

---

## Sentiment Analysis of Twitter Cases of Riots at Kanjuruhan Stadium Using the Naive Bayes Method

Bryan Jerremia Katiandhago<sup>1)</sup>, Akhmad Mustolih<sup>2)</sup>, Wachyu Dwi Susanto<sup>3)</sup>, Pungkas Subarkah<sup>4)\*</sup>,  
Chendri Irawan Satrio Nugroho<sup>5)</sup>

<sup>1)2)3)4)</sup>Program Studi Informatika, Universitas Amikom Purwokerto, Indonesia

<sup>5)</sup>Program Studi Informatika, Institut Teknologi Tangerang Selatan, Indonesia

<sup>1)</sup>[jerremia.bryan777@gmail.com](mailto:jerremia.bryan777@gmail.com), <sup>2)</sup>[tolih282@gmail.com](mailto:tolih282@gmail.com), <sup>3)</sup>[wahyudsx@gmail.com](mailto:wahyudsx@gmail.com),

<sup>4)</sup>[subarkah@amikompurwokerto.ac.id](mailto:subarkah@amikompurwokerto.ac.id), <sup>5)</sup>[chendri@itts.ac.id](mailto:chendri@itts.ac.id)

---

### ABSTRACT

Sentiment analysis is a process carried out to analyze opinions, sentiments, judgments, and emotions from the riot case at the Kanjuruhan stadium. The purpose of this research is to find out public opinion about the tragedy that is currently happening at the Kanjuruhan Stadium. The data was obtained from social media Twitter using the Twitter API, then after that, an analysis was carried out. Data from the results of the analysis will be classified using the Naive Bayes method. The classification process is divided into 7 (seven) stages, namely Crawling, Cleansing, pre-processing, labeling, classification, data training, and data testing. In the labeling process, data is classified into 2 (two) classes, namely the positive class and the negative class. The data obtained before the preprocessing process was 1963 tweets, after the preprocessing the data obtained was 1001 tweets. The data will be trained and tested using the naive Bayes classification method. Classification results obtained precision values of 82% for negative data and 65% for positive data, recall values obtained 74% for negative data and 75% for positive data, F1-score values obtained 78% for negative data and 70% for positive data, while accuracy value obtained 74%.

**Keywords:** *sentiment, sentiment analysis, kanjuruhan, naive bayes, twitter.*

---

### 1. INTRODUCTION

The development of information technology in Indonesia is growing rapidly. The development of information technology is increasingly making it easier for people to get the information they need. This advancement in information technology makes users more curious about their new world. This new world is known as social media. Social media is increasingly used and increasingly dominates the world. Indonesian people, both in cities and in villages, use social media. In 2019 alone, based on the news portal [tribunnews.com](http://tribunnews.com), it was noted that India was ranked 4th after India, the United States, and Brazil (R. Gunawan, Septiadi, Apri Wenando, Mukhtar, & Syahril, 2022).

Social media is growing rapidly since the internet is easily accessible. One of the popular social media is Twitter. Twitter is a social media site that was developed in 2006. This site was first discovered by Jack Dorsey and Evan Williams. Twitter is a social network that allows its users to communicate with each other through a feature called tweet. With the tweet feature, users can write or text up to 280 characters (Fairuz, Ramadhani, & Tanjung, 2021).

Sentiment analysis is a process that is carried out to analyze opinions, sentiments, judgments, and emotions from one's statements about a domain. There are various possible domains ranging from services, products, or an event. Sentiment analysis can be interpreted as a process carried out to identify the sentiments contained in a text. The process of sentiment analysis is generally carried out to classify text contained in sentences including positive or negative sentiments (Dharmawan, Arwani, & Ratnawati, 2020).

The classification method used is the Naive Bayes method. Naive Bayes is a classification that is supervised learning because it has a supervisor (a human manually classifies the data used in training) as a teacher in the learning process. In addition, the performance of Naive Bayes has a short classification time, thereby speeding up the process of the sentiment analysis system (B. Gunawan, Pratiwi, & Pratama, 2018).

Based on the events that are currently busy in the community, many social media users have provided their opinions, opinions, and thoughts on the riot case at the Kanjuruhan Stadium on the Twitter social media platform. This is interesting to study to find out public opinion about the tragedy at the Kanjuruhan Stadium that is happening right

\* Corresponding author



now. To support research on this matter, an algorithm is needed to classify public comments on social media Twitter, both positive comments and negative comments. Several algorithms can be used to classify comments on Twitter social media, including Naive Bayes, K-Nearest Neighbor, Decision Tree, and also the Support Vector Machine. In this study, Naive Bayes will be selected as the algorithm that will be used to classify comments on Twitter social media (Puspita & Widodo, 2021).

## 2. LITERATURE REVIEW

Several previous studies related to Sentiment Analysis used the Naive Bayes method. First, research conducted by Lian Ardiani, et al, with the research title "Implementation of Sentiment Analysis of Community Responses to Development in Pontianak City", in this study, data was classified into 3 (three) classes namely, positive, negative, and neutral classes. The process of Sentiment Analysis is done by doing text processing. after that, the results of text processing are classified using the Naive Bayes classification method into positive, negative, and neutral classes. Test results using the Naive Bayes classification method to process sentiment information contained in tweet data automatically with a total of 450 training data and 50 test data, obtained an accuracy of 72% (Ardiani, Sujaini, & Tursina, 2020).

Second, research conducted by Lingga Aji Andika, et al, with the research title "Analysis of Public Sentiment on the Quick Count Results of the 2019 Indonesian Presidential Election on Twitter Social Media Using the Naive Bayes Classifier Method". In this study, the data were classified into 2 (two) classes, namely positive and negative classes. this study resulted in an accuracy of 82.90%. The classification obtained is 34.5% with data of 471 positive tweets and 65.5% with data of 895 negative tweets from the quick count results (Andika, Azizah, & Respatiwan, 2019).

Third, research was conducted by Mujaddid Izzul Fikri, et al, with the research title "Comparison of the Naive Bayes Method and Support Vector Machine in Twitter Sentiment Analysis". In this study, analysis was carried out by classifying tweets containing public sentiment about UMM (University of Muhammadiyah Malang). There are 2 (two) classification methods used in this study, namely, Naive Bayes and Support Vector Machine (SVM) using TF-IDF weighting. The results of the comparison of the two methods show that Naive Bayes gets better accuracy results than SVM with an accuracy of 73.65% (Fikri, Sabrila, & Azhar, 2020).

Fourth, research was conducted by Muhammad Syarifuddin, entitled "Analysis Of Public Opinion Sentiments Regarding Covid-19 On Twitter Using The Naive Bayes And Knn Methods". This study focuses on the results of the comparison of the Naive Bayes and KNN methods. Data were taken using the Twitter API as many as 1098 opinions with the keyword "COVID-19". The stages carried out in this study are classification which is divided into 2 (two) classes, namely, Positive and Negative, data cleansing, and preprocessing, to get the final result. The results of the classification using the Naive Bayes method obtained an accuracy of 63.21% while the classification using the KNN method obtained an accuracy of 58.10% and it was also found that the tendency of public opinion on Twitter to be positive, this can be seen from the number of positive opinions of 610 while negative 488, supported by precision test results in the Naive Bayes method with positive values higher than negative, namely 66.40%: 58.94% (Syarifuddin, 2020).

Fifth, research was conducted by Fransisca Vina Sari, entitled "Analysis Of Jd. Id Online Store Customer Sentiment Using The Naive Bayes Classifier Method Based On Emotional Icon Conversion". This study uses the Naive Bayes Classifier (NBC) method with tf-idf weighting along with the addition of an emotion icon (emoticon) conversion feature to find out the existing sentiment class from tweets about JD.id stores. The results show that the Naive Bayes method without adding features can classify emotions with an accuracy of 96.44%, whereas if the tf-idf weight function is added with the transformation of emotion symbols, the accuracy value can be increased to 98% (Sari & Wibowo, 2019).

## 3. METHOD

In this study, the method used is the classification method with the Naive Bayes Algorithm there are many classification methods such as Decision Trees, Neural Networks, Support Vector Machines, and Machine Learning (Syarifuddin, 2020). However, researchers only focused on 1 (one) method only. The research stages are in Figure 1. The following:

\* Corresponding author



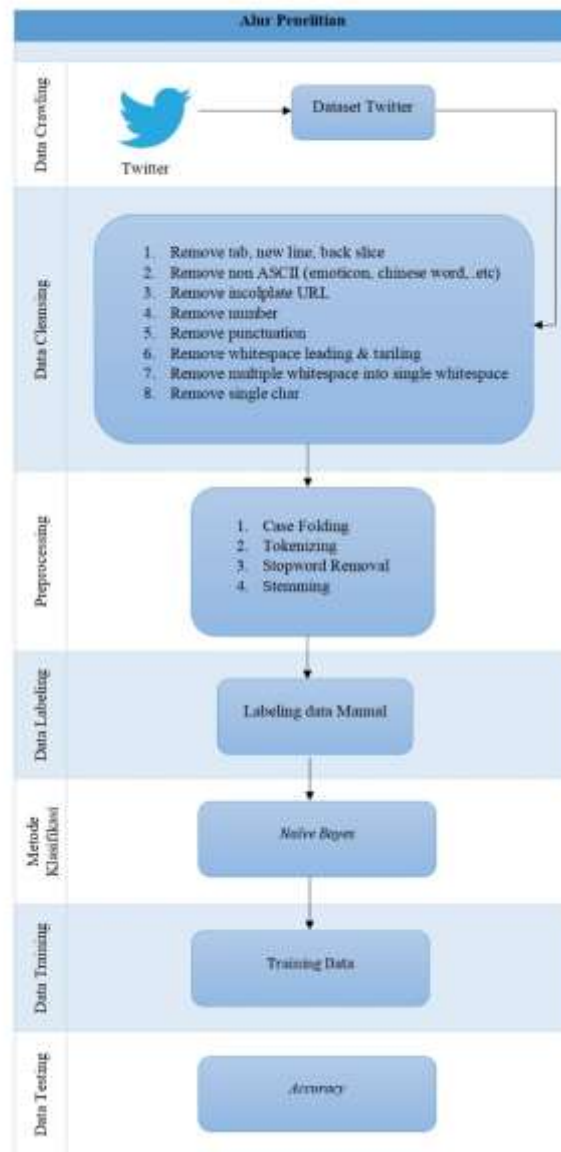


Figure 1. Research Flow

\* Corresponding author



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

### 3.1 Data Retrieval

From the flowchart in Figure 1 above, it can be seen that the research started by searching data on Twitter using the Twitter API. The data used in this study were tweets containing the word "Kanjuruhan". The dataset obtained is 1963 tweets.

### 3.2 Cleansing

The process of identifying and correcting or removing incorrect, incomplete, or irrelevant data from a dataset. This is an important step in the preparation of data for further analysis or use because inaccuracies or inconsistencies in the data can lead to erroneous and imprecise conclusions.

### 3.3 Text Preprocessing

The process for preparing data to be used in text mining. So that the dataset is clean and ready to use (Fitri, 2020). The text preprocessing stage in this study uses several stages, namely:

3.3.1 *Case Folding*: The process of converting the entire text in a tweet into standard form (Kurniawan & Apriliani, 2020). Not all words in the text consistently use uppercase, this is where case folding is done to change every character in the word to lowercase (Ardiani et al., 2020).

3.3.2 *Tokenizing*: The process of cutting a document into parts, which are called tokens. At the same time, tokenizing also serves to remove certain characters that are considered punctuation (Alita & Isnain, 2020).

3.3.3 *Stopword Removal*: The process of eliminating words that are considered to have no effect in the classification process (Natasuwarna, 2019).

3.3.4 *Stemming*: Words that have been changed to lowercase need to be checked. Stemming is used to standardize words to reduce the list of words in the training data (B. Gunawan et al., 2018).

### 3.4 Labelling

The process of assigning labels or categories to the data to be analyzed. Data labeling is usually done to facilitate the analysis or modeling of the data. In this study, the data that was given labeling was tweets data that had been carried out by Text Preprocessing as many as 1001 data tweets. Labeling is divided into 2 (two) types, namely positive and negative.

### 3.5 Klasifikasi

Classification belongs to the learning area of machine learning where machines learn data that have been labeled or data that has been classified. The classification used in this study is Naive Bayes, which is a very efficient linear classification. The probability model and Naive Bayes classification are based on Bayes' theorem, and the adjective naive comes from the assumption that the features in the dataset are mutually independent (Muzaki & Witanti, 2021).

### 3.6 Data Training

To create machine learning models or other predictive algorithms that can predict labels from new data with a high degree of accuracy. The data training process is usually carried out using a dataset that has been labeled, where the model will try to find patterns from the data provided, and use these patterns to predict labels from new data.

### 3.7 Data Testing

Aims to find out how well the model that has been made can predict or classify new data. Data testing is used to evaluate the model that has been made by measuring the accuracy, thoroughness, and how well the model can handle previously unknown data. Data testing can also be used to compare different data or to determine the best hyperparameters for the model.

\* Corresponding author



4. RESULT

The data was taken from the official website [www.twitter.com](http://www.twitter.com) with the Twitter API and the keyword used was "Kanjuruhan" in October 2022, taken starting October 1 and obtained data from as many as 1963 tweets that had not been processed, then after processing again, only 1001 tweets remained. regarding the kanjuruhan tragedy, then there were 596 negative opinions and 405 positive opinions. Then the data will be processed and tested and training the data set to get the accuracy value, as well as precision and recall values. Figure 2 below is a representation of the number of positive and negative opinions on the training data used.

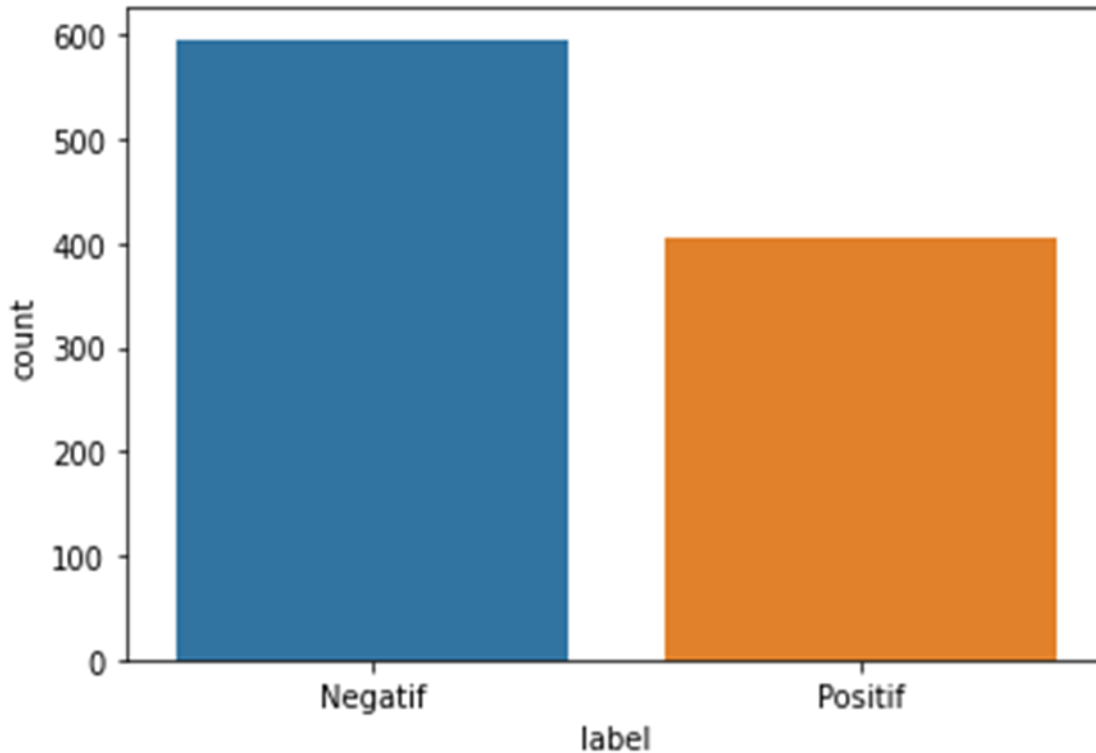


Figure 2. Comparison of the Number of Positive and Negative Opinions

4.1 Cleansing

The process of data cleansing or data cleansing is the process of cleaning or processing data with the aim of improving or improving data quality. The purpose of this process is usually to ensure that the data to be used meets established standards, meets user needs, and is easy to understand. This process can include a variety of activities, such as deleting unnecessary data, removing duplicate data, changing inappropriate data formats, and filling in missing or invalid data. Below is the table after cleansing:

Content
b'@baniimron Senangnya beda dg yg dulu...ya hanya menyampaikan informasi saja. Tapi sekali lagi tim ini tidak ada salve2vx80'xa6 <a href="https://t.co/sU5rzc2NG">https://t.co/sU5rzc2NG</a> '
b'RT @more80s: Hari ini puncaknya pesta sepak bola dunia, dengan sengaja saya mengucapkan USUT TUNTAS TRAGEDI KANJURUHAN!
b'RT @okkymadasari: Malang, 2 bulan lebih setelah Tragedi Kanjuruhan. Kemarahan dan kesedihan masih terus bergema di setiap sudut kota.'n'nSiap'xe2'v80'xa6'
b'RT @kiminigyu: Buka Tabir Kelam Tragedi Kanjuruhan: Banyak Stadion Tidak Layak. 'n'n#TabirKanjuruhan 'n'Stadion Aman Nyaman'n-UTAS-'n' <a href="https://t.co/ve2vx80'xa6">https://t.co/ve2vx80'xa6</a> '
b'RT @kiminigyu: Buka Tabir Kelam Tragedi Kanjuruhan: Banyak Stadion Tidak Layak. 'n'n#TabirKanjuruhan 'n'Stadion Aman Nyaman'n-UTAS-'n' <a href="https://t.co/ve2vx80'xa6">https://t.co/ve2vx80'xa6</a> '

Figure 3. Data that has been Cleansing

\* Corresponding author



4.2 Text Preprocessing

The process of text preprocessing or text processing is the process of cleaning and processing text in a dataset to prepare the text for input into a text processing model or algorithm. This process can include various activities, such as case folding (turning all text into lowercase or uppercase), stemming (removing affixes to words), stopword removal (removing meaningless words such as "and", "or", etc. .), and others. The purpose of this process is to reduce the dimensionality of text data and increase the efficiency of the model or algorithm that will use it. In addition, this process is also useful for removing noise or distractions that may exist in the text so that it is easier for the model or algorithm to understand.

4.3 Case Folding

The case folding process is also useful for eliminating differences that might occur due to differences in the use of uppercase and lowercase letters. For example, the words "apple" and "apple" might be perceived as different words by a text processing model or algorithm, even though they are semantically the same. By carrying out the case folding process, the two words will be considered the same word. Below is the table after case folding is done:

Content
b'@baniimron senangnya beda dg yg dulu...ya hanya menyampaikan informasi saja. tapi sekali lagi tim ini tidak ada sal'xe2'x80'wa6 https://t.co/su5rzjc2ng'
b'rt @more80s: hari ini puncaknya pesta sepak bola dunia, dengan sengaja saya mengucapkan usut tuntas tragedi kanjuruhan!
b'rt @oklymadasari: malang, 2 bulan lebih setelah tragedi kanjuruhan. kemarahan dan kesedihan masih terus bergema di setiap sudut kota.'n'nsiap'xe2'x80'wa6'
b'rt @kiminigy: buka tabir kelam tragedi kanjuruhan: banyak stadion tidak layak. 'n'n#tabir'kanjuruhan 'n'stadion aman nyaman'n-utas-'nhttps://t.co'xe2'x80'wa6'
b'rt @kiminigy: buka tabir kelam tragedi kanjuruhan: banyak stadion tidak layak. 'n'n#tabir'kanjuruhan 'n'stadion aman nyaman'n-utas-'nhttps://t.co'xe2'x80'wa6'

Figure 4. Data that has been done Case Folding

4.4 Tokenizing

The process of tokenizing or splitting text is the process of breaking a text into small tokens or basic words. The purpose of this process is to facilitate a text-processing model or algorithm in analyzing and understanding the text. Tokens can be words, phrases, or even sentences depending on your needs and desired level of granularity. Below is the table after tokenization:

Content	tweet_tokens
senangnya beda dg yg duluya hanya menyampaikan informasi saja tapi sekali lagi tim ini tidak ada sal'xe2'x80'wa6 https://t.co/su5rzjc2ng'	senangnya,beda,dg,yg,duluya,hanya,menyampaikan,informasi,saja,tapi,sekali,lagi,tim,ini,tidak,ada,sal'xe2'x80'wa6,https://t.co/su5rzjc2ng'
b'rt hari ini puncaknya pesta sepak bola dunia dengan sengaja saya mengucapkan usut tuntas tragedi kanjuruhan	b'rt,hari,ini,puncaknya,pesta,sepak,bola,dunia,dengan,sengaja,saya,ucapkan,usut,tuntas,tragedi,kanjuruhan
b'rt malang bulan lebih setelah tragedi kanjuruhan kemarahan dan kesedihan masih terus bergema di setiap sudut kota. siap'xe2'x80'wa6'	b'rt,malang,bulan,lebih,setelah,tragedi,kanjuruhan,kemarahan,dan,kesedihan,masih,terus,bergema,di,setiap,sudut,kota,siap'xe2'x80'wa6'
b'rt buka tabir kelam tragedi kanjuruhan banyak stadion tidak layak stadion aman nyaman utas t'co'xe2'x80'wa6'	b'rt,buka,tabir,kelam,tragedi,kanjuruhan,banyak,stadion,tidak,layak,stadion,aman,nyaman,utas,t'co'xe2'x80'wa6'
b'rt buka tabir kelam tragedi kanjuruhan banyak stadion tidak layak stadion aman nyaman utas t'co'xe2'x80'wa6'	b'rt,buka,tabir,kelam,tragedi,kanjuruhan,banyak,stadion,tidak,layak,stadion,aman,nyaman,utas,t'co'xe2'x80'wa6'

Figure 5. Data that has been Tokenized

4.5 Stopword Removal

The process of stopping word removal or removing meaningless words is the process of removing words that have no meaning or do not have useful information from a text. These words are usually referred to as stopwords or junk words. Examples of frequently used stopwords are words like "and", "or", "the", "in", etc. Below is a data comparison table before the stopwords removal process is carried out and after the stopwords removal process:

\* Corresponding author



tweet_tokens	tweet_tokens_fdlist	tweet_tokens_WSW
senangnya beda dg yg dulunya hanya menyampaikan informasi saja tapi sekali lagi tim ini tdk ada salvexxa tcosurjcnng	<FreqDist with 18 samples and 18 outcomes>	senangnya beda duluya informasi tim salvexxa tcosurjcnng
brt,malang,bulan,lebih,selelah,tragedi,kanjuruhan,kemarahan,dan,kesedihan,masih,terus,bergema,di,setiap,sudut,kota,siapxxa	<FreqDist with 18 samples and 18 outcomes>	brt,malang,tragedi,kanjuruhan,kemarahan,kesedihan,bergema,sudut,kota,siapxxa
brt,hari,ini,puncaknya,pesta,sepak,bola,dunia,dengan,sengaja,saya,mengucapkan,usut,tuntas,tragedi,kanjuruhan	<FreqDist with 16 samples and 16 outcomes>	brt,puncaknya,pesta,sepak,bola,dunia,sengaja,usut,tuntas,tragedi,kanjuruhan

Figure 6. Comparison after and before Stopword Removal

4.6 Stemming

The process of steaming or cutting affixes is the process of removing affixes or endings to a word in a text. The goal is to reduce the dimensionality of text data and increase the efficiency of the model or algorithm that will use it. Word affixes or suffixes usually do not have a specific meaning and do not provide useful information for text-processing models or algorithms. In addition, affixes or suffixes can also be a source of noise or interference which can reduce the accuracy of the model or algorithm used. Below is a data comparison table before the steaming process and after stemming:

tweet_tokens_WSW	tweet_tokens_stemmed
senangnya beda duluya informasi tim salvexxa tcosurjcnng	senang beda duluya informasi tim salvexxa tcosurjcnng
brt,puncaknya,pesta,sepak,bola,dunia,sengaja,usut,tuntas,tragedi,kanjuruhan	brt,puncak,pesta,sepak,bola,dunia,sengaja,usut,tuntas,tragedi,kanjuruhan
brt,malang,tragedi,kanjuruhan,kemarahan,kesedihan,bergema,sudut,kota,siapxxa	brt,malang,tragedi,kanjuruhan,marah,sedih,gema,sudut,kota,siapxxa

Figure 7. Comparison before and after stemming

4.7 Labelling

The labeling process can be done in various ways depending on the type of data to be labeled and the purpose of the labeling process. For example, the labeling process can be done by giving a "spam" or "not spam" label to an email dataset and giving a "positive" or "negative" label to a sentiment analysis dataset as was done in this study. Below is an example table of data that has been labeled:

\* Corresponding author



index	text	label
0	senangnya beda dg yg dulunya hanya menyampaikan informasi saja tapi sekali lagi tim ini tidak ada salivexxa tcosurjcong	Negatif
1	brt hari ini puncaknya pesta sepak bola dunia dengan sengaja saya mengucapkan usut tuntas tragedi kanjuruhan	Positif
2	brt malang bulan lebih setelah tragedi kanjuruhan kemarahan dan kesedihan masih terus bergema di setiap sudut kota siaprexxa	Negatif
3	brt buka tabir kelam tragedi kanjuruhan banyak stadion tidak layak stadion aman nyaman utas tcoveexxa	Negatif
4	brt buka tabir kelam tragedi kanjuruhan banyak stadion tidak layak stadion aman nyaman utas tcoveexxa	Positif

Figure 8. Data that has been labeled

#### 4.8 Klasifikasi

In the classification process, the data to be classified is referred to as an instance, while the groups to be formed are referred to as classes. The classification process usually involves two main stages, namely the training stage and the testing stage. The training stage is the stage where a machine learning model or algorithm is created and trained using data that the data group or category is known to be. The testing phase is the stage where the model or algorithm is tested for its ability to predict new groups or categories of data that were not previously known. The following is a table of classification results from this study using the Naïve Bayes classification method:

```

accuracy: 0.743
              precision    recall  f1-score   support

   negatif      0.82      0.74      0.78         61
   positif      0.65      0.75      0.70         40

 accuracy              0.74         101
 macro avg           0.74         0.74         0.74         101
 weighted avg        0.75         0.74         0.74         101
    
```

Figure 9. Classification Results

#### 4.9 Word Cloud

The word cloud is a data visualization that displays words in shades of color and text size. Its function is to visually display the distribution of words in a text or group of texts. By using the word cloud, we can see the words that appear most often in the text and know the trend of the most popular words. A word cloud can also be used to assist in the data exploration process and initial understanding of the text to be analyzed. Following are the Word cloud results of the Kanjuruhan riot case tweet dataset.

\* Corresponding author





Figure 10. Word Cloud results

#### 4.10 Confusion Matrix

The function of data visualization using a confusion matrix is to assist in understanding the performance of a classification model. The confusion matrix provides information about how much data the model has classified correctly and incorrectly. This allows you to see the success and failure rates of the model in grouping the data into the right classes. The confusion matrix can also provide information about what classes are difficult for the model to classify so that it can assist in determining the next direction in model development.

\* Corresponding author



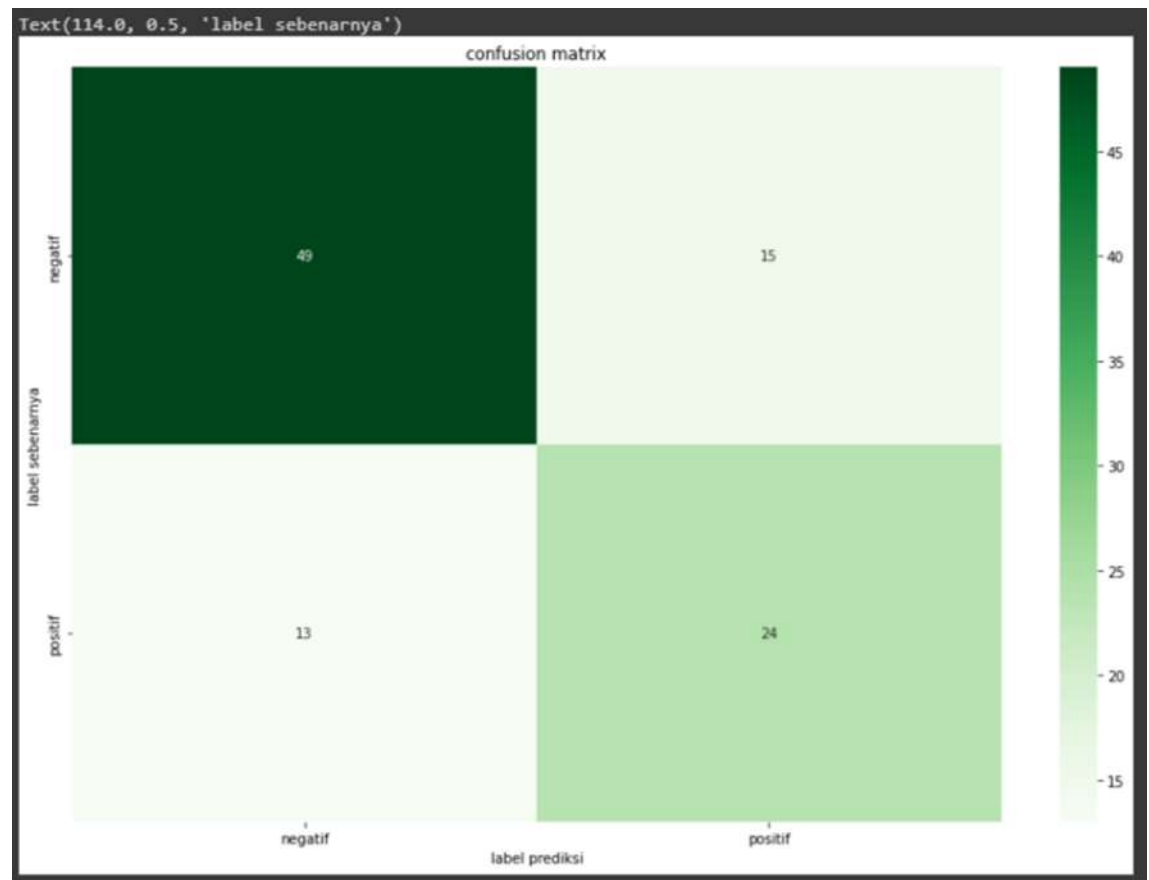


Figure 11. Confusion Matrix Results

## 5. CONCLUSION

From the research that has been done, namely Twitter sentiment analysis about the riot case at Kanjuruhan stadium using the Naïve Bayes algorithm, from the crawling process with the keyword "Kanjuruhan" 1963 data were obtained, and 1001 after the preprocessing process was carried out, with a comparison of 596 negative opinions and 405 positive opinions after manual labeling, then from the classification results obtained precision values of 82% for negative data and 65% for positive data, recall values obtained 74% for negative data and 75% for positive data, the f1-score value obtained 78% for data negative and 70% for positive data, while the accuracy value is 74%.

## 6. REFERENCES

- Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, 8(2), 50–58. <https://doi.org/10.23960/komputasi.v8i2.2615>
- Andika, L. A., Azizah, P. A. N., & Respatiwan, R. (2019). Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier. *Indonesian Journal of Applied Statistics*, 2(1), 34. <https://doi.org/10.13057/ijas.v2i1.29998>
- Ardiani, L., Sujaini, H., & Tursina, T. (2020). Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak. *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 8(2), 183. <https://doi.org/10.26418/justin.v8i2.36776>
- Dharmawan, L. R., Arwani, I., & Ratnawati, D. E. (2020). Analisis Sentimen pada Sosial Media Twitter Terhadap

\* Corresponding author



- Layanan Sistem Informasi Akademik Mahasiswa Universitas Brawijaya dengan Metode K- Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(3), 959–965. Retrieved from <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/7099>
- Fairuz, A. L., Ramadhani, R. D., & Tanjung, N. A. F. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 1(1), 42–51. <https://doi.org/10.20895/dinda.v1i1.180>
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *Smatika Jurnal*, 10(02), 71–76. <https://doi.org/10.32664/smatika.v10i02.455>
- Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71. <https://doi.org/10.26623/transformatika.v18i1.2317>
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(2), 113. <https://doi.org/10.26418/jp.v4i2.27526>
- Gunawan, R., Septiadi, R., Apri Wenando, F., Mukhtar, H., & Syahril. (2022). K-Nearest Neighbor (KNN) untuk Menganalisis Sentimen terhadap Kebijakan Merdeka Belajar Kampus Merdeka pada Komentar Twitter. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(2), 152–158. <https://doi.org/10.37859/coscitech.v3i2.3841>
- Kurniawan, R., & Apriliani, A. (2020). 13686-34825-1-SM.pdf. *Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scraper*, Vol. 5.
- Muzaki, A., & Witanti, A. (2021). Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier. *Jurnal Teknik Informatika (Jutif)*, 2(2), 101–107. <https://doi.org/10.20884/1.jutif.2021.2.2.51>
- Natasuwarna, A. P. (2019). Analisis Sentimen Keputusan Pemindahan Ibukota Negara Menggunakan Klasifikasi Naive Bayes. *SEMINAR NASIONAL SISTEM INFORMASI Dan TEKNIK INFORMATIKA SENSITIF*, 47–53.
- Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*, 5(4), 646. <https://doi.org/10.32493/informatika.v5i4.7622>
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Syarifuddin, M. (2020). Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn. *INTI Nusa Mandiri*, 15(1), 23–28. <https://doi.org/10.33480/inti.v15i1.1347>

\* Corresponding author

