) is obtained by multiplying the TF and IDF values[7]

$$W_t = TF_t \, . \, IDF_t \qquad (3)$$

TF-IDF functions to convert data into numerical form so that the features that have been processed can be loaded into the equation. So this method was chosen to extract the Twitter data.

D. *Support Vector Machine*

The algorithm chosen is the Support vector machine because the Support vector machine is a text mining algorithm that has high accuracy [8]. As seen in this journal, the Support vector machine algorithm has higher accuracy, precision, and recall compared to the Naïve Bayes algorithm, then for journal [9] it is known that the accuracy, precision, and recall of the Support vector machine algorithm has higher accuracy than the K -Nearest Neighbor (K-NN) algorithm and Naïve Bayes. Support Vector Mahine was created by Boser, Guyon, and Vapnik [10] which is a classification method for finding the best hyperplane to separate 2 classes [11]. This algorithm will produce a hyperplane that separates a plane into 2 parts in each class which can be seen in figure 2.
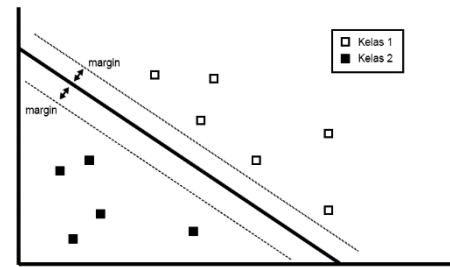

Figure 2. Hyperplane Support Vector Machine Example

Several kernels that can be applied to the SVM algorithm:

a. Linear Kernel is a kernel function that separates 2 classes linearly with the following equation:

$$[(w^t . x_i) + b] \geq 1 \, for \, y_i = 1 \, dan \, [(w^t . x_i) + b] \leq -1 \, for \, y_i = -1 \qquad (4)$$

Description of the following equation is x_i as the train data, followed by i as ammount of data (ex: 1,2, …, n), then y_i as the label of train data [10].

b. Polynomial Kernel is a kernel function that is not linearly separated, the equation is as follows:

$$K(x_i, x_j) = (x_i, x_j) + c)^d \qquad (5)$$

Description of the following equation is x_i, x_j as the train data, followed by c and d as constant [10].

c. Radial Basis Function (RBF) The kernel is a function for analyzing data that is not linearly separated and requires 2 types of parameters (Gamma and Cost).

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (6)$$

Description of the following equation is $\|x_i - x_j\|$ as euclidean distance and 2σ as independent parameter which define substraction level of RBF toward 0 [10].

To use the Support vector machine algorithm, it is necessary to find the best $\vec{w}$ to separate each class. Normally with data that is not spread out, the hyperplane is either linear or polynomial. But if it comes to the author's case, the data distribution is very wide and the feature dimensions are unlimited, then it is necessary to use a kernel trick, the implementation that the writer has done to overcome this problem is to use a Gaussian RBF (Radial Basis Function) which is able to overcome the problem above.

E. 10-Fold Cross Validation

K-Fold Cross Validation is a form of testing that serves to assess the performance of the algorithm by dividing the sample data randomly and then dividing the data as much as the K value of the total fold. Furthermore, one of the K groups will be used as data testing and the rest of the group will be used as training data [12].

After the SVM machine is trained, the best model is obtained by finding the best f1-score value from the SVM model, 10-Fold Cross Validation method divides a dataset into 10 parts with a data ratio 10:90.

Next, it will be seen from the 10 data divisions, which division has the best f1-score to take the model from the fold

F. Testing

The Confusion Matrix is a table for evaluating classification models, the f1-score value is the value used to determine the best model [13]. So I apply evaluation to research using the Confusion Matrix method. The application of the Confusion Matrix can be seen in the following table:

Table 8. Confusion Matrix Table

	Prediksi Positif	Prediksi Netral	Prediksi Positif
Label Positif	True Positif	False Netral	False Negatif
Label Netral	False Positif	True Netral	False Negatif
Label Negatif	False Positif	False Netral	True Negatif

After that, look for precision, recall, and f1-score values obtained by using the equation:

$$Precision_{(A)} = \frac{TP_A}{TP_A + FP_A} \quad (7)$$

$$Recall_{(A)} = \frac{TP_A}{TP_A + FN_A} \quad (8)$$

$$f1_{(A)} = \frac{2 * Precision_{(A)} * Recall_{(A)}}{Precision_{(A)} + Recall_{(A)}} \quad (9)$$

Description of the following equation is TP_A as the True Positive data, FP_A as the False Positive data, and FN_A as False Negative.

G. Electability Correlation

Pearson Correlation is a method that aims to find a linear relationship between 2 or more variables. The results of the Pearson Correlation are correlation coefficients ranging from numbers 0 to 1 which apply to negative numbers [10]. The Pearson Product Moment Correlation is shown in equation as follows:

$$r_{x,y} = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{\{n \sum x^2 - (\sum x)^2\}\{n \sum y^2 - (\sum y)^2\}}} \quad (10)$$

Description of the following equation is n as the amount of data, $\sum xy$ as sum of pair X and Y, $\sum x$ as sum of X variable, $\sum y$ as sum of Y variable, $\sum x^2$ as sum of square X, and $\sum y^2$ as sum of square Y.

III. RESULTS AND DISCUSSION

The implementation of the SVM algorithm is carried out exactly as previously attached, the development of the SVM system begins with the data cleaning stage, then transforms the data into numeric form, then conducts model training by dividing the data 10:90 using the 10-Fold Cross Validation method, and testing the best model by taking f1-score value with the Confusion Matrix method. The best fold is determined based on the f1-score when the False Negative (FN) and False Positive (FP) data are not very close [14]. Also, to find the F1-score value in models with more than 2 labels, I use the macro f1, where the value taken is the average of the f1-score values for all classes, the authors use the macro value because the values will be more even for datasets that are not balanced [15].

Next the results are obtained of the training process by finding the model with the best fold, followed by an analysis of the linkages of the predicted model results using the Pearson Moment Product Correlation. The test results of the 3 presidential candidates are shown in the following table:

Table 9. Best Fold Test Results for Every Presidential Candidate

Candidate Name	Best Fold	F1-score
Anies Baswedan	8	66,36%
Ganjar Pranowo	5	71,72%
Prabowo Subianto	4	77,99%

Then there are the results of electability survey data every month for each candidate in the following table:

Table 10. Electability Survey and Sentiment

Candidate Name	Month	Sentiment		Electability Score
		Positive	*Negative*	
	January	88,11%	11,89%	15%

Candidate Name	Month	Sentiment		Electability Score
		Positive	*Negative*	
Anies Baswedan	February	90,69%	9,31%	16,95%
	March	91,92%	8,08%	23,5%
Ganjar Pranowo	January	96,22%	3,78%	27,58%
	February	97,30%	2,70%	23,65%
	March	94,32%	5,68%	28,83%
Prabowo Subianto	January	99,05%	0,95%	21,68%
	February	99,23%	0,77%	24,62%
	March	99,79%	0,21%	23,12%

Result of Pearson Product Moment Correlation of data above shown in the following table:

Candidate Name	Correlation	
	Positive	*Negative*
Anies Baswedan	0,876	-0,876
Ganjar Pranowo	0,894	-0,894
Prabowo Subianto	0,97	-0,97

The result obtained by the correlation has a strong unidirectional relationship. It can be concluded that there is a relationship between sentiment and news portal electability surveys, from the 3 presidential candidates that the higher the positive sentiment, the higher the electability of the candidate.

IV. CONCLUSION

Based on the results of the research that has been done, it can be concluded that the implementation of the Sentiment Analysis Engine with the Support Vector Machine algorithm using the Gaussian RBF kernel found the best fold for each presidential candidate, namely the 8th fold with an f1-score of 0.66 for Anies Baswedan's candidate with a total 2,554 training data and 283 testing data, the 5th fold with an f1-score of 0.72 for the Ganjar Pranowo candidate with a total of 3,330 training data and 370 testing data, and the 4th fold with an f1-score of 0.78 for the Prabowo Subianto candidate with a total of 3487 training data and 387 testing data.

Also, the correlation between the electability survey and the results of the positive sentiment analysis of presidential candidate Anies Baswedan is 0.876 and the negative sentiment is -0.876. So that the conclusion obtained by the correlation has a strong relationship in the same direction. Then for the correlation between the electability survey and the results of the analysis of positive sentiment for presidential candidate Ganjar Pranowo is 0.894 and negative sentiment is -0.894. So that the conclusion obtained by the correlation has a strong relationship in the same direction. Furthermore, the correlation between the electability survey and the results of positive sentiment analysis for presidential candidate Prabowo Subianto is 0.97 and negative sentiment is -0.97. So that the conclusion obtained by the correlation has a strong relationship in the same direction. So it can be concluded that there is a link between sentiment and news

portal electability surveys. It was concluded from the 3 presidential candidates that the higher the positive sentiment, the higher the electability of the candidate.

REFERENCES

[1] BPS, "Hasil Perhitungan Suara Sah Pemilu Presiden dan Wakil Presiden Menurut Provinsi Tahun 2004 , 2009 , 2014, 2019." https://www.bps.go.id/statictable/2009/03/04/1574/hasil-perhitungan-suara-sah-pemilu-presiden-dan-wakil-presiden-menurut-provinsi-tahun-2004-2009-2014-2019.html (accessed Jul. 25, 2023).

[2] B. W. Sari and F. F. Haranto, "IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PELAYANAN TELKOM DAN BIZNET," *Jurnal Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 171–176, Sep. 2019, doi: 10.33480/pilar.v15i2.699.

[3] Fatihah Rahmadayana and Yuliant Sibaroni, "Sentiment Analysis of Work from Home Activity using SVM with Randomized Search Optimization," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 5, pp. 936–942, Oct. 2021, doi: 10.29207/resti.v5i5.3457.

[4] S. Fendyputra Pratama, R. Andrean, and A. Nugroho, "Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 4, no. 2, pp. 2541–3619, 2019, doi: 10.31328/jo.

[5] D. W. Seno and A. Wibowo, "Analisis Sentimen Data Twitter Tentang Pasangan Capres-Cawapres Pemilu 2019 Dengan Metode Lexicon Based Dan Support Vector Machine," *Jurnal Ilmiah FIFO*, vol. 11, no. 2, p. 144, Nov. 2019, doi: 10.22441/fifo.2019.v11i2.004.

[6] D. Darwis, E. Shintya Pratiwi, A. Ferico, and O. Pasaribu, "PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA," 2020.

[7] A. S. Arief, "SENTIMENTANALYSIS REVIEW APLIKASI MENGGUNAKAN ALGORITMA SVM PADA APLIKASI MYPERTAMINA," 2023.

[8] S. Diantika, W. Gata, H. Nalatissifa, and M. Lase, "Komparasi Algoritma SVM Dan Naive Bayes Untuk Klasifikasi Kestabilan Jaringan Listrik," *JURNAL ILMIAH ELEKTRONIKA DAN KOMPUTER*, vol. 14, no. 1, pp. 10–15, 2021.

[9] R. Risnantoyo, A. Nugroho, and K. Mandara, "Sentiment Analysis on Corona Virus Pandemic Using Machine Learning Algorithm," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 4, no. 1, pp. 86–96, Jul. 2020, doi: 10.31289/jite.v4i1.3798.

[10] D. B. Rarasati and J. C. A. Putra, "Correlation Between Twitter Sentiment Analysis with Three Kernels Using Algorithm Support Vector Machine

(SVM) Governor Candidate Electability Level," *COIESE*, pp. 249–256, 2021.

[11] F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan, "IMPLEMENTASI TWITTER SENTIMENT ANALYSIS UNTUK REVIEW FILM MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," *Jurnal Informatika Polinema*, vol. 4, no. 2, pp. 93–100, 2018.

[12] A. Hutapea and M. Tanzil Furqon, "Penerapan Algoritme Modified K-Nearest Neighbour Pada Pengklasifikasian Penyakit Kejiwaan Skizofrenia," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, pp. 3957–3961, 2018, [Online]. Available: http://j-ptiik.ub.ac.id

[13] F. Istighfarizkya, N. A. S. ER, I. M. Widiarthaa, L. G. Astutia, I. G. N. A. C. Putra, and I. K. G. Suhartana, "Klasifikasi Jurnal menggunakan Metode KNN dengan Mengimplementasikan Perbandingan Seleksi Fitur," *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 11, pp. 167–176, 2022, [Online]. Available: https://scholar.google.com

[14] F. Satria, Zamhariri, and M. A. Syaripudin, "Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung," *Jurnal Ilmiah MATRIK*, vol. 22, pp. 28–35, 2020.

[15] O. H. Anidjar, A. Barak, B. Ben-Moshe, E. Hagai, and S. Tuvyahu, "A Stethoscope for Drones: Transformers Based Methods for UAVs Acoustic Anomaly Detection," *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3262702.