JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

SENTIMENT ANALYSIS OF PUBLIC OPINION ON APPLICATION X (TWITTER) IN INDONESIA AGAINST CHATGPT USING NAÏVE BAYES ALGORITHM

Yayak Kartika Sari*[1)], Fahrur Rozi[2)], Sulton Muhyiddin[3)], Farid Sukmana[4)]

1. Informatics, Faculty of Sains and Technology, University of Bhinneka PGRI, Indonesia
2. Information Technology Education, Faculty of Sains and Technology, University of Bhinneka PGRI, Indonesia
3. Informatics, Faculty of Sains and Technology, University of Bhinneka PGRI, Indonesia
4. Informatics Engineering, Faculty of Engineering, University of Muhammadiyah Gresik, Indonesia

Article Info
Keywords: ChatGPT; Google Collabs; Naïve Bayes; Recall.

Article history:
Received 13 September 2024
Revised 5 Oktober 2024
Accepted 15 Oktober 2024
Available online 1 December 2024

DOI :
https://doi.org/10.29100/jipi.v9i4.7052

* Corresponding Author
Yayak kartika Sari
E-mail address:
Yayakkartikasari93132042@gmail.com

ABSTRACT
In the era of technological development and information is increasingly widespread. Data and information are easier to obtain using current technology, especially using social media such as Instagram, Facebook, (x) Twitter and others. In social media, information can be in the form of public opinions containing praise, hate speech, and hoaxes which can result in arguments against the information presented, especially on the x (twitter) application. Therefore, research was conducted on sentiment analysis of positive and negative opinions of Indonesian people on application x (twitter) about ChatGPT using the naive bayes method. Basically, Naive Bayes looks for the largest conditional probability value for each class. The technique used to explore public opinion data on application x (twitter) about ChatGPT is google collabs with the results of data mining as much as 1012 data. of these 1012 data cleaning and sentiment analysis using the naive bayes method. Naïve Bayes method classification results with a total of 762 twitter comments about ChatGPT. 100 are used as training data modeled using the naïve bayes method. The accuracy value is 99.00%, positive prediction precision is 100%, negative prediction precision is 96.43%, positive data recall is 98.63%, and negative data recall is 100%.

.

I. INTRODUCTION

IN the era of technological development and information is increasingly widespread. Data and information are easier to obtain using current technology, especially using social media such as Instagram, Facebook, Twitter and others. The information presented in social media generally varies in the form of text, images, videos and so on based on the content grouping of the data and information presented in the social media. In social media, information can be in the form of public opinions containing praise, hate speech, and hoaxes that can result in arguments against the information presented. One of the social media that contains public opinion is twitter. [1]. According to the results of survey data on social media users according to We Are School until January 2024 as many as 139 million user identities. This number is equivalent to 49.9% of the population in Indonesia. There are 27.05 million Twitter users in Indonesia until October 2024. That number puts Indonesia in fourth place globally. The internet plays a very important role in accessing information quickly and efficiently, especially on social media. Most people express their opinions on social media, especially Twitter [2]. Twitter is a social media that disseminates data and information most quickly and accurately compared to other social media [3]. Twitter users make posts, called tweets, of up to 140 characters. Tweets allow Twitter users to interact with each other and exchange thoughts and opinions. Twitter allows researchers to incorporate diverse individual opinions into emotional data for processing and analysis. Sentiment analysis helps develop systems to analyze, identify, and express opinions in textual form. It is also a process used to identify opinions and sentiments from the content of textual data sets about positive and negative topics and events[4][5]. In this case, the application of sentiment analysis can be an analysis of opinions, feelings, evaluations, emotions, assessments or attitudes towards products, personalities, organizations, services, issues or events that occur in a person's community. In addition, sentiment analysis is also always relevant to the community because the source of information comes from social media where the community is the user [1]. Status processing and sentiment analysis of Twitter users' tweets show people's

opinions about artificial intelligence technology that is widely used by the general public, namely ChatGPT. ChatGPT is a Natural Language Processing (NLP) technology that enables answering human questions in the form of text entered into the app. ChatGPT surprised many people, as the answers provided by ChatGPT seemed well-structured, with consistent relationships between words and sentences, and were quite accurate and could recall previous conversations. Even with proper prompting techniques, scientific papers and even books can be written in a much shorter time than traditional methods [6]. However, applications that use artificial intelligence and include large amounts of data can generate negative responses. As ChatGPT increasingly generates human-like text responses, there is always a risk that people will rely on AI-generated content without critically evaluating and validating it. ChatGPT also raises concerns among some educators about the feasibility and usefulness of teaching basic writing and critical thinking skills to students. related to copying assignments directly, the ease with which students can access the ChatGPT app. This raises concerns about education. Using the keyword ChatGPT, many public opinions were expressed on twitter. This research uses a classification algorithm, namely the Naive Bayes algorithm. There are advantages of this classification algorithm in classifying text data, which is the most appropriate algorithm for classifying data on responses from Indonesian people on social media such as Twitter. Basically, Naive Bayes looks for the largest conditional probability value for each class. According to previous research, sentiment analysis on the comparison of Naïve Bayes and C4.5 methods against the president of 3 periods resulted in a difference in accuracy that naïve bayes was higher at 85% compared to C4.5 which was only 78%[7]. Support Vector Machine (SVM) and Naïve Bayes methods on textual reviews on Google Play resulted in classification with the highest accuracy[8]. Sentiment analysis on the comparison of SVM and KNN methods on the gojek application found that SVM has higher accuracy than KNN[9]. The number of ChatGPT users causes ChatGPT to get a lot of opinions written on twitter, then these opinions will be classified into 2 groups, namely positive and negative opinions. Therefore, the author raises the research theme in the form of a thesis with the title analisys the public opinion sentiment on application X (Twitter) in Indonesia on ChatGPT using Naïve Bayes algoritma.

II. RESEARCH METHODS

A. *ChatGPT*

ChatGPT has the ability to generate text that matches the context of the conversation. ChatGPT uses a Transformer architecture called selfattention model. This architecture allows ChatGPT to handle complex conversations by referring to the context of previous conversations. In addition, since ChatGPT is trained to handle sizable data, it can easily understand user questions and expectations [6]. The way ChatGPT works is by collecting various information from journals, articles, newspapers that have been published on the internet and then ChatGPT absorbs it all so that when there is a person or user who is looking for information about what he wants to know, ChatGPT will conclude an answer based on the information it has collected in a concise time [10].

B. *X (Twitter)*

Twitter, a social media platform, was founded by Jack Dorsey in 2006 [11]. Twitter is a social media platform with various formats, focusing on the dissemination of short, concise, and real-time messages with sentences of less than 140 characters to readers worldwide [12] on November 7, 2017, it increased to 280 characters in tweets. Initially, Twitter was known for its bright blue and white logo and its cute bird named "Larry Bird." In the end, on April 22, 2022, Twitter's editorial board approved the purchase of the app by Elon Musk, CEO of SpaceX and Tesla, Inc., for $44 billion, or around Rp 635 trillion, and agreed to transform Twitter from a social media platform into a private ownership. After that, Elon Musk changed the Blue Bird Twitter logo to a simple X-shaped logo, signaling the end of the Twitter era. After Twitter was taken over by Elon Musk in 2023, its iconic "Blue Bird" logo was removed. On the contrary, Twitter uses a simple black and white X logo.

C. *Sentimen Analisys*

Sentiment analysis is an activity used to analyze someone's opinion or perspective on a topic. The basic task of sentiment analysis is to classify several sentences or texts from documents, sentences or features, and sentences that can be positive, negative, or neutral [13]. Sentiment analysis is one of the most popular examples in the field of Natural Language Processing (NLP). Natural Language Processing (NLP) is a scientific field that discusses how to make computers work and think like humans. Natural Language Processing (NLP) is a part of Artificial Intelligence or artificial intelligence. In the development of data mining, Artificial Intelligence (AI) is one of the four branches of data mining, namely statistics, databases, and information retrieval. In its application, Artificial Intelligence (AI) also requires machine learning as a solution algorithm. The presence of machine learning is used to replace humans in decision-making. Machine learning does not have feelings like humans, so the decisions made

are based on processed data [14]. The results of sentiment analysis can also serve as a reference for companies, public figures, and governments to determine their next steps [15]. There are several types of sentiment analysis, namely emotion detection, aspect-based sentiment analysis, and fine-grained sentiment analysis. Fine-grained sentiment analysis is a type of analysis that has specific assessments and is commonly used in the e-commerce field. Emoticon detection is a type of analysis aimed at identifying the emotions present in a message, such as happiness, sadness, anger, and so on. Aspect-based sentiment analysis is a type of analysis to identify influential aspects and customer evaluations. [16]

D. Text Preprocessing

The data set must go through the text preprocessing stage first because the data set cannot be used without going through the data management stage. Text preprocessing is a process of managing the data set before the data is processed. In reality, there are still many datasets that are not clean, such as system errors during recording that result in duplicate data. Unprocessed or unclean data falls into categories such as irregular data formats, the presence of empty data, different data types, the presence of unimportant attributes, and so on. The cleaner the preprocessing conducted, the more likely the resulting data will be accurate[14]. There are stages of text preprocessing, as follows [17]:

1) Cleaning

Cleaning is the process of removing words that are superfluous, meaningless, or that influence sentiment from the data collection, including HTML, links, mentions, and hashtags

2) Tokenize

Tokenization is the process of breaking down a data set into tokens or word segments to facilitate the next stages. For example, the sentence "saya ingin makan" is segmented into ["saya", "ingin", "makan"]. [18]

3) Transform Case

The process of converting the text data sentences into uniform text is known as the transform case. Because the current data is not always structured in its letter usage, this step is always included in the text preprocessing process. Standardizing the use of capital letters can be aided by this stage. For instance, the words "Data" and "data" will be interpreted as two distinct words, allowing the system to read efficiently.

4) Stopword removal

Stopwords are words that frequently appear but don't have any important meanings or effects on the system, such as "oh," "di," "pada," and so on. [19]

5) Filter

Filter is the stage for removing words that are too short and too long, with a minimum of 3 letters and a maximum of 25 letters.[20]

E. Naïve Bayes

The Naive Bayes algorithm is based on the simplifying assumption that attribute values are conditionally independent of each other given an output value. One type of statistical model is the concept of probability. one approach that uses the idea of probability is the naïve bayes algorithm. In this way, each attribute will contribute to the decision making, with each attribute weighted equally in importance, and each attribute independent of each other [21].

In this study, we will compare using the Naïve Bayes method from these algorithms to determine the level of accuracy and the results of twitter user sentiment whether positive or negative which can affect people's views on the ChatGPT application using a quantitative approach. Microsoft Word and rapidminer are tools that help in report preparation and data processing from twitter.

A. Research Stages

The research stages are used by researchers in carrying out research. The research stages are used as procedures that explain the process of running a study. Figure 1 below is the research stage:

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

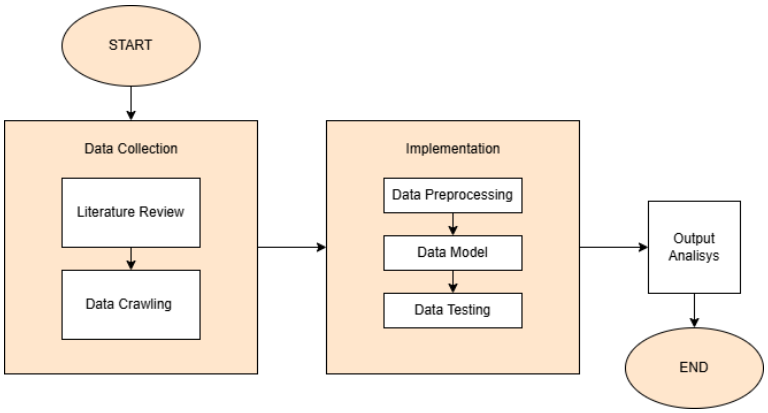

Fig. 1. Research Stages

B. *Data Collection*

In the data collection method used in this research, there are 2 ways, namely literature review and data crawling.

1) *Literature Review*

The literature study method is a data collection technique by collecting data or seeking information from books, journals, internet media, and the results of previous research related to the problem under study, in this study the authors collected data by searching and studying from various sources, such as books, journals, and internet articles related to the issues raised. In this study, researchers collected articles in scientific journals related to the topic taken, namely sentiment analysis on twitter on ChatGPT using machine learning algorithms. Various machine learning algorithms, namely naïve bayes, SVM, KNN, C4.5. and looking for book references related to the theories taken in the theoretical basis chapter.

2) *Data Crawling*

Crawling data in this research aims to retrieve data from X social media sources using twitter technology. The investigated data was taken from social media X with the keyword "ChatGPT" and was taken in the form of text, which is a Twitter user's tweet. This data was collected from June 29, 2024 to September 30, 2024. Figure 2 below shows the program code in crawling data:

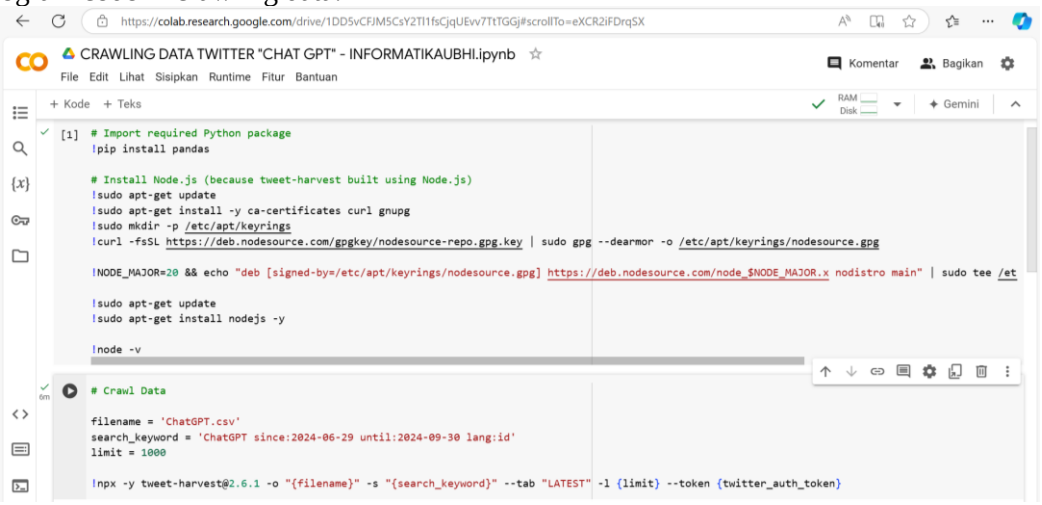

Fig. 2. Crawling data using Google Collabs

This research uses Google Collabs which is stored in the form of a .csv file for crawling data. Researchers connect Google Collabs to X social media by entering X social media tokens using the @Ubhinformatics account. Search tweets using the keyword "ChatGPT" from June 29, 2024 to September 30, 2024 with a limit of 1000. The crawling data file is in .CSV format.

C. *Implementation Simulation*

1) *Data Preprocessing*

Pada tahap preprocessing data, dilakukan pengolahan data setelah crawling data dilakukan. Data yang diambil dalam tahap crawling data adalah opini masyarakat terhadap ChatGPT. Pada tahap ini, dilakukan pengolahan data untuk mendapatkan data yang bisa digunakan untuk proses yang setelahnya yaitu pengujian data. Pada preprocessing melibatkan cleaning, transform cases, tokenize, stopword removal, dan filtering sehingga menjadi data terstruktur. Setelah itu, dilakukan pelabelan data berdasarkan kelas data tersebut untuk menentukan opini positif dan opini negatif menggunakan pelabelan manual sebagian dan automatis mengacu pada dictionary.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

2) Data Model

After the data preprocessing stage, data labeling is carried out based on the data class to determine positive opinions and negative opinions using partial manual and automatic labeling referring to the dictionary. Then processed using both methods, namely Naïve Bayes.

3) Data testing

After the data preprocessing stage, the data set in the form of train data and test data is processed based on both methods, namely Naïve Bayes. Figure 3 below is the implementation of the naive bayes method:

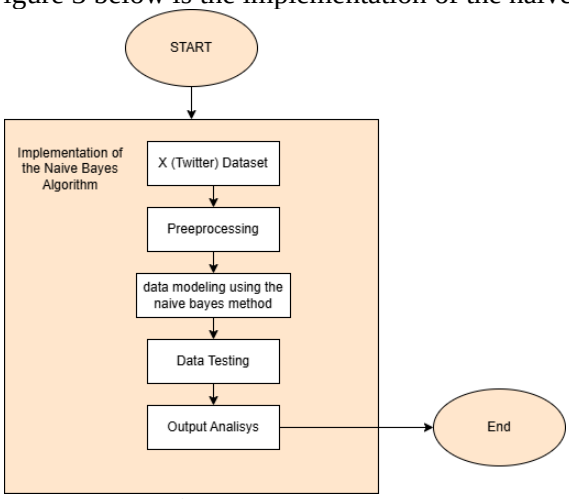

Fig. 3. Stages of Sentiment Analysis Using Naïve Bayes

D. Output Analisys Simulation

The last stage in this research is to analyze the output results of sentiment analysis that has been done. At this stage comparing the results of the classification of the test data with the sentiment label on the test data that was previously obtained. The result of this evaluation is the accuracy value obtained after validating based on the confusion matix value. In the confusion matrix, accuracy, precision, and recall are calculated.

III. RESULT AND DISCUSSION

A. Creating a Sentiment Analysis Dataset in Rapidminer

The first step is to prepare the Rapid Miner application which will be used in processing tweet data. On the operators panel in Rapid miner, find the Read .CSV operator and drag and drop it to the process page. Then the Read .CSV operator is used in uploading the tweet result data that has been crawled on google collabs. To make it easier to analyze the data, of the 14 columns, only the text attribute is displayed. Figure 4 below is the result of the select attributes filter:

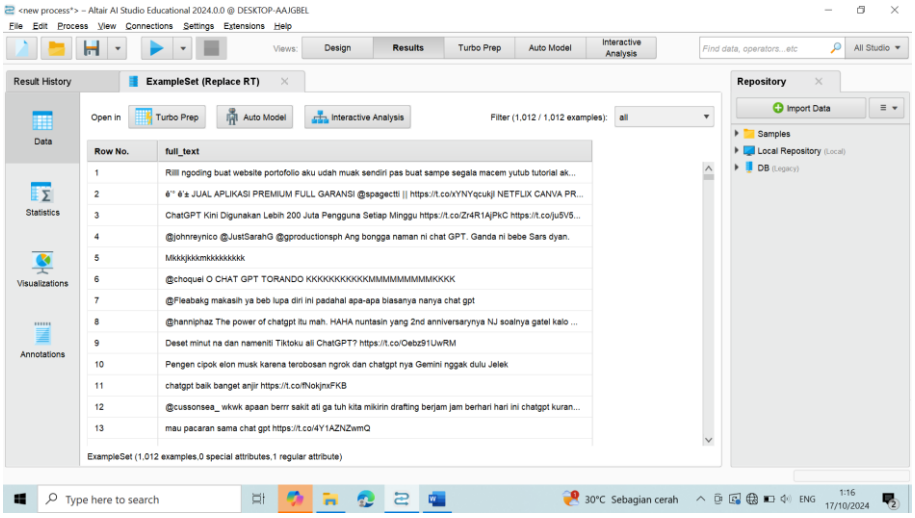

Fig. 4. Result of the Select Attributes Filter

1) Modify

The data set consisting of text preprocessing is modified. This stage is intended to improve the computer system's understanding of the shape of the data set. This activity can also turn unorganized data into well-organized data. There are five processes in the text preprocessing process including cleaning, tokenizing, transform cases, and

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

filtering.

2) Cleaning

In this process, data cleaning is performed. Operators used are replace RT used to remove reetwet, replace URL 1 used to remove URLs at the beginning and middle of the sentence, replace URL 2 used to remove URLs at the end of the sentence, replace hastag used to remove hastags, replace mention used to remove mentions, and replace symbols used in removing symbols that have no meaning. The trim operator is used to remove whitespace. An example of cleaning on a data set is presented in Figure 5:

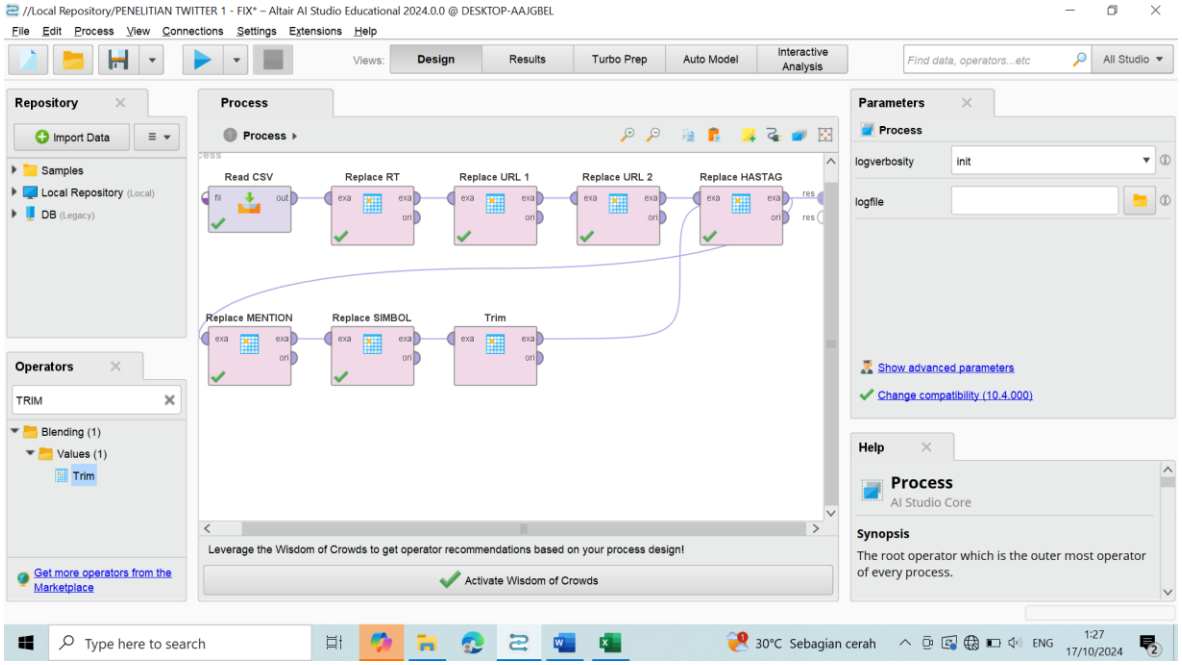

Fig. 5. Data Cleaning Process

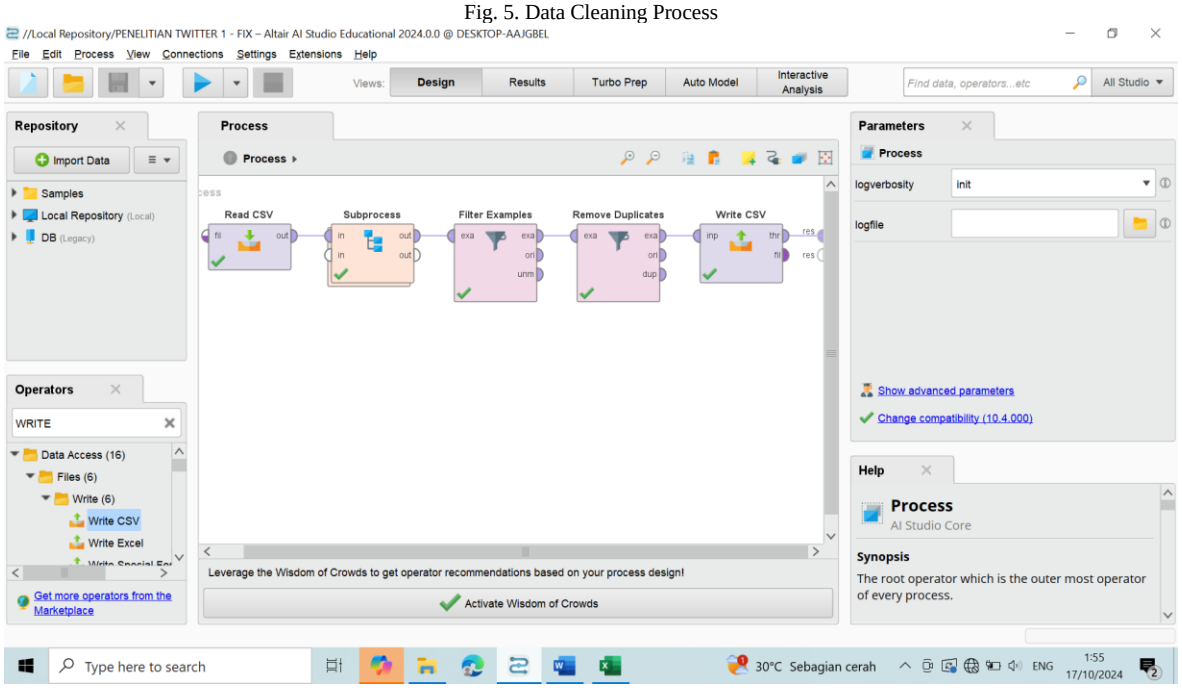

Fig. 6. Filter Examples and Remove Duplicated Process

Figure 6 describes the advanced cleaning process. Operators replace mention, hastag, retweet, URL, symbol and trim are put together into subprocess operators to make it look more efficient. Then the results of cleaning mention, hastag, retweet, URL, symbol and trim are continued with filter examples operators used to filter out empty data. Remove duplicated is used to remove duplicate data. After the data is filtered, continue with the use of write CSV operators. Write CSV is used to store the results of the entire data cleaning process. The data after going through the remove duplicate process amounted to 1001 tweets. Figure 7 below is an image of the overall cleaning results:

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

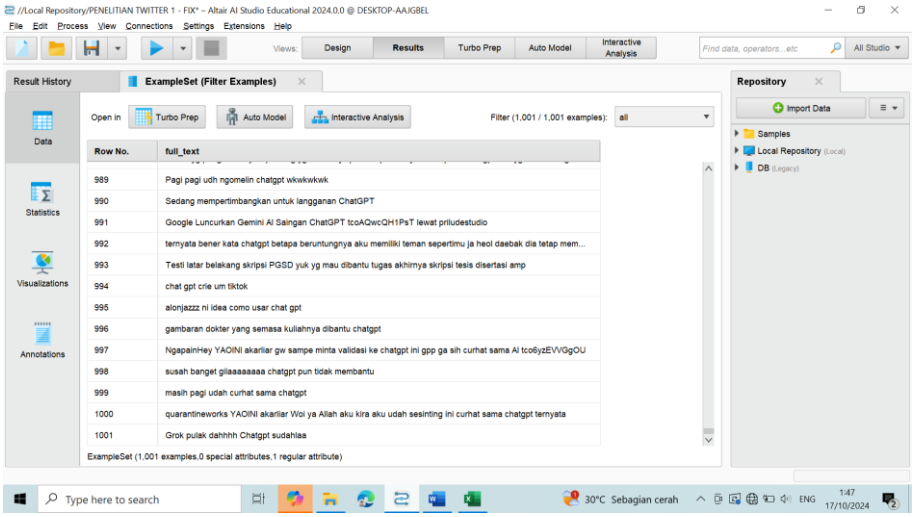

Fig. 7. Figure of Cleaning Filtering Examples and Remove Duplicated Results

Furthermore, after using the write CSV operators, when run the data will be saved in excel. From the 1001 data, the author analyzes the categories that fall into negative or positive sentences. In addition to analyzing positive and negative sentences, the author also cleaned the data again for the 1001 data that did not use Indonesian and contained advertising elements. This activity takes place manually using excel. The results of the analysis will later be put back into Rapid Miner for further processing. The total cleaning data becomes 762 and is divided into 2 data, namely data that is made into positive and negative sentences totaling 100, and those that are not labeled are 662. Figure 8 below is the result of analyzing positive and negative sentences:

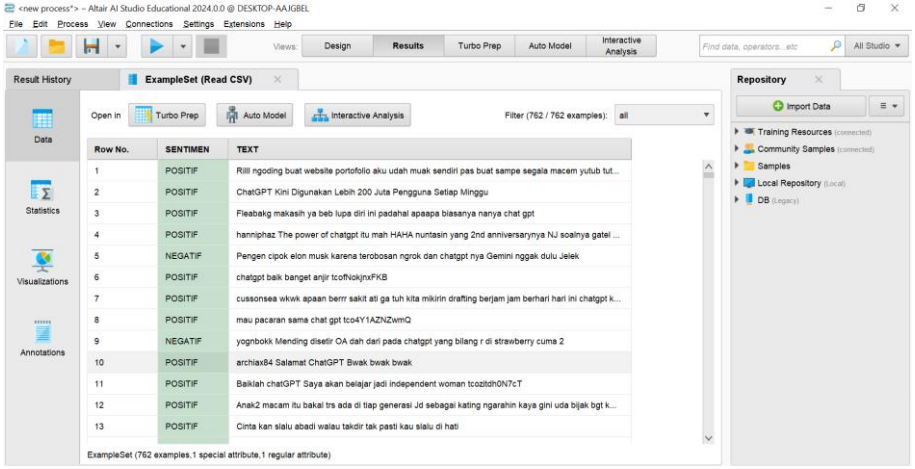

Fig. 8. Results of Positive and Negative Sentence Labeling

Next, the author will divide the overall data which will be processed for sentiment analysis using a machine learning algorithm, namely naïve Bayes. The first data division is taken from data that has positive and negative labels, because the data will be used as testing data. Figure 9 below is the data division process:

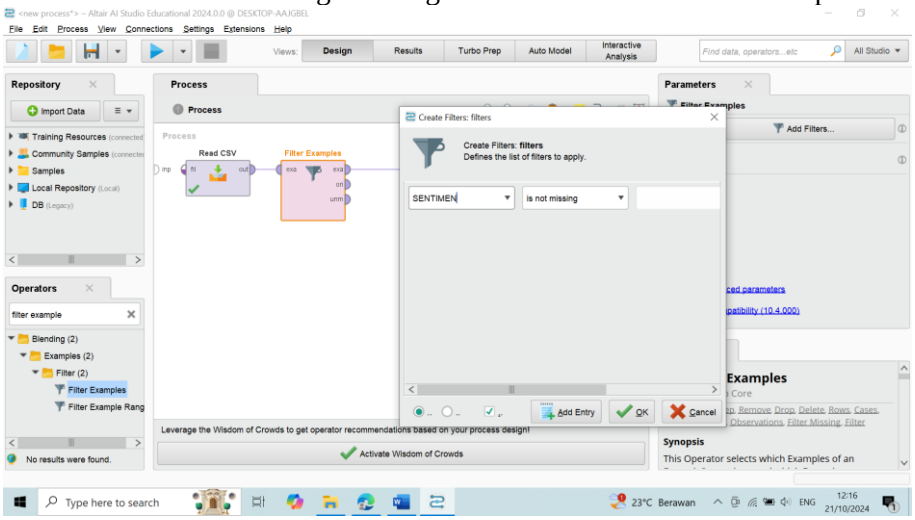

Fig. 9. Filter Examples process

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

In Figure 9 when sharing data using filter examples operators. In its settings, the setiment label is selected with "is not missing" data, meaning that the data is taken data that has a label. Figure 10 is the result of dividing data that has a label:

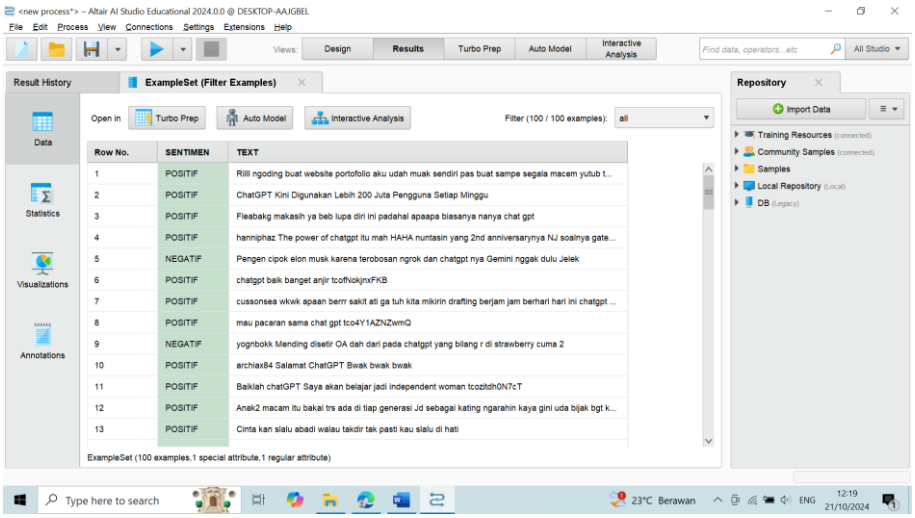

Fig. 10. Results of Dividing Positive and Negative Label Data

3) Tokenize

After the data is filtered examples, the sentences in the data will be separated into words so that Rapid Miner can recognize the data. The process uses Process Documents From Data operators, which when clicked twice later the author can enter the process in it. Furthermore, when the process documents form data operator is clicked 2 times, it will enter a new page. Next, the author will use tokenize operators. The tokenization process is the process of breaking the sentence into pieces of words or tokens to find out the origin of the word. From 100 data, 788 different words are generated. The meaning of the value 0 is that the data does not appear in the sentence. Here is figure 11 showing the tokenization results on the dataset.

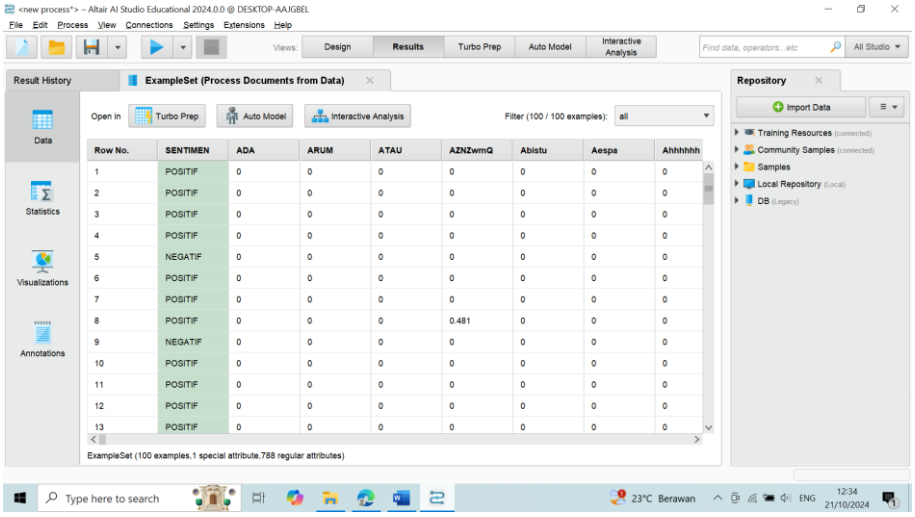

Fig. 11. Tokenized Data

4) Transform Case

Transform case is the process of equalizing letters from capital letters to lowercase letters or vice versa, because rapid miners do not analyze the same data when there are differences in the shape of capital letters. In this research, the data set is converted into all lowercase letters because the tweets that have been taken mostly use lowercase letters.

5) Stopword Removal

This process is done to remove words that are widely used but do not affect the sentiment of a sentence. Stopword removal in this study uses a dictionary which is an Indonesian-language stopword corpus. After doing the stopword filter, the resulting data amounted to 593.

6) Filter Tokens (by Length)

In this process, words that have letters that are too short or abbreviated and letters that are too long are removed. The following figure 12 is the Tokens Filter process (by Length):

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

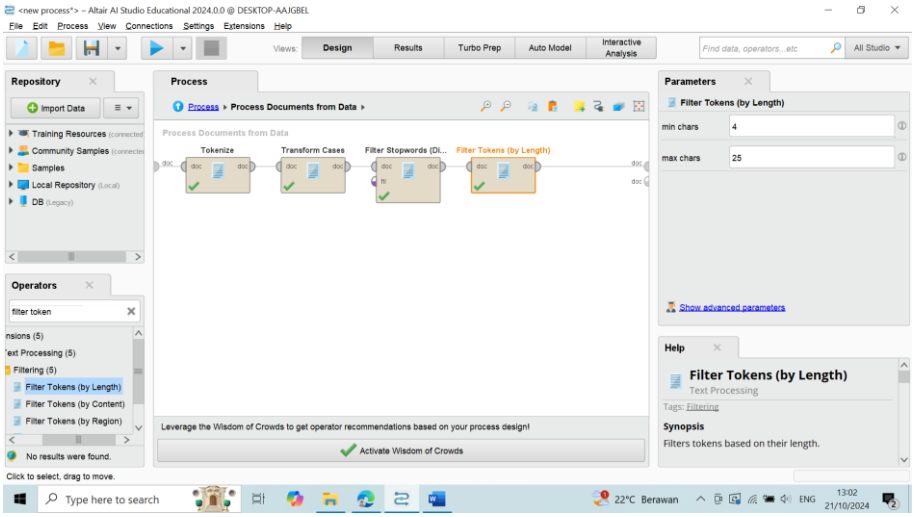

Fig. 12. Filter Tokens (by Length) Process

7) WordCloud

Wordcloud is a visualization of data sets taken from words that often appear in tweets. In the results of data visualization, the biggest result is the ChatGPT writing, followed by Chat. Figure 13 is an example of data visualization with wordcloud:

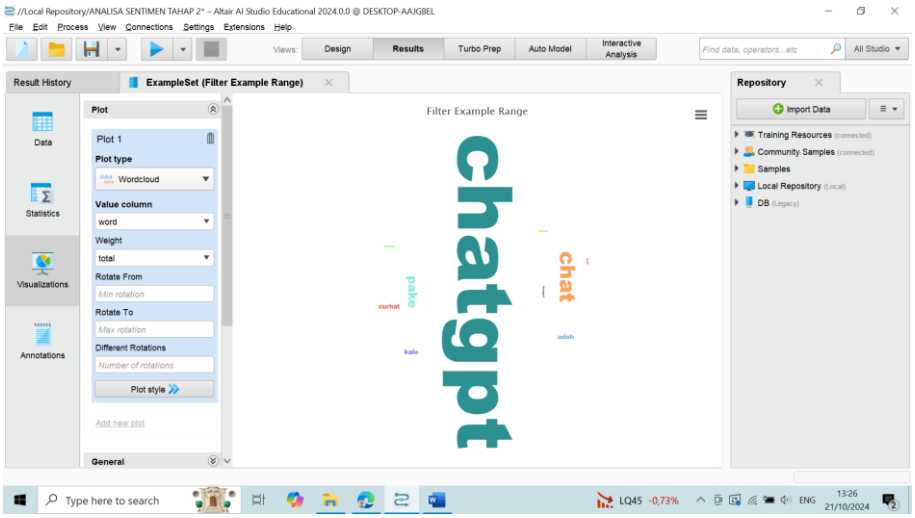

Fig. 13. Visualization Results Using Wordcloud

B. Data Model

The modeling stage in this research includes manual labeling and two classification modeling with the Naïve Bayes method.

1) Naïve Bayes Process Training Data

In this study, the data set is divided into two parts, namely test data and training data with a test and train ratio of 100 training data and 662 test data. The total data is 762 data which is then carried out the filter examples process for sharing training data to produce 100 data, then processed the process document and then connected with naïve bayes operators. Naïve Bayes is used to determine the sentiment analysis of sentences that have no data label.

After the naïve bayes process data will be stored in the store operators. Store model is divided into 2 namely store model and store training data. Store model is used to store the model of training data that is processed using naïve bayes, while the training data store is used to store the results of training data that has been filtered. The process can be seen in Figure 14 below:

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

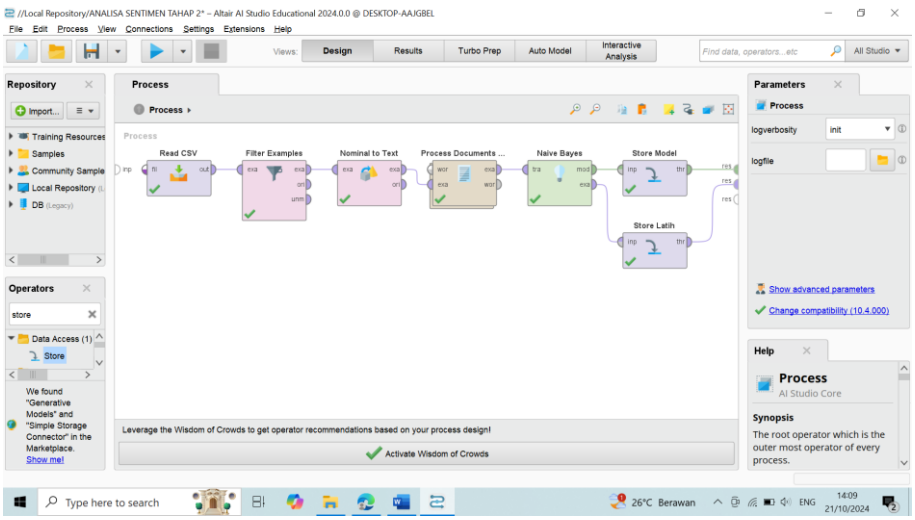

Fig. 14. Process of Storing the Model and Training Data

2) Naïve Bayes Sentiment Analysis Process

In the next stage, a new page is created for sentiment analysis using the naïve bayes model. The data in the cleaning stage is then re-entered using the read csv operator. After the data is entered, filter examples are used to divide the test data by selecting the "missing value" option which is used to display empty sentiment data labels. After that, process document is done to break sentences into more specific words. Union operators are used to combine training data attributes with test data. The incoming training data will be filtered to fulfill the test data. Figure 15 is the result of sentiment analysis of positive comments and negative comments:

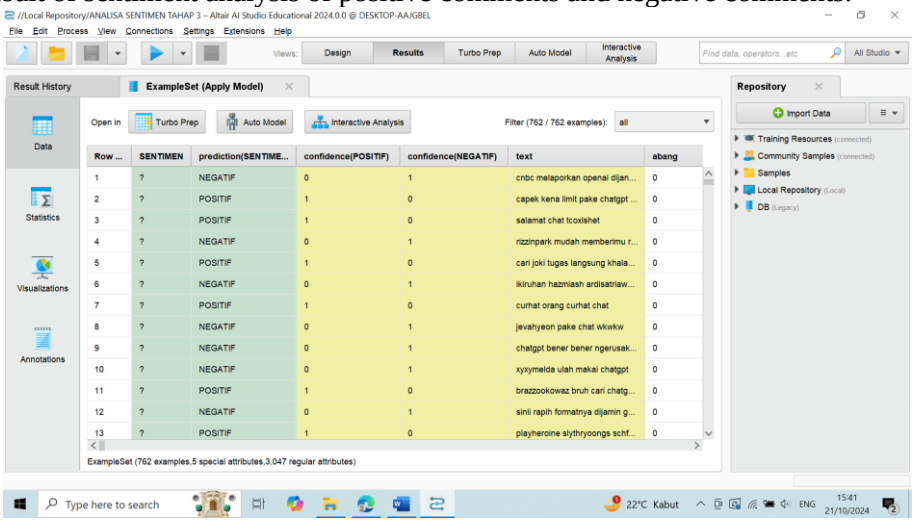

Fig. 15. Positive and Negative Sentiment Analysis Results

The results of the positive and negative sentiment analysis then find the accuracy level of the calculation. The accuracy value obtained by the Naïve Bayes algorithm is 99.00%, meaning that a total of 99.00% of the Naïve Bayes model can classify the correct data. The margin of error means the amount of error in sampling is +- 58 9.10%. There are 72 positive sentiments and 27 negative sentiments. Precision is the ratio between the prediction results and the requested data, which is 100% in positive predictions and 96.43% in negative predictions. The recall value describes the success of the model in retrieving information with a comparison between the ratio of positive correct predictions and the overall positive prediction data. The results obtained on positive data are 100% and on negative data 96.63%. Figure 16 is the result of calculating accuracy, precision, and recall using rapid miner:

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 9, No. 4, Desember 2024, Pp. 2473-2484

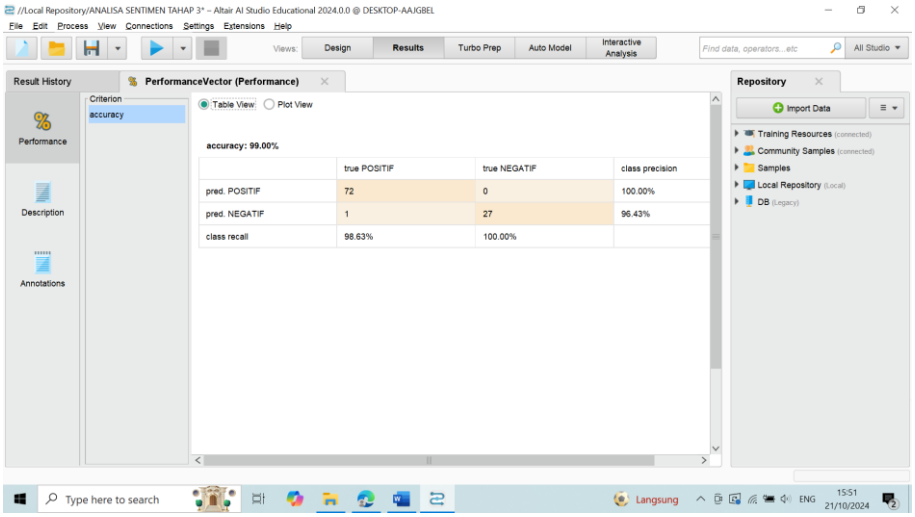

Fig. 16. Accuracy Values of the Naïve Bayes Algorithm

IV. CONCLUSIONS AND SUGGESTIONS

A. *Conclusion*

Naïve Bayes method classification results with a total of 762 twitter comments about ChatGPT. 100 were used as training data modeled using the naïve bayes method. The accuracy value is 99.00%, positive prediction precision is 100%, negative prediction precision is 96.43%, positive data recall is 98.63%, and negative data recall is 100%.

B. *Suggestions*

The author's suggestions for further research:

1. Future research can use other social media such as Facebook or Instagram.
2. Further research can use other applications with programming languages such as python or r language.
3. Future research can add neutral sentiment.
4. Further research can use other than Indonesian language

REFERENCES

[1] Rusdiaman, D., & Rosiyadi, D. (2019). *Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naïve Bayes Classifier Dan Support Vector Machine* (Vol. 4, Issue 2).
[2] Anjani, A. M., Chamid, A. A., & Murti, A. C. (2022). *Analisis Sentimen Kaum Lgbt Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes* (Vol. 1, Issue 2). Halaman 1~8.
[3] Fitra Rizki, M., Auliasari, K., & Primaswara Prasetya, R. (2021). *Analisis Sentiment Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine.* In Jurnal Mahasiswa Teknik Informatika) (Vol. 5, Issue 2).
[4] 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.
[5] Berliana, D. R., & Santoso, B. (2022). *Elektabilitas Ridwan Kamil Dan Anies Baswedan Dalam Simulasi Pilpres 2024 Di Twitter (Analisis Jaringan Media Sosial Dan Analisis Sentimen Pengguna Twitter Terhadap #Ridwankamil Dan #Aniesbaswedan).* Mediakom : Jurnal Ilmu Komunikasi, 6(2), 150–162. https://doi.org/10.35760/mkm.2022.v6i2.6962
[6] Setiawan, A., & Luthfiyani, U. K. (2023). *Penggunaan ChatGPT Untuk Pendidikan di Era Education 4.0: Usulan Inovasi Meningkatkan Keterampilan Menulis*. Jurnal PETISI, 04(01).
[7] Albasithu, F., & Wibowo, A. (2022). *Perbandingan Algoritma Naïve Bayes Dan C4.5 pada Analisis Sentimen Presiden 3 Periode di Twitter*. In Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia. https://senafti.budiluhur.ac.id/index.php/
[8] Ilmawan, L. B., & Mude, M. A. (2020). *Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store*. ILKOM Jurnal Ilmiah, 12(2), 154–161. https://doi.org/10.33096/ilkom.v12i2.597.154-161
[9] Muttaqin, M. N., & Kharisudin, I. (2021). *Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor.* UNNES Journal of Mathematics, 10(2), 22–27. http://journal.unnes.ac.id/sju/index.php/ujm
[10] Suharmawan, W. (2023). *Pemanfaatan Chat Gpt Dalam Duniaa Pendidikan*. Education Journal : Journal Education Research and Development, 7(2), 158–166
[11] Pasaribu, B. M., Winoto, Y., Lies, U., & Khadijah, S. (n.d.). *Twitter Sebagai Media Pemenuhan Kebutuhan Informasi Penggemar Musik Korea di Indonesia.* 10, 7–15. http://ejournal.unp.ac.id/index.php/iipk/index.
[12] Annisa Batu Bara, E., Amelia Nasution, K., Zahara Ginting, R., & Studi Ilmu Perpustakaan, P. (n.d.). (2023). *Penelitian tentang Twitter*
[13] Fahlevvi, M. R. (2022). *Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine*. Jurnal Teknologi Dan Komunikasi Pemerintahan, 4(1), 1–13. http://ejournal.ipdn.ac.id/JTKP,
[14] Irwansyah Saputra, D. A. K. (2022). *Machine Learning untuk Pemula*. Informatika Bandung.
[15] Natasuwarna, A. P. (2020). *Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring*. Techno.Com, 19(4), 437–448. https://doi.org/10.33633/tc.v19i4.4044
[16] Arviana, G. N. (2021). *Sentiment Analysis, Teknik untuk Pahami Maksud di Balik Opini Pelanggan*. 1 Februari.
[17] Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). *Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm.* Jdmsi, 2(1), 31–37.

[18] Baihaqi, W. M., Pinilih, M., Rohmah, M., Informasi, T., Purwokerto, U. A., Informasi, S., Purwokerto, U. A., & Korespondensi, P. (2020). *Kombinasi K-Means Dan Support Vector Machine (Svm) Untuk K-Means and Support Vector Machine (Svm) Combination To Predict Sara Elements on Tweet.* Jurnal Teknologi Informasi Dan Ilmu Komputer, 7(3), 501–510. https://doi.org/10.25126/jtiik.202072126

[19] Taufiqurrahman, F., Faraby, S. Al, & Purbolaksono, M. D. (2021). *Klasifikasi Teks Multi Label pada Hadis Terjemahan Bahasa Indonesia Menggunakan Chi Square dan SVM. E-Proceeding of Engineering*, 8(5), 10650–10659.

[20] Hajiati, S. R. I., Informatika, P. S., Komunikasi, F., Informatika, D. A. N., & Surakarta, U. M. (2021). *Perbandingan Tingkat Akurasi Algoritma Restricted Boltzmann Machine Dan Backpropagation Dalam.*

[21] Rahmawati, N. (2019*). Klasifikasi Kondisi Gizi Balita Menggunakan Metode Naive Bayes (Studi Kasus Posyandu Melati IV).* Jurnal Ilmiah Mahasiswa Teknik Informatika, 2(3).http://repository.unama.ac.id/id/eprint/980