Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. 5. No. 1; e-ISSN: 2808-4519

Analysis of Football Supporters' Sentiment on Social Media on PSSI'S Performance Using the K-Nearest Neighbor Method

Chintya Dwi Putri Br Ginting [1]*, Imran Lubis[2], Rusmin saragih[3]

[1,2,3]*STMIK Kaputama*

Chintyadwiputri5@gmail.com[1*], imran.loebis.medan@gmail.com[2], evitha12014@gmail.com[3]

Abstract

The performance of the Football Association of Indonesia (PSSI) often receives public scrutiny, especially from football supporters. The dynamics of Indonesian football, which are frequently colored by controversy, have generated a large number of opinions on social media. This study aims to analyze the sentiment of football supporters on social media regarding PSSI's performance using the K-Nearest Neighbor (KNN) method. The research data were collected from Twitter through a crawling process, with word weighting performed using the TF-IDF method, while the KNN model was tested with the parameter value of k = 3. The results show that the K-Nearest Neighbor (KNN) model achieved an accuracy of 93.5%, with a precision of 63.2%, recall of 52.9%, and an f1-score of 56.5%. However, the model's performance was influenced by data imbalance, where neutral sentiment comments were far more dominant than positive or negative ones. The sentiment distribution indicates that public opinion on social media was largely neutral, while the proportion of positive and negative sentiments was relatively smaller. These findings suggest that although criticisms of PSSI's performance were quite prevalent, most supporters tended to remain neutral in expressing their opinions.

Keywords: *Sentiment Analysis, K-Nearest Neighbor, PSSI, Twitter*

1. Introduction

Football is one of the most loved sports and has many fans in many countries, including Indonesia. 77% of Indonesians love football and Indonesia ranks second in the world as the most football fans, according to Nielsen Sport research [1]. The Indonesian Football Association (PSSI) is responsible for the development and development of football in Indonesia as a National football organization [2] However, PSSI's performance is often in the public spotlight, especially on social media, where football fans are active in voicing their opinions regarding PSSI's performance.

Along with the increase in the use of social media, the number of opinions uploaded by netizens regarding PSSI's performance is also increasing. The number of netizens' opinions on social media, especially football supporters, regarding PSSI's performance certainly makes it difficult to distinguish positive, negative and neutral responses from social media users. Therefore, an analysis is needed to find out the differences in responses that netizens have systematically, namely sentiment analysis [3] Sentiment analysis is the process of processing text to determine a person's attitude or feelings towards a certain topic. For sentiment analysis, it can be done using *the K-Nearest Neighbor* (KNN) method.

The *K-Nearest Neighbor* (KNN) method is one of the techniques in machine learning and is widely used for data classification, including sentiment analysis. The KNN method was chosen because of its ability to classify simply without having to perform complex calculations [4] The basic concept of the KNN method is to find the closest distance between the data to be evaluated and the nearest neighbor in the training data. Distance calculation was carried out using the Euclidean concept [5] So the KNN method is suitable for analyzing the sentiment of football supporters on social media on PSSI's performance.

2. Method

Problem solving methods are a series of systematic processes used to analyze and solve a problem contained in a research.

2.1. Analisis Sentimen

Sentiment analysis is a technique used by internet users on social media to provide their personal assessments or opinions by extracting information about positive, neutral, or negative sentiments from text data [6]

2.2. CRISP-DM

CRISP-DM (Cross-Industry Standard Process for Data Mining) is a standard methodology widely used in data mining project development. This model provides a systematic and structured framework for efficient data analysis, from problem understanding to implementation of the results [7] .

2.3. K-Nearst Neigbor(KNN)

K-Nearest Neighbor (KNN) is one of the methods used in data classification. The working principle of *K-Nearest Neighbor* is to classify data based on the proximity of the distance of one data to other data. To use the *K-Nearest Neighbors algorithm,* it is necessary to determine the number of *K-Nearest* Neighbors used to classify new data. The number of *K* should be an odd number, for example K = 1, 2, 3, and so on[8] [9].

2.4. Social Media

Social media is an internet-based platform that allows users to interact, share information, and form networks virtually. Social media is an online medium that allows users to easily participate, share, and create content, such as blogs, social networks, wikis, forums, and virtual worlds. Social media is not only a means of social interaction, but also a platform to build professional networks, promote business, and disseminate information widely [10]

2.5. PSSI

The Indonesian Football Association (PSSI) is the parent organization responsible for regulating, fostering, and developing football in Indonesia. Founded on April 19, 1930 in Yogyakarta by Soeratin Sosrosoegondo and other national figures, PSSI was born as part of the spirit of resistance against Dutch colonialism through sports. In its journey, PSSI became the nation's representative in the international football arena, as well as the main driving institution in the management of national football competitions, both professional and amateur. PSSI has the main function of drafting regulations, managing domestic leagues, coaching young players, and representing Indonesia in FIFA and AFC membership [11]

3. Discussions

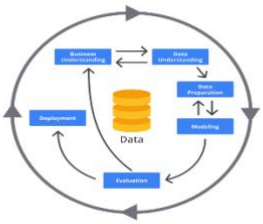

Fig. 1: Research Workflow

Based on the image above, it can be explained that there are several stages contained in CRISP-DM, which are as follows:

1. Business Understanding

   In this first stage , the main focus is to understand the purpose of a study and how the analysis of football supporters' sentiment on PSSI's performance can provide useful insights. Business Understanding aims to set priorities and define the problem to be solved, namely how football supporters' perception of the performance of PSSI can be analyzed into a new knowledge that is useful for researchers.
   In this study, the goal is to analyze the sentiment of football supporters towards the performance of the parent organization of Indonesian football, PSSI. The results of this analysis are expected to be new insights and knowledge to find out the weaknesses and strengths of PSSI's performance through the views of the country's football supporters.

2. Data Understanding

   Once the purpose of the research is understood, the next stage is to understand the data that will be used in the research. At this stage, the data is sourced from social media that is often used by supporters, namely Twitter(X). The data collected is in the form of comments or opinions from supporters' views which will later be analyzed to classify their sentiments towards PSSI's performance.
   The data that has been collected is then analyzed to understand the patterns in a comment text. This analysis includes identifying the distribution of sentiment (neutral, positive, and negative comments), as well as examining the relevance of comments to the performance that has been built by PSSI. At this stage, an initial exploration of the data was also carried out, such as the amount of data received and the type of sentiment that existed.

3. Data Preparation

   This stage involves the process of cleaning and transforming the data to prepare it to be ready for use in modeling. Data obtained through the results of crawling from social media X will be processed through several Text Preprocessing steps as follows:
   a) Cleaning: Removes irrelevant elements such as URLs, emoticons, numbers, and special characters. Here is an example of the cleaning stages in the following table:

Table 1: Cleaning

Data before cleaning	After cleaning
Bisa naturalisasi pengurus federasi gak sih sprti pemain?	Bisa naturalisasi pengurus federasi gak sih sprti pemain

b) Case Folding: Converts the entire text to lowercase to equalize the format and make analysis easier. Here is an example of the stages of case folding in the following table:

Table 2: Case Folding

Data	Case Folding *Results*
Bisa naturalisasi pengurus federasi gak sih sprti pemain?	bisa naturalisasi pengurus federasi gak sih sprti pemain

c) Normalization: Standardizing the writing of non-standard words into standard words, such as correcting typos and changing abbreviations to standard forms. The following is an example of the normalization stages in the following table:

Table 3: Normalization

Data Before Normalization	Normalization Results
Bisa	bisa
Naturalisasi	naturalisasi
Pengurus	pengurus
Federasi	federasi
Gak	Tidak
Sih	sih
Sprti	seperti
Pemain	pemain

d) Tokenizing: Breaking text into smaller words or tokens. Here is an example of the tokenizing stages in the following table

Table 4: Tokenizing

Data Before Tokenizing	Tokenizing Results
bisa naturalisasi pengurus federasi tidak sih seperti pemain	bisa naturalisasi pengurus federasi tidak seperti pemain

e) Stopword Removal: Removes common words that don't provide important information, such as "which", "and", "etc". Here is an example of the stopword removal stages in the following table:

Table 5: Stopword Removal

Data Before Stopword Removal	Hasil Stopword Removal
Bisa	bisa
Naturalisasi	naturalisasi
Pengurus	pengurus
Federasi	federasi
Yang	
Seperti	seperti
Pemain	pemain

f) Stemming: Changing a word to its basic form (root). Here is an example of the stemming stages in the following table:

Table 6: Stemming

Data Before Voting	Voting Results
Bisa	bisa
naturalisasi	natural
pengurus	urus
Federasi	federasi
Seperti	seperti
Para	para
Pemain	main

3.1. TF-IDF

Table 7: Sample Data

ID Document	Before Text Preprocessing	After Text Preprocessing	Sentiment
D1	sampe harus berhentiin liga padahal bukan jadwal resmi fifa federasi rongsok	henti liga padahal bukan jadwal resmi fifa federasi dongok	Negatif
D2	dulu ada oranh mujamuja jadi ketua aja udah goblok dan aneh	orang muja ketua goblok aneh	Negatif
D3	pssi jeda kompetisi untuk kebaikan timnas maupun klub liga	pssi jeda kompetisi baik timnas klub liga	Positif
D4	timnas akan berprestasi itu dimulai dari liganya itu sendiri tolol	timnas prestasi mulai liga sendiri tolol	Negatif
D5	seharus setiap individu pemain timnas dilatih ini perhari	individu main timnas latih hari	Netral

First, the value of TF (d,t) word per word will be calculated. In this calculation, the first 10 words are used sequentially as a representative sample of the totality of the data. The calculation of the value of TF (d,t) is as follows:

1. Cover “henti”:

$$D1 = TF_{(d,t)} = \frac{1}{8} = 0{,}125$$
$$D2 = TF_{(d,t)} = \frac{0}{5} = 0$$
$$D3 = TF_{(d,t)} = \frac{0}{7} = 0$$
$$D4 = TF_{(d,t)} = \frac{0}{6} = 0$$
$$D5 = TF_{(d,t)} = \frac{0}{5} = 0$$

Table 7: TF Results

Term	Frekuensi					TF (t,d)				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
henti	1	0	0	0	0	0,125	0	0	0	0
liga	1	0	1	1	0	0,125	0	0,143	0,167	0
bukan	1	0	0	0	0	0,125	0	0	0	0
jadwal	1	0	0	0	0	0,125	0	0	0	0
resmi	1	0	0	0	0	0,125	0	0	0	0
fifa	1	0	0	0	0	0,125	0	0	0	0
federasi	1	0	0	0	0	0,125	0	0	0	0
dongok	1	0	0	0	0	0,125	0	0	0	0
orang	0	1	0	0	0	0	0,2	0	0	0
puja	0	1	0	0	0	0	0,2	0	0	0

The IDF(t) value for each word will be calculated on a per-word basis. In this calculation, the first 10 words are sequentially selected as a representative sample of the overall data. The calculation of the IDF(t) value is as follows:

2. Cover “henti”:

$$\text{IDF}_{(t)} = \text{Log}\left(\frac{5}{1}\right) = 5$$

Table 8: IDF Results

Term	Frekuensi					TF (t,d)					IDF
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	
henti	1	0	0	0	0	0,125	0	0	0	0	5
liga	1	0	1	1	0	0,125	0	0,143	0,167	0	1,67
bukan	1	0	0	0	0	0,125	0	0	0	0	5
jadwal	1	0	0	0	0	0,125	0	0	0	0	5
resmi	1	0	0	0	0	0,125	0	0	0	0	5
fifa	1	0	0	0	0	0,125	0	0	0	0	5
federasi	1	0	0	0	0	0,125	0	0	0	0	5
dongok	1	0	0	0	0	0,125	0	0	0	0	5
orang	0	1	0	0	0	0	0,2	0	0	0	5
puja	0	1	0	0	0	0	0,2	0	0	0	5

At this stage, the values of TF and IDF will be multiplied to get the value of TF-IDF for each word. In this calculation, the first 10 words are sequentially selected as a representative sample of the overall data. The calculation of the TF-IDF value is as follows:

3. Cover “henti” at D1:

$$\text{TF} - \text{IDF}_{(\text{d},t)} = 0{,}125 \text{ x } 5 = 0{,}625$$

Table 9: TF-IDF Results

Term	IDF	TF-IDF				
	D5		D1	D2	D3	D4	D5
henti	0	5	0,63	0	0	0	0
liga	0	1,67	0,208	0	0,238	0,278	0
bukan	0	5	0,625	0	0	0	0
jadwal	0	5	0,625	0	0	0	0
resmi	0	5	0,625	0	0	0	0
fifa	0	5	0,625	0	0	0	0
federasi	0	5	0,625	0	0	0	0
dongok	0	5	0,625	0	0	0	0
orang	0	5	0	1	0	0	0
puja	0	5	0	1	0	0	0

3.2. Klasifikasi KNN

The classification process will use the results of the TF-IDF calculation on the test data. For example, the classification of the test data can be seen in the table below.

Table 10: Test Data

ID Document	Before Text Preprocessing	After Text Preprocessing	Sentiment
D6	aneh demi timnas berprestasi liga berhentikan	aneh timnas prestasi liga henti	negatif
D7	lanjutin liga woy pssi	lanjut liga pssi	netral

D6 Calculation Process:

$$D1 = \sqrt{\begin{matrix}(0{,}625-1)^2+(0{,}2083-0{,}333)^2+(0{,}625-0)^2+(0{,}625-0)^2+\\(0{,}625-0)^2+(0{,}625-0)^2+(0{,}625-0)^2+(0{,}625-0)^2+(0-0)^2+\\(0-0)^2+(0-0)^2+(0-0)^2+(0-1)^2+(0-0)^2+(0-0)^2+(0-0)^2+\\(0-0)^2+(0-0{,}333)^2+(0-0)^2+(0-1)^2+(0-0)^2+(0-0)^2+\\(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2\end{matrix}}$$

$$= \sqrt{\begin{matrix}0{,}1406+0{,}0156+0{,}3906+0{,}3906+0{,}3906+0{,}3906+0{,}3906+0{,}3906+0+\\0+0+0+1+0+0+0+0+0{,}1111+0+1+0+0+0+0+0+0+0\end{matrix}}$$

$$= \sqrt{4{,}611}$$

$$= 2{,}147$$

$$D2 = \sqrt{(0-1)^2+(0-0{,}333)^2+(0-0)^2+(0-0)^2+(0-0)^2+\cdots+(0-0)^2}$$

$$= \sqrt{1+0{,}1111+0+0+0+\cdots+0}$$

$$= \sqrt{6{,}222}$$

$$= 2{,}494$$

$$D3 = \sqrt{(0-1)^2+(0{,}2381-0{,}333)^2+(0-0)^2+(0-0)^2+(0-0)^2+\cdots+(0-0)^2}$$

$$= \sqrt{1+0{,}0091+0+0+0+\cdots+0}$$

$$= \sqrt{5{,}569}$$

$$= 2{,}36$$

$$D4 = \sqrt{(0-1)^2+(0{,}2778-0{,}333)^2+(0-0)^2+(0-0)^2+(0-0)^2+\cdots+(0-0)^2}$$

$$= \sqrt{1+0{,}0031+0+0+0+\cdots+0}$$

$$= \sqrt{4{,}117}$$

$$= 2{,}029$$

$$D5 = \sqrt{(0-1)^2 + (0-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \cdots + (0-0)^2}$$

$$= \sqrt{1 + 0{,}1111 + 0 + 0 + 0 + \cdots + 0}$$

$$= \sqrt{7{,}111}$$

$$= 2{,}667$$

Table 11: Nearest Jaral Results D6

ID Document	Before Text Preprocessing	After Text Preprocessing	Sentiment	Distance
D4	timnas akan berprestasi itu dimulai dari liganya itu sendiri tolol	timnas prestasi mulai liga sendiri tolol	Negatif	2,029
D1	sampe harus berhentiin liga padahal bukan jadwal resmi fifa federasi rongsok	henti liga bukan jadwal resmi fifa federasi dongok	Negatif	2,147
D3	pssi jeda kompetisi untuk kebaikan timnas maupun klub liga	pssi jeda kompetisi baik timnas klub liga	Positif	2,36
D2	dulu ada oranh mujamuja jadi ketua aja udah goblok dan aneh	orang puja ketua goblok aneh	Negatif	2,494
D5	seharus setiap individu pemain timnas dilatih ini perhari	individu main timnas latih hari	netral	2,667

D7 Calculation Process:

$$D1 = \sqrt{\begin{matrix}(0{,}625-0)^2 + (0{,}2083-0{,}556)^2 + (0{,}625-0)^2 + (0{,}625-0)^2 + \\ (0{,}625-0)^2 + (0{,}625-0)^2 + (0{,}625-0)^2 + (0{,}625-0)^2 + (0-0)^2 + \\ (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1{,}667)^2 + (0-0)^2 + (0-0)^2 + \\ (0-0)^2 + (0-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\ (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2\end{matrix}}$$

$$= \sqrt{\begin{matrix}0{,}3906 + 0{,}1206 + 0{,}3906 + 0{,}3906 + 0{,}3906 + 0{,}3906 + 0{,}3906 + 0{,}3906 + 0 + \\ 0 + 0 + 0 + 0 + 2{,}778 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0\end{matrix}}$$

$$= \sqrt{5{,}6327}$$

$$= 2{,}3733$$

$$D2 = \sqrt{(0-1)^2 + (0-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \cdots + (0-0)^2}$$

$$= \sqrt{1 + 0{,}1111 + 0 + 0 + 0 + \cdots + 0}$$

$$= \sqrt{8{,}0864}$$

$$= 2{,}8437$$

$$D3 = \sqrt{(0-1)^2 + (0{,}2381-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \cdots + (0-0)^2}$$

$$= \sqrt{1 + 0{,}0091 + 0 + 0 + 0 + \cdots + 0}$$

$$= \sqrt{3{,}1053}$$

$$= 1{,}762$$

$$D4 = \sqrt{(0-1)^2 + (0{,}2778-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \cdots + (0-0)^2}$$

$$= \sqrt{1 + 0{,}0031 + 0 + 0 + 0 + \cdots + 0}$$

$$= \sqrt{5{,}7099}$$

$$= 2{,}3895$$

$$D5 \quad = \sqrt{(0-1)^2 + (0-0{,}333)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \cdots + (0-0)^2}$$

$$= \sqrt{1 + 0{,}1111 + 0 + 0 + 0 + \cdots + 0}$$

$$= \sqrt{7{,}1975}$$

$$= 2{,}6828$$

Table 12: Nearest Jaral Results D7

ID Document	Before Text Preprocessing	After Text Preprocessing	Sentiment	Distance
D3	pssi jeda kompetisi untuk kebaikan timnas maupun klub liga	pssi jeda kompetisi baik timnas klub liga	Positif	1,762
D1	sampe harus berhentiin liga padahal bukan jadwal resmi fifa federasi rongsok	henti liga bukan jadwal resmi fifa federasi dongok	Negatif	2,373
D4	timnas akan berprestasi itu dimulai dari liganya itu sendiri tolol	timnas prestasi mulai liga sendiri tolol	Negatif	5,71
D5	seharus setiap individu pemain timnas dilatih ini perhari	individu main timnas latih hari	netral	2,39
D2	dulu ada oranh mujamuja jadi ketua aja udah goblok dan aneh	orang puja ketua goblok aneh	Negatif	2,844

Table 13: Prediction Results

ID Document	Before Text Preprocessing	After Text Preprocessing	Sentiment	Sentiment (Predicted Results)
D6	aneh demi timnas berprestasi liga berhentikan	aneh timnas prestasi liga henti	negatif	negatif
D7	lanjutin liga woy pssi	lanjut liga pssi	netral	negatif

3.3. Results of the KNN Model Evaluation

After the training process is complete, the next step is to test the performance of the K-Nearest Neighbor model on the test data to determine the model's ability to classify comments into negative, positive, or neutral sentiment categories. The evaluation was carried out using two approaches, namely prediction probability analysis to see the model's tendency in determining the class, and performance measurement using accuracy, precision, recall, and f1-score metrics.

Classification Report for KNN:

	precision	recall	f1-score	support
negatif	0.000	0.000	0.000	7.000
netral	0.932	1.000	0.965	344.000
positif	0.964	0.587	0.730	46.000
accuracy	0.935	0.935	0.935	0.935
macro avg	0.632	0.529	0.565	397.000
weighted avg	0.920	0.935	0.921	397.000

Fig. 2: KNN Model Report Classification Results

4. Conclusion & Suggestions

4.1 Conclusion

Based on the sentiment analysis carried out on the opinions of football supporters on social media regarding PSSI's performance using the K-Nearest Neighbors (KNN) method, several conclusions can be drawn as follows:

1. Sentiment Analysis Successfully Carried Out: This study successfully implemented the stages of sentiment analysis, starting from data collection, pre-processing of text (such as *cleaning*, *case folding*, *tokenizing*, and *stemming*), weighting of features using TF-

IDF, to classification using the KNN method. This process shows that unstructured public opinion on social media can be transformed into data that can be systematically analyzed.

2. Effectiveness of the KNN Method: The KNN method showed a fairly good performance in classifying sentiment, with a model accuracy of 93.5%. This shows that the KNN method is effectively used for sentiment analysis, especially on well-prepared text data.

3. Data Imbalance: Significant imbalances were found in the data used, where the majority of comments were neutral sentiment (1724 comments), followed by positive sentiment (220 comments), and negative sentiment (40 comments). This imbalance affects the performance of the model, which is shown to have a tendency to have difficulty accurately recognizing negative and positive sentiment classes.

4. Model Prediction Results: The model has an overall accuracy of 93.5%, but the classification report shows weaknesses in predicting both negative and positive classes. This can be seen from the low *precision* and *recall* values for both classes, while the neutral class has very high values. The model tends to predict most data as "neutral" due to the dominance of data in that class.

4.1. Suggestion

Based on the conclusions and limitations found in this study, here are some suggestions for future research:

1. Addressing Data Imbalances: To address the problem of data imbalance, it is recommended to use unbalanced data handling techniques, such as oversampling methods (e.g. SMOTE) or undersampling. This technique can help the model learn better than the minority class (negative and positive) thereby increasing the accuracy and recall value of the model.

2. Trials with Other Algorithms: Although KNN shows high accuracy, it is recommended to compare its performance with other classification algorithms, such as Naive Bayes, Support Vector Machine (SVM), or deep learning-based models (e.g. Long Short-Term Memory/LSTM or BERT). This comparison can provide further insight into which algorithm is the most optimal for analyzing sentiment regarding sports issues.

3. Increased Data Coverage: Subsequent research may collect data from other social media platforms (such as Facebook or Instagram) or with a wider time span. This aims to obtain a more representative and varied dataset, which can reflect public opinion more comprehensively.

4. Additional Extraction Features: In addition to TF-IDF, it is recommended to consider other feature extraction methods, such as Word2Vec or GloVe, which can capture the semantic meaning and context of words. The use of this feature can help models understand the more complex nuances of language, such as sarcasm and irony, that often appear on social media.

References

[1] M. F. Asshiddiqi and K. M. Lhaksmana, "Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI," in *e-Proceeding of Engineering*, M. F. Asshiddqi and Kemas Muslim Lhaksmana, Eds., Universitas Telkom, 2020, pp. 9936–9948.

[2] M. Ulinnuha, L. Nurrachmad, and U. N. Semarang, "Peran Dan Model Strategi Komunikasi Asosiasi Provinsi (Asprov) Pssi Jawa Tengah Dalam Menyelenggarakan Kompetisi Sepakbola Liga 3 Jawa Tengah 2023," vol. 4, 2024.

[3] A. Aliyudin, "Analisis Sentimen Terhadap Piala Dunia U-20 Yang Batal Diselenggarakan Di Indonesia Pada Media Youtube Dengan Menggunakan Metode Naïve Bayes," *Informasi Interaktif: Jurnal Informatika dan Teknologi Informasi*, vol. 8, no. 2, pp. 71–78, 2023.

[4] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," *Journal of Intelligent System and Computation*, vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.

[5] J. Supriyanto, D. Alita, and A. R. Isnain, "Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 74–80, 2023, doi: 10.33365/jatika.v4i1.2468.

[6] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *Jurnal SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.

[7] S. Navisa, Luqman Hakim, and Aulia Nabilah, "Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM," *Jurnal Sistem Cerdas*, vol. 4, no. 2, pp. 114–125, 2021, doi: 10.37396/jsc.v4i2.162.

[8] E. Laksono, A. Basuki, and F. Bachtiar, "Optimization of K Value in KNN Algorithm for Spam and Ham Email Classification," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 2, pp. 377–383, 2020, doi: 10.29207/resti.v4i2.1845.

[9] "79-Article Text-148-1-10-20170314 (3)".

[10] "JOURNAL OF INTELLIGENT SYSTEMS AND COMPUTATION 43." [Online]. Available: https://t.co/9WloaWpfD5

[11] I. Ataqwa, "Indonesian Journal for," *Journal.Unnes*, vol. 1, no. 1, pp. 188–196, 2020.