

Sentiment Analysis of Delivery Services Using the Sastrawi Library and Biobii Naive Bayes Text Classifier

Eko Budi Setiawan¹, Muhammad Lutfie Zakaria²
Universitas Komputer Indonesia, Indonesia

Corresponding Author: Eko Budi Setiawan
Department Informatics Engineering
Universitas Komputer Indonesia, Indonesia
Email: eko@email.unikom.ac.id

Article History

Received Dec 13, 2025
Accepted Jan 22, 2026
Published June 2026

Keywords

Sentiment Analysis, Text
Processing, Naïve Bayes
Text Classifier, Sastrawi

Abstract

Sentiment analysis from social media is an important aspect to find out people's views. This research produces a web-based system that can analyze public sentiment towards a delivery service based on social media Twitter (currently X). The technology used is the Twitter(X) API, Sastrawi library and Biobii Naïve Bayes text classifier. The Twitter(X) API is used to retrieve data from the Twitter(X) social media platform, while the Sastrawi library is used to process and analyze texts in Indonesian. In this system users can enter certain keywords or hashtags for analysis. The collected tweet data is then processed using a customized Naïve Bayes text classifier Library to identify positive and negative sentiments towards the services offered by a delivery service company. The results of sentiment analysis are displayed in the form of informative graphical visualizations, providing a better understanding of customer views regarding the service. The accuracy of the sentiment analysis results produced was based on testing with one of the delivery service companies, obtaining results with a value of 93.75% accurate.



This is an open access article licensed a Creative Commons Attribution-ShareAlike 4.0 International License.

1 INTRODUCTION

Technology is basically created to make it easier for humans to do things. Many people use this information technology to process, process and analyze data, which produces data or information that is relevant, fast, clear and accurate. Government organizations, private companies, and others have made extensive use of information technology. Information technology has brought new changes in society, including in the business world. Entrepreneurs are now using this technology to develop their businesses in the form of e-business in order to increase competitive advantage [1].

E-business carried out in various business sectors. This includes delivery services which really need support from the use of technology. Many businesses in delivery services are integrated into E-commerce. Today's e-commerce is very superior. All necessary needs easily available without having to bother going to a shopping place first. With the high number of transactions via E-commerce, it will also affect the goods delivery service business resulting from transactions carried out via E-commerce [2].

These technological developments encourage the trading industry to adopt new trading methods via the internet [3]. Increasing the delivery of goods, both from within the country and from abroad, is one of the main drivers of a country's economic growth

[4]. Many sellers who transact on e-commerce platforms need delivery services to deliver their customers' orders to their destination, so it is not surprising that there are many delivery service companies [5].

The number of private companies in the courier services sector registered as members of the Association of Indonesian Express, Postal and Logistics Distribution Services Companies is 560. PT IDEXpress Logistik Indonesia, or better known as IDEXpress, is a company operating in the delivery and logistics services sector which is headquartered in Jakarta, Indonesia.

Twitter(X) is the social media with the most users among several social media with a total of 328 million users [6]. Twitter(X) in Indonesia is not only used to express feelings, but is also used to suggest and criticize policies, goods or services [7], according to reports from Hootsuite and We Are Social, in January 2023 there were 556 million. Twitter(X) users worldwide, an increase of 27.4% from the same period the previous year. Twitter(X) active users in Indonesia reached 50 million users [8].

There are many comments on Twitter(X) from IDEXpress customers which can be used to find out public opinion about the expedition service, due to the difficulty and lack of knowledge of branch heads regarding easier and more efficient ways to obtain opinions circulating on Twitter(X) social media which have positive or negative meanings from the customer's point of view [9] which is on Twitter(X) social media for IDEXpress, which is useful for improving the office's performance [10].

Collecting opinion data from Twitter(X) users obtained through the use and utilization of the Twitter(X) Application Programming Interface (API) [11]. By using the Twitter (X) API, software developers can utilize data sourced from Twitter(X) social media to process it into more useful information. The data exchange used in the Twitter(X) API uses a REST architecture which can make data management easier [12], because it can provide data formats in the form of Extensible Markup Language or data formats in the form of JavaScript Object Notation [13].

Much research on sentiment analysis has been conducted, most using comment data from Facebook, Twitter, online news, blogs, and others. However, not many have concentrated on managing Indonesian language texts [14]. Sentiment analysis is a method for understanding, extracting opinion data, and processing text data automatically to obtain the sentiment contained in opinions [15]. The aim of sentiment analysis is to find patterns in an article, both positive and negative opinions [16].

Research related to Naïve Bayes has also been carried out in various a journal. Sentiment analysis and text mining methodologies, as well as the use of Naïve Bayes classification in opinion mining via Twitter(X) data [17]. The Naive Bayes method chosen in this research because it has simple calculations and is very accurate and fast in large amounts of data [18].

The Sastrawi Library is a library that is used to carry out the stemming process for texts that use Indonesian. This library offers a stemming algorithm derived from the Nazief and Adriani method, which is then strengthened with the Confix Stripping and Enhanced Confix Stripping algorithms. In that research the Sastrawi library provides increased stemming accuracy, faster processing time, and increased document classification accuracy in preprocessing student complaint documents [19].

2 METHODS

Research methods are the working methods used to collect and process data in order to achieve research objectives. The software development method used in this research is the Software Development Life Cycle (SDLC) method with a prototype model.

A. System Architecture

The system that will be implemented is a website that can retrieve tweet data based on keywords and services which will later be processed and classified using libraries. System architecture is a general description of systems that work and are interconnected with each other. Figure 1 is diagram system architecture in this research.

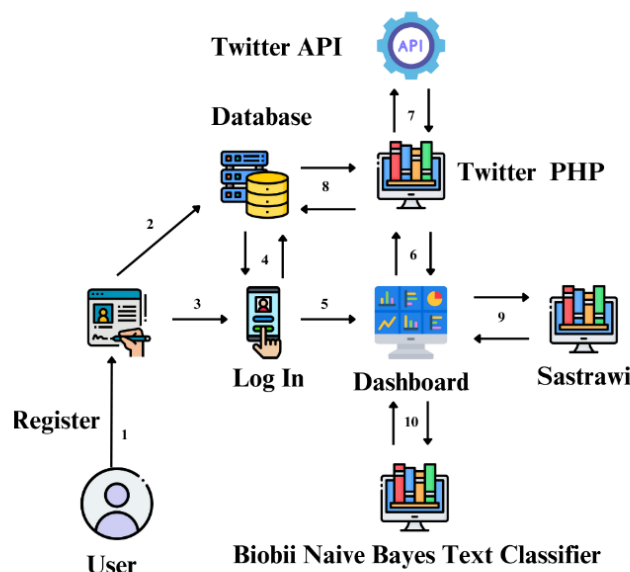


Figure 1. Diagram System architecture

The explanation of system architecture is:

1. Users go to the register page to register an account.
2. The account will be entered into the database as user auth.
3. User logs in with credential identity.
4. Authenticate the user and log in to the application system.
5. Users enter the dashboard page and search based on service keywords.
6. The Twitter(X) PHP library will request data based on keywords.
7. Twitter(X) API responds and then sends data information.
8. The library saves the data to a database, then displays the data to the user.
9. The tweet data that has been retrieved will be cleaned by the literary library or text processing.

B. Collecting Data Method

In this research, the data collection process carried out by taking several random tweets on Twitter(X) accounts using certain keywords related to delivery-related services and then the data classified into two sentiments, namely positive sentiments and negative sentiments. This data is used as training data during the classification process. The training data used is 100 tweet data with 50 positive sentiments and 50 negative sentiments. This training data is stored in the database as learning data from the library used. The following is an example of training data which can see in table 1.

Table 1. Example Data of Collection Training

Text Data Training	Classification
Text Data Training	Positive
I am very satisfied with their delivery service. My package arrived on time.	Negative
My delivery was delayed for two days. Very unprofessional!	Positive
The customer service team's response was very fast and helpful. They helped me resolve my shipping issues quickly.	Negative
My package was damaged when it arrived. Very disappointed with the condition of the item.	Positive

C. Twitter(X) API Technology

The Twitter(X) API is a programming interface for applications that allows developers to access and integrate data from the Twitter(X) platform into third-party applications or services. These APIs allow developers to do things like send tweets, send direct messages, and manage user lists as well as retrieve data such as followers, lists, and tweets [20]. Developers can use the Twitter(X) API to connect their applications to the platform and take advantage of the features and data it provides. Overall, the Twitter(X) API allows access, analysis, and interaction with Twitter data, making it a useful tool for a variety of purposes.

Twitter endpoint has unique functionality that allows users to retrieve or send certain data to or from Twitter. Endpoint is the URL or endpoint used to access and send requests to the Twitter(X) API. There are endpoints that function to retrieve tweet data, send new tweets, search for tweets based on keywords, and so on. Users can use these endpoints by sending an HTTP GET or POST request to the appropriate URL [21].

In this research, the Twitter(X) API is used by the system to be able to search for Tweets and also provide information and retrieve Tweet data on Twitter(X) based on keywords and conditions that have been designed in the application system that was built. The Figure 2 is the Twitter(X) API workflow:

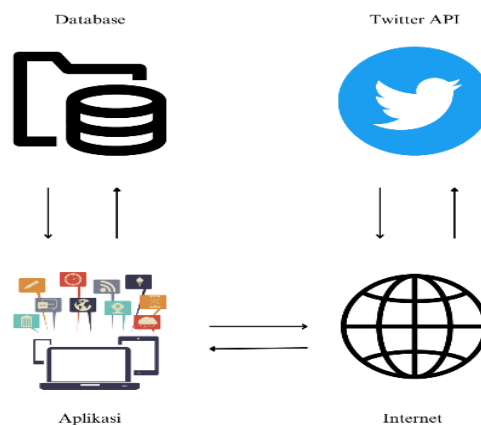


Figure 2. Workflow of Twitter(X) API

D. Tweet Data Crawling Stages

Crawling is the process of automatically collecting data from one or various data sources [22]. In this research, crawling was carried out on the Twitter(X) platform to obtain tweet data related to the entered keywords. The Twitter(X) API is used to search for Tweets based on keywords that the user has entered. The Twitter(X) API will respond to requests from the application and then send tweet data which is automatically saved to the database and displayed in the application. Figure 3 is the process flow for the search feature using the Twitter(X) API.

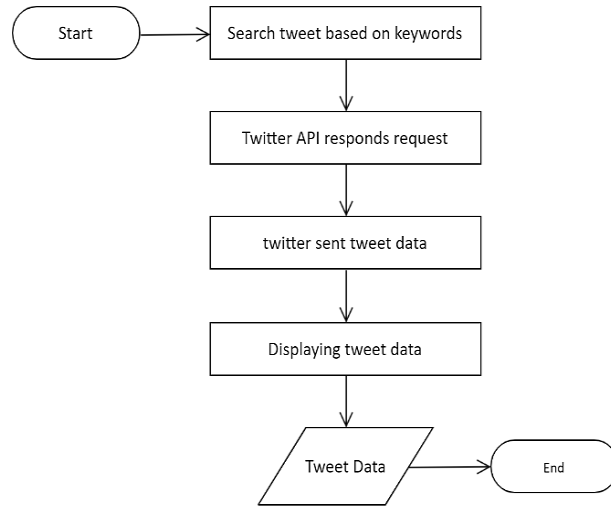


Figure 3. Twitter(X) data search and crawling flowchart

E. Preprocessing Process

Preprocessing is the sentiment analysis phase where the captured tweet data is selected and cleaned. Cleaning, tokenization, case folding, filtering (or removal of stopwords), and stemming are some of the preprocessing stages [23]. Punctuation, unnecessary characters, URLs, hashtags, mentions, emojis, and irrelevant characters are removed during the cleanup stage. In the tokenization stage, sentences from tweets are divided into word chunks. Next, the collection of sentences from the tweet is folded into lowercase as a whole at the case folding stage [24]. In the filtering or removal of stop words step, irrelevant words or stop words are removed. The stemming stage changes affixed words into basic words [25].

Libraries that focus on natural language processing in Indonesian are called literary libraries. Stemming or changing words with affixes into their basic forms or base words is one of the functions done in this library. The Sastrawi Library is very useful for analyzing texts, including sentiment analysis, text classification, and other text processing in Indonesian. Figure 4 is stage of sastrawi library:

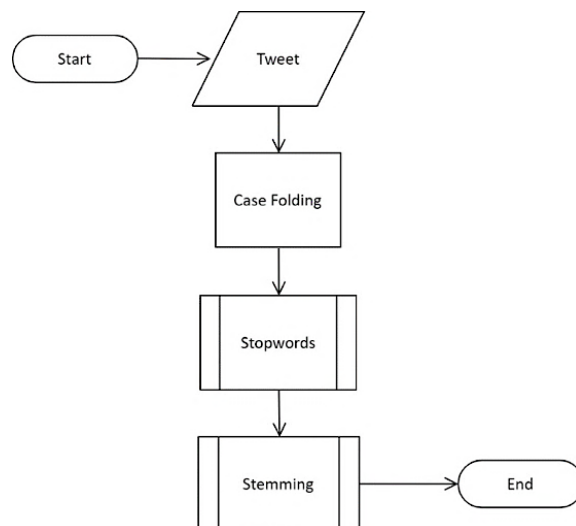


Figure 4. Preprocessing process flowchart

F. Stopword Process

The Stopwords process checks whether the words collected from case folding results with Literary library are important or meaningful words. Important words will be saved, while meaningless words will be deleted. The list of unimportant words found in Sastrawi for example 'the', 'on', 'according to', 'between', 'to', 'him', 'two', 'for', 'like', 'if', 'however', 'if'. Figure5 is a stopword process:

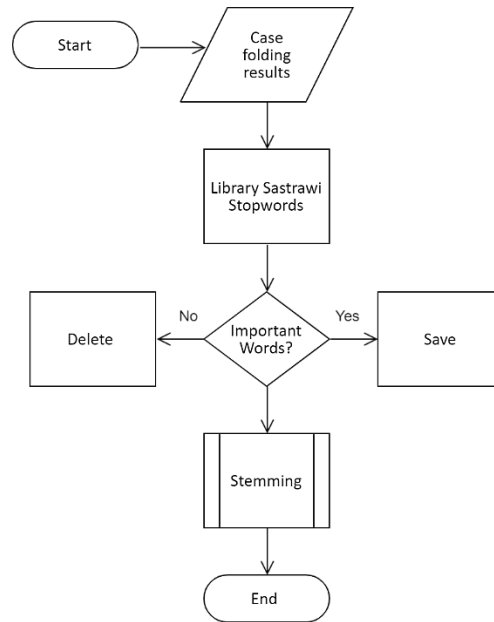


Figure 5. Process flowchart stopwords

F. Stemming Process

The stemming process aims to get the basic word of a sentence that has affixes by removing all prefixes, insertions and suffixes. In the stemming process at this research, the words will be compared with the Sastrawi Library. If the words or tweets are not basic words, the library will automatically change them to basic words. Figure 6 is the process flow of stemming.

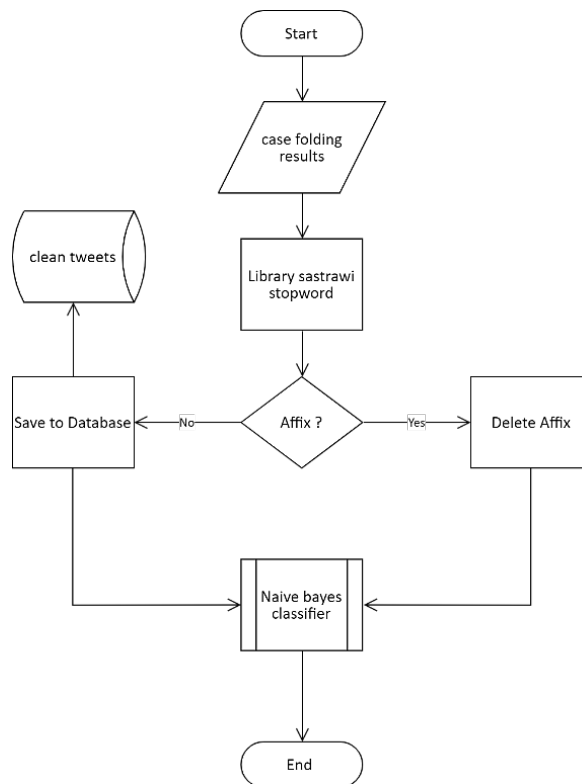


Figure 6. Stemming Process Flowchart

G. Naïve Bayes Classifier Classification Process

Classification is the process of dividing or grouping data into predetermined categories or classes based on their characteristics. The purpose of classification is to find patterns or characteristics in the data and group them into appropriate categories [26].

This research uses the Naïve Bayes classification method when using text classification technology. Naïve Bayes uses probability and statistical methods to estimate future probabilities by looking at previous experience calculations. The main characteristic of this method is the very strong (naive) assumption that the text will depend on every condition or event [27].

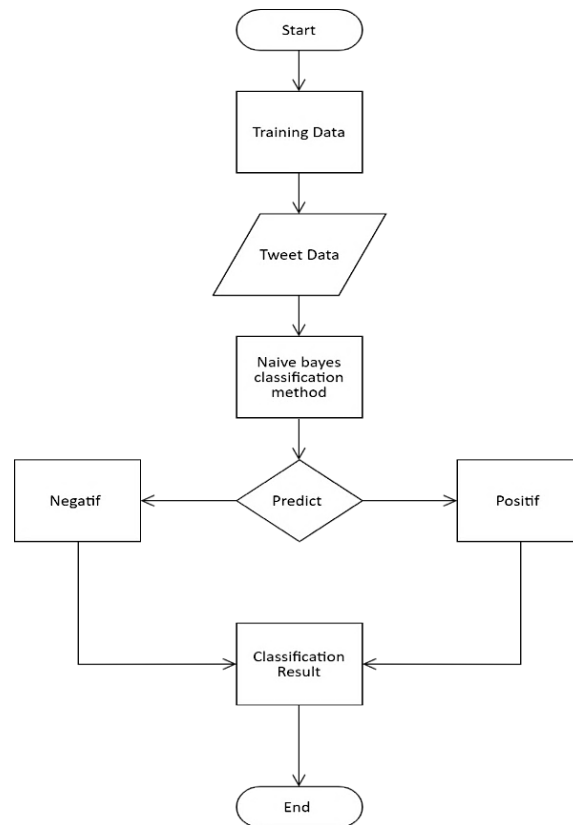


Figure 7 is the process flow of the Naïve Bayes classifier classification.

3 RESULTS AND DISCUSSION

The application of the results of this research is implemented with hardware specifications, namely using an Intel Pentium Dual core Series @ 2.0 Ghz processor, 256 GB hard disk, 2 GB RAM, Intel UHD VGA, 22inch monitor. The computer uses a Windows 10 operating system and Chrome web browser.

A. Results of preprocessing implementation

Figure 8 and Figure 9 explain the implementation results of the application and the results of the preprocessing carried out.

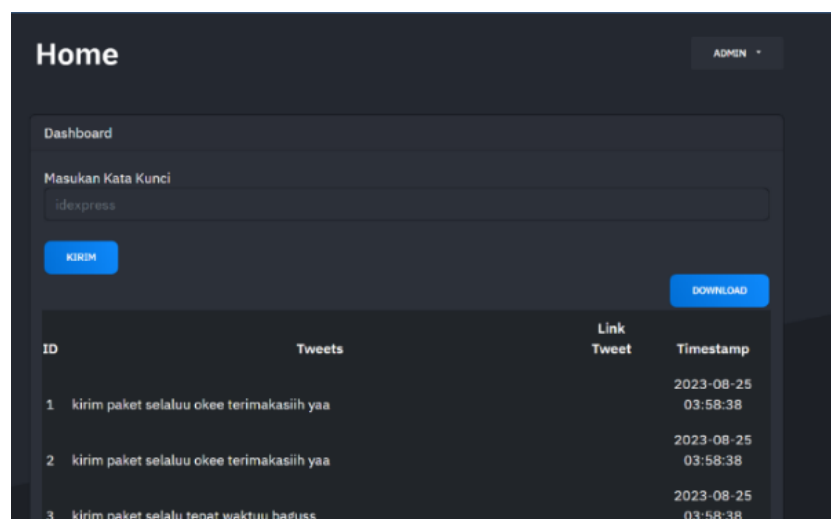


Figure 8. Implementation of user interface pages

Figure 8 is the website display for searching keywords or keywords for which data will be retrieved. The resulting system takes data in real time from Twitter (X) via the Application Programming Interface and saves the data to a database. Users can also directly download or save the tweets data.

The interface implementation display Figure 9 is a preprocessing of the keyword results that have been entered according to the service. The data taken will be cleaned from affixes, emotes and casefolding before classifying the tweet data.

Home			ADMIN
Preprocessing			
ID	Kata Kunci	Tweets	
1	idexpress pengiriman paket	irim paket selalu okee terimakasih yaa	
2	idexpress pengiriman paket	irim paket selalu okee terimakasih yaa	
3	idexpress pengiriman paket	irim paket selalu tepat waktu baguss	
4	idexpress pengiriman paket	irim paket cepatt murahh sekali	
5	idexpress pengiriman paket	irim paket sangat cepat ongkirnya selalu murah makasii	
Showing 1 to 5 of 38 results			1 2 3 4 5 6 7 8

Figure 9. Results of preprocessing implementation.

B. Twitter(X) Data Classification Results

Tweet data classification testing was carried out in August 2023 at the IDExpress Cikajang branch. Testing is carried out directly by the branch head office as shown in Figure 10.



Figure 10. User testing process

Testing is carried out by entering keywords or keywords related to existing services, namely delivery and pick up. Visual testing is also carried out by looking directly at the sentiment labels that have been predicted by the library to see whether they are appropriate for giving positive and negative labels to each tweet data that has been taken according to the keywords entered. The following is the output produced during testing which can see in Figure 11 and Figure 12.

idexpress			
ID	Kata Kunci	Tweets	Sentimen
6	idexpress pengiriman	irim cepat	Positif
7	idexpress pengiriman	reply nyata komplain semua bener jasa kirim masalah duuuh kapok boz	Negatif
8	idexpress pengiriman	min mau sratus kirim id express bandung babakan ciparay ga ubah	Negatif
9	idexpress pengiriman	yaelahnab aku cilacap jasa kirim daerah aku ga cuma sampe tegal abis ganti jadi	Positif

Figure 11. Test Results with the Keyword Pengiriman

35	idexpress pickup	gimana nih brow idexpressid barang pickup pickup kacau xmytaxut6z	Negatif
36	idexpress pickup	sbt ekspedisi mau pickup ga sama minimal paket lalu free bayar klo mau pick up	Negatif
37	idexpress pickup	kurir systemnya amburadul begini mau pakai tokopedia tokopediacare ylki id 8 hari sejak jual request pickup skrg paket buat main bola dioper2 terus qnw4ywuhiif	Negatif

Figure 12. Test Results with the Keyword Pengiriman

Based on Figure 11 and Figure 12 above, the level of accuracy of predictions will be more accurate if a lot of training data is entered. The more training data that is entered, the more accurate the prediction level of the library used will be. After labeling each tweet data, there is also an interface implementation that displays visuals and calculations for each service-related keyword or key word so that branch heads can easily determine service improvement priorities.

The following is an implementation of the data visualization interface from tweets which can see in Figure 13 and Figure 14.

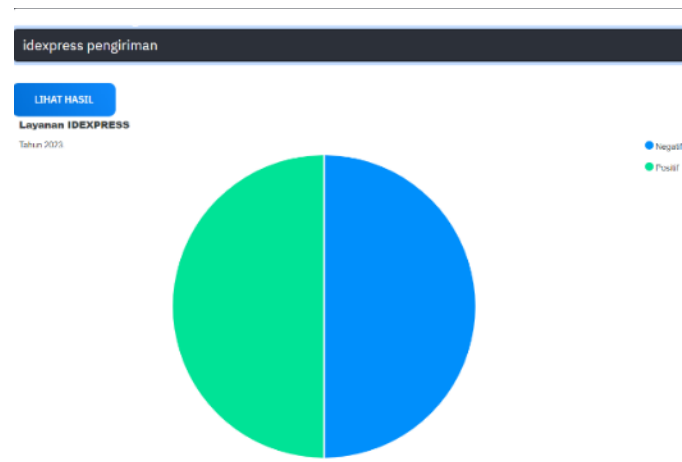


Figure 13. Visualization results of Twitter(X) data classification with "pengiriman" keyword

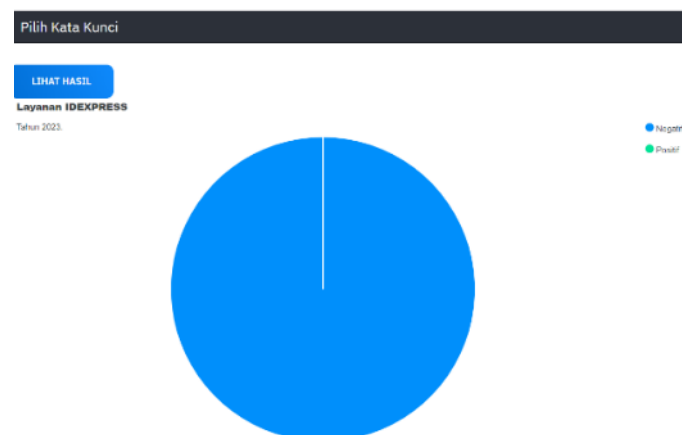


Figure 14. Visualization results of Twitter(X) data classification with pickup keyword

For the "pengiriman" service keyword, there were two positive sentiment tweets and two negative sentiment tweets, while for the "pickup" keyword there were four negative sentiment tweets.

Based on Figure 13 and Figure 14, it shows that tweets with the keyword delivery service have more positive sentiment compared to the keyword pickup service, which is where the pick-up service gets the most negative sentiment. So branch heads can easily determine priority service improvements, namely for pickup services.

C. Accuracy Test Results

Accuracy testing is carried out to find out how accurate the library used by the system is to classify tweets that have been taken from Twitter(X). In this test, the keyword "delivery" will be tested as a reference for the library's level of accuracy. The following is a training data table with the keyword "delivery" which has been labeled as training data used by the system and

library to classify tweets and direct observation data along with sentiment resulting from human observations to see the level of accuracy of predictions from the library. Table 2 is accuracy process result testing.

Table 2. Accuracy Process Result

Tweet Data Observation	Sentiment	Result
Idexpress package delivery is always okay, thank you 🍪	Positive	Accurate
Idexpress package delivery is always okay, thank you♥	Positive	Accurate
ID Express delivers packages always on time and well 🍪	Positive	Accurate
Idexpress packages delivery is very fast and cheap 🍪	Positive	Accurate
Idexpress sends packages very fast, and the postage is always cheap, thanks♥	Positive	Accurate
IDexpress Fast delivery 🍪	Positive	Accurate
Really, all the complaints are true, the delivery service has a problem, give up, boss	Negative	Accurate
@idexpress_care min want this, why doesn't the delivery status of ID Express Bandung Babakan Ciparay change?	Negative	Accurate
@yaelahnab I'm Cilacap, there's no delivery service in my area, only until Tegal and then it was changed to IDExpress	Positive	Accurate
@ @idexpress_care Please help with the receipt number IDB900071832488, the delivery status has failed for 2 days, even though the number is always active, courier Alfitho Iksyan never contacted the recipient at all regarding delivery, please help	Negative	Accurate
@idexpress_care Why is there a delivery problem? what's inside the contact lens? What happens if it breaks? CS can reply to me pls	Negative	Accurate
This one is more absurd, since when can I as a buyer ask to reschedule delivery? How do you understand it, min @idexpress_care	Negative	Accurate
@idexpress_care I want to complain about the delivery being delayed beyond the estimate. please process	Negative	Accurate
@idexpress_care min want this, why doesn't the delivery status of ID Express Bandung Babakan Ciparay change?	Positive	No
@yaelahnab I'm Cilacap, there's no delivery service in my area, only until Tegal and then it was changed to IDExpress	Positive	Accurate
@idexpress_care I want to complain about the delivery being delayed beyond the estimate.	Negative	Accurate

The following is an explanation of the percentage accuracy level predictions based on 100 training data and 16 tweet data that have been taken directly from Twitter(X).

Accuracy = (Total of correct predictions / Total Tweet data) x 100 %

The calculation is:

Total training data: 100 (50 positive, 50 negative)

Total data Tweet: 16

Total correct prediction: 15

Total incorrect predictions: 1

The prediction accuracy calculated as follows:

Accuracy percentage level = (15 / 16) x 100 = 93.75%.

Based on the calculations and data presentation above, the resulting prediction accuracy is around 93.75%. This means that from 16 tweet data, 15 of them were predicted correctly, and only 1 prediction was incorrectly predicted and it can concluded that the resulting prediction accuracy level is accurate.

As a discussion point, getting a high accuracy value will be based on the amount of training data entered into the previous system. This is because the system requires comparative data to be able to draw conclusions about positive or negative statements. The less training data, the lower the accuracy value. Thus, this system for sentiment analysis needs to be used by users who know how to input training data and who always update the training data.

CONCLUSIONS

Sentiment analysis can help users to easily obtain analysis from other people's words conveyed via social media regarding a service provided. Based on the accuracy testing, the Sastrawi and Biobii can analyze the sentiment of sentences originating from the social media Twitter(X). Further research carried out by combining various existing social media so that there will be more data as a source for sentiment analysis.

REFERENCES

- [1] M. S. Mustofa, T. Wijayanti, N. A. Putri, and D. Hermawan, "Strategies for increasing the competitiveness of micro-businesses in using e-commerce amidst the Covid-19 pandemic," *Stud. Appl. Econ.*, vol. 40, no. 2, 2022, doi: 10.25115/eea.v40i2.5541.
- [2] M. A. Camilleri, "E-commerce websites, consumer order fulfillment and after-sales service satisfaction," *J. Strategy Manag.*, vol. 15, no. 3, 2022, doi: 10.1108/JSMA-02-2021-0045.
- [3] S. G. Ajiniyazovna, "Implementation of e-commerce security methods and tools," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 5, 2020, doi: 10.30534/ijeter/2020/12852020.
- [4] M. Nashar, A. Hariadi D. P., and R. D. Parashakti, "E-commerce, economic growth and gross domestic product to the demand for delivery goods," *Int. J. Organ. Innov.*, vol. 12, no. 3, 2020.
- [5] G. Battumur, K. Gantumur, and W. Kim, "Factors affecting satisfaction with the postal delivery service," *Int. J. Adv. Appl. Sci.*, vol. 10, no. 1, 2023, doi: 10.21833/ijaas.2023.01.015.
- [6] E. Febriyani and H. Februariyanti, "Analisis sentimen terhadap program Kampus Merdeka menggunakan Naive Bayes classifier di Twitter," *J. Tekno Kompak*, vol. 17, no. 1, 2023.
- [7] M. Murni, I. Riadi, and A. Fadlil, "Analisis sentimen hatespeech pada pengguna layanan Twitter," *JURIKOM*, vol. 10, no. 2, 2023, doi: 10.30865/jurikom.v10i2.5984.
- [8] Y. A. Rohman and R. Kusumaningrum, "Twitter storytelling generator using LDA and HMM POS-tag," in *Proc. ICICOS*, 2019, doi: 10.1109/ICICoS48119.2019.8982411.
- [9] N. Garg, "Annotated corpus creation for sentiment analysis in Hinglish social network data," *Indian J. Sci. Technol.*, vol. 13, no. 40, 2020, doi: 10.17485/ijst/v13i40.1451.
- [10] N. Ilk and S. Fan, "Combining textual cues with social clues for sentiment analysis," *Decis. Sci.*, vol. 53, no. 2, 2022, doi: 10.1111/decis.12490.
- [11] S. W. Ritonga, M. Fikry, and E. P. Cynthia, "Klasifikasi sentimen masyarakat di Twitter terhadap Ganjar Pranowo," *BITS*, vol. 5, no. 1, 2023, doi: 10.47065/bits.v5i1.3535.
- [12] C. L. Rujiani, E. R. Syahputra, and S. D. Andriana, "Implementation of API using REST architecture," *Int. J. Data Sci. Vis.*, vol. 1, no. 1, 2022.
- [13] D. A. S. G. Putra Kusuma, "Designing REST architecture for school MIS," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 7, 2020, doi: 10.30534/ijeter/2020/124872020.
- [14] A. Syaifuddin and M. Muslimin, "Analisis sentimen implementasi kebijakan PSE Kominfo," *Seminar Nas. Fak. Tek.*, vol. 1, no. 1, 2022, doi: 10.36815/semastek.v1i1.2.
- [15] W. Bourequat and H. Mourad, "Sentiment analysis for iPhone release using SVM," *Int. J. Adv. Data Inf. Syst.*, vol. 2, no. 1, 2021, doi: 10.25008/ijadis.v2i1.1216.
- [16] U. Nawaz, A. Ali, and K. S. U. A. Raza, "Survey on sentiment analysis using ML techniques," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 2, 2021, doi: 10.30534/ijatcse/2021/1091022021.
- [17] H. Suroso, I. Budi, A. B. Santoso, and P. K. Putra, "Sentiment analysis on homecoming restriction policy," in *Proc. IC2IE*, 2020, doi: 10.1109/IC2IE50715.2020.9274609.
- [18] G. Sailasya and G. L. A. Kumari, "Stroke prediction using ML algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, 2021, doi: 10.14569/IJACSA.2021.0120662.
- [19] M. A. Rosid *et al.*, "Improving text preprocessing for complaint classification," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012017.
- [20] M. Syarifuddin, "Analysis of public opinion sentiment on Covid-19 using NB and KNN," *Inti Nusa Mandiri*, vol. 15, no. 1, 2020.
- [21] J. Pfeffer *et al.*, "Assessing coverage and temporal reliability of Twitter's Academic API," *Proc. ICWSM*, vol. 17, 2023, doi: 10.1609/icwsml.v17i1.22182.
- [22] M. Heak and S. H. Choi, "Dynamic crawling system for news extraction," *Int. J. Adv. Sci. Technol.*, vol. 28, no. 3, 2019.
- [23] A. A. I. A. Maharani, S. S. Prasetyowati, and Y. Sibaroni, "Classification of public sentiment on fuel price increases using CNN," *Sinkron*, vol. 8, no. 3, 2023, doi: 10.33395/sinkron.v8i3.12609.
- [24] R. Kosasih and A. Alberto, "Sentiment analysis of game products on Shopee," *ILKOM J. Ilm.*, vol. 13, no. 2, 2021, doi: 10.33096/ilkom.v13i2.721.101-109.
- [25] A. S. Rizki, A. Tjahyanto, and R. Trialih, "Comparison of stemming algorithms for Indonesian text," *TELKOMNIKA*, vol. 17, no. 1, 2019, doi: 10.12928/TELKOMNIKA.v17i1.10183.

- [26] C. F. Hasri and D. Alita, "Penerapan NBC dan SVM pada analisis sentimen dampak Covid-19," *JATIKA*, vol. 3, no. 2, 2022.
- [27] D. F. Risa, F. Pradana, and F. A. Bachtiar, "Implementasi Naive Bayes untuk mendeteksi stres siswa berdasarkan tweet," *J. Tek. Inf. Ilmu Komput.*, vol. 8, no. 6, 2021, doi: 10.25126/jtiik.2021864372.