

Implementation of User Sentiment with Naïve Bayes Algorithm to Analyze LinkedIn Application Regarding Job Vacancies in the Play Store

Dhiyauddin Al Ghozi, Firman Noor Hasan*

Faculty of Industrial Technology and Informatics, Informatics Engineering, Universitas Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia

Email: ¹ghozialko@gmail.com, ^{2,*}firman.noorhasan@uhamka.ac.id

Correspondence Author Email: firman.noorhasan@uhamka.ac.id

Abstract–Mobile applications have become an important part, one of which is the LinkedIn application which is a mobile application that focuses on the recruitment process, job search and as a professional networking platform which is now increasingly relevant, especially in Indonesia. The methodology involves data collection, data preprocessing, data labeling, and application of the Naïve Bayes algorithm. Sentiment analysis can be used as a reference to improve the quality of an application and the level of user satisfaction as well as knowing the number of positive and negative sentiments in user feedback. The 999 data obtained were then divided into 60% training data and 40% test data. In this analysis, negative sentiment outweighs positive sentiment, with a total of 539 negative reviews and 460 positive reviews. Based on evaluation using the confusion matrix, accuracy results were 95.74%, precision was 100%, and recall was 91.46%. This research aims to provide insight into the communication and interaction patterns of LinkedIn users in relation to job opportunities and overall sentiment towards the platform.

Keywords: Sentiment Analysis; Naive Bayes; LinkedIn; Job Vacancy; Play Store

1. INTRODUCTION

A In an era where everything is digital with the rapid growth of technology having an influential impact, one of which is that mobile applications have become an integral part of people's lives [1]. With mobile applications you can access everything quickly and comfortably, users can also connect with the world around them [2]. Mobile applications have made significant contributions in various contexts, including in the fields of education, health and work. A mobile application that can help recruitment and job searching is LinkedIn which provides job information [3].

The LinkedIn application is the third largest online professional networking application that is most widely used in the world [4]. Founded in 2002 by Reid Hoffman and launched in 2003, it is a professional social media platform that allows users to build closeness and professional networks [5]. It is a very important aspect in career management and recruitment to create job opportunities, internships and experiences that strengthen professional relationships. As well as pursuing the skills needed for future career success [6]. Play Store is an official Google application that runs on Android-based mobile devices [7]. According to data from Napoleon Cat, there has been an increase in LinkedIn users in Indonesia by 24.9 million users from January to June 2023. The high number of LinkedIn users makes it easier to access the information needed. However, the availability of a lot of information with uncertainty can cause information overload, where users find it difficult to sort out relevant information [8].

Based on Trading Economy data, Indonesia is one of the countries with the highest unemployment rate in Southeast Asia. In February 2023, the unemployment rate in Indonesia was recorded at 5.45%. According to the Central Statistics Agency (BPS), there are 7.86 million unemployed people in Indonesia in August 2023, a decrease of around 6.77% compared to August 2022 [9]. Even though it has decreased, if we look at the previous year, unemployment in Indonesia has still not recovered to pre-Covid 19 pandemic levels in 2019. This means that there are still many Indonesians who need jobs to support their families. The Covid-19 pandemic has had a significant impact on the global economy, including Indonesia. Many business sectors have been forced to stop operations or reduce the number of employees due to decreased demand and social restrictions implemented to control the spread of the virus. As a result, many workers have lost their jobs and are having difficulty finding new jobs amidst an uncertain economic situation.

In facing the challenge of unemployment, the Indonesian government has launched various initiatives and programs to create new jobs and improve workforce skills. One prominent program is the Pre-Employment Card, which provides training and financial incentives to job seekers to improve their skills and increase their chances of finding work. However, government efforts cannot work alone. The role of the private sector and educational institutions is also very important in creating a work ecosystem that is conducive and relevant to market needs. It is hoped that collaboration between government, industry and educational institutions can overcome the problem of unemployment effectively and sustainably. This means that there are still many Indonesian people who need jobs to support their families. With cooperation and commitment from all parties, it is hoped that the unemployment rate in Indonesia can continue to decline and the economy can fully recover from the impact of the pandemic.

Sentiment analysis is a data processing technique for collecting comment data from user reviews in text form, which is used to analyze the content of the data, in the form of negative or positive views or opinions

Volume 8, Nomor 3, Juli 2024, Page 1647-1655



ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

regarding an event. The data obtained in the form of comments or opinions is then analyzed to determine individual sentiment. The results of this analysis can be used as a benchmark for evaluating an issue or service in the future [10]. The Naïve Bayes method is used to classify user reviews into positive and negative sentiment because this method is simple, fast, and has a high level of accuracy. This method is based on Bayes' Theorem, which allows calculating the probability of a hypothesis based on new information. The main advantage of this method is its ability to handle large datasets with high efficiency, which is especially useful in sentiment analysis where the number of user reviews can be very large [11].

Previous research has been carried out regarding the application of the Naïve Bayes algorithm in analyzing sentiment. Research conducted by Bobby Kurniadi Widodo, Nur Hafifah Matondang, and Desta Sandya Prasvita regarding the application of the Naïve Bayes algorithm for using the jobstreet application using 1000 datasets resulted in an accuracy value of 96%, precision 98% and recall 94% [12]. In research conducted by Artanti Inez Tanggraeni, and Melkior N Sitokdana regarding reviews of E-Government applications with 674 datasets, it resulted in 89% accuracy, 83% precision, and 87% recall [13]. The Naïve Bayes method was also used in research conducted by Rizky Kurnia Pratama and Putry Wahyu Setyaningsih, who analyzed comments related to employment using 3929 tweet data. The results of this research show an accuracy of 97%, precision of 92%, and recall of 99%, confirming the Naïve Bayes method as a very effective tool for classifying sentiment in user reviews [6]. Other research conducted by Oktavia Putri Zusrotun, Alif Catur Murti, and Rina Fiati conducted sentiment analysis about online learning using the Naïve Bayes algorithm with an accuracy of 74.08%. In the model testing process, the size of the test data is taken as 30% of the training data which is carried out randomly. The model evaluation carried out in this study used 15 fold cross validation with accuracy results of 76.39% [14].

Based on previous research, this research will investigate public opinion on LinkedIn regarding job vacancies and user satisfaction levels by applying the Naïve Bayes algorithm. This research aims to provide an insight into the communication and interaction patterns of LinkedIn users in relation to job and career opportunities and overall sentiment towards the platform. In addition, to determine whether there is positive sentiment or negative sentiment in user reviews of the LinkedIn application.

2. RESEARCH METHODOLOGY

The research methodology section will be explained systematically as the stages carried out to solve the research problem. Using the Naïve Bayes model, data analysis was carried out as a tool for analyzing sentiment [15]. Figure 1 research flow, below shows the research methodology flow.

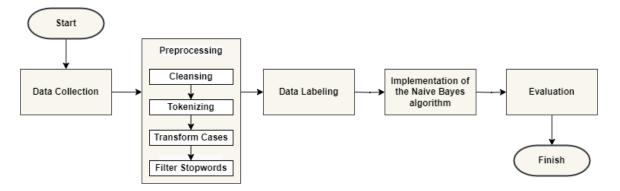


Figure 1. Research Flow

The initial stage of the research methodology starts from collecting data using web scrapping techniques, where the data taken is comments from users of the LinkedIn application on the Play Store with the help of Google Colab [16]. The data obtained from the scrapping results in CSV form is continued to the preprocessing stage for processing. The preprocessing stage is the data cleaning stage which includes cleansing, tokenizing, transform cases, and stopwords filter which aims to make it easier for researchers to carry out labeling. After preprocessing, we then carry out labeling. There are two parts of data in this research, namely training data and test data [17]. Manual data labeling carried out on training data is used to train the model before the model carries out classification. Meanwhile, test data is used in the Naïve Bayes implementation process. The stages of the process of implementing Naïve Bayes involve forecasting data and determining whether the data should be classified as positive or negative sentiment based on its similarity to past experiences [18]. The final stage of this research is evaluation, which is to measure the accuracy of the results of the process that has been carried out. By using the confusion matrix equation you can indicate the accuracy, precision and recall values. Below is the calculation formula used in the confusion matrix.

Accuracy =
$$\frac{TP+TN}{TP+TN+FP+FN}$$

JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 8, Nomor 3, Juli 2024, Page 1647-1655 ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879



(2)

(3)

 $Precision = \frac{TP}{TP+FP}$ $Recall = \frac{TP}{TP+FN}$

3. RESULT AND DISCUSSION

3.1 Data Collection

Data collection was the first stage in this research which was taken using web scrapping techniques using the Python programming language. This data was obtained on November 21, 2023. Web scrapping is a method used to collect information in the form of review data with the aim of extracting information [19]. As shown in Figure 2, the data scrapping process is a data retrieval process which begins by accessing the Play Store, then going to the LinkedIn application to get the link which will later be entered into Google Colab. After that, the data resulting from the scrapping process is saved in CSV file format and downloaded to proceed to the preprocessing stage.



Figure 2. Data Scrapping Process

Researchers carried out the process of scraping data from user reviews of the LinkedIn application. Figure 3 Scrapping Data LinkedIn Application Review, below shows data scraping carried out using Google Colab, the first stage is entering 2 libraries, namely numpy and pandas. Then enter the destination link for the LinkedIn application, namely 'com.linkedin.android' and use the latest review category [20]. The data taken was 1000 review data containing user name labels, ratings, dates and reviews. Then the scrapping data is saved in CSV format. After the scrapping process is complete, the CSV file will be downloaded to proceed to the preprocessing stage.

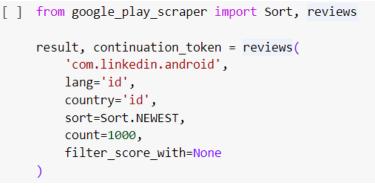


Figure 3. LinkedIn Application Review Data Scrapping

3.2 Data Preprocessing

The preprocessing process aims to clean the data so that the data analysis steps can run smoothly. The tool used in the preprocessing process is RapidMiner. At this stage, several stages in preprocessing are carried out, including:

a. The first stage is cleansing. The purpose of this operator is to sterilize documents from components that have nothing to do with document data [21]. In figure 4 of the cleansing stage, the following shows the cleansing process where the data is tidied up by using the read CSV operator to enter the scrapping data and then connecting it with the select attributes operator to select the part to be cleaned. Then it is connected with the replace operator to delete parts of the data such as punctuation, characters, hashtags, mathematical symbols, emoticons, and unnecessary symbols ([.?,:!#\$+TM;*^(a嚾🇰@ '¤]) will be omitted.

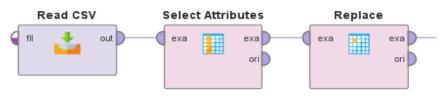


Figure 4. Cleansing Stage

Volume 8, Nomor 3, Juli 2024, Page 1647-1655



ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

Table 1, cleansing results, is the result of the cleansing stage where the data is clean of characters, symbols and punctuation in the review data.

Table	1.	Clea	nsing	Results
-------	----	------	-------	---------

Before Cleansing	After Cleansing
Good for job seekers, continue	Good for job seekers continue
Very helpful in developing my business.	Very helpful in developing my business
Apk@ Really Helps the Young Generation Who	Apk Really Helps the Young Generation Who Don't Have
Don't Have a Job	a Job
The application is quite good and really helps me as a company owner. However, it is a shame that you cannot use digital payments such as GoPay,	The application is quite good and really helps me as a company owner However it is a shame that you cannot use digital payments such as GoPay Shopee Pay Dana etc
Shopee Pay Dana, etc. to pay for advertising or paid posts.	to pay for advertising or paid posts
Login image verification error occurs several times, please fix it	Login image verification error occurs several times please fix it
Sorry in advance to LinkedIn, why now if you	Sorry in advance to LinkedIn why now if you add an
add an attachment it often doesn't work. Is this a	attachment it often doesn't work Is this a bug from
bug from LinkedIn please fix it Thank you	LinkedIn please fix it Thank you

The preprocessing stages shown in Figure 5 are an important foundation in text data processing because each step plays a role in converting raw data into data that is ready to be analyzed. By going through the process of cleansing, tokenizing, transform cases, and stopwords filters, text data becomes more structured and meaningful, allowing for more accurate and effective analysis in subsequent stages. This process is a crucial step before the data can be used for advanced analyzes such as sentiment analysis, text classification, or other natural language processing tasks.



Figure 5. Operators in Preprocessing

b. The second stage is tokenizing. The aim of this operator is to separate or break down words per word in a sentence into individuals, as in table 2 of tokenizing results [22].

Table 2. Tokenizing results

Before Tokenizing	After Tokenizing	
Good for job seekers continue	Good, for, job, seekers, continue	
Very helpful in developing my business	Very, helpful, in, developing, my, business	
Apk Really Helps the Young Generation Who Don't	Apk, Really, Helps, the Young, Generation, Who, Don't,	
Have a Job	Have, a Job	
The application is quite good and really helps me as	The application, is quite, good, and, really, helps, me, as,	
a company owner However it is a shame that you	a company, owner, However, it, is a shame, that, you,	
cannot use digital payments such as GoPay Shopee	cannot, use, digital, payments, such, as GoPay, Shopee,	
Pay Dana etc to pay for advertising or paid posts	Pay, Dana, etc, to, pay, for, advertising, or, paid, posts	
Login image verification error occurs several times	Login, image, verification, error, occurs, several, times,	
please fix it	please, fix, it	
Sorry in advance to LinkedIn why now if you add an	Sorry, in, advance, to, LinkedIn, why, now, if, you, add,	
attachment it often doesn't work Is this a bug from	an attachment, it often, doesn't, work, Is this, a bug,	
LinkedIn please fix it Thank you	from, LinkedIn, please, fix it, Thank you	

c. The third stage is transform cases. The purpose of this operator is to change every word containing uppercase letters in the dataset to all lowercase, using the lower case alternative so that all data is uniformly small in size, as in table 3 of the transform cases results [17].

Table 3. Results	of Transform Cases
------------------	--------------------

Before Transform Cases	After Transform Cases
Good, for, job, seekers, continue	good, for, job, seekers, continue
Very, helpful, in, developing, my, business	very, helpful, in, developing, my, business

Volume 8, Nomor 3, Juli 2024, Page 1647-1655



ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

Before Transform Cases	After Transform Cases	
Apk, Really, Helps, the Young, Generation, Who,	apk, really, helps, the young, generation, who, don't,	
Don't, Have, a Job	have, a job	
The application, is quite, good, and, really, helps, me,	the application, is quite, good, and, really, helps, me,	
as, a company, owner, However, it, is a shame, that,	as, a company, owner, however, it, is a shame, that,	
you, cannot, use, digital, payments, such, as GoPay,	you, cannot, use, digital, payments, such, as gopay,	
Shopee, Pay, Dana, etc, to, pay, for, advertising, or,	shopee, pay, dana, etc, to, pay, for, advertising, or,	
paid, posts	paid, posts	
Login, image, verification, error, occurs, several,	login, image, verification, error, occurs, several,	
times, please, fix, it	times, please, fix, it	
Sorry, in, advance, to, LinkedIn, why, now, if, you,	sorry, in, advance, to, linkedIn, why, now, if, you, add,	
add, an attachment, it often, doesn't, work, Is this, a	an attachment, it often, doesn't, work, Is this, a bug,	
bug, from, LinkedIn, please, fix it, Thank you	from, linkedIn, please, fix it, thank you	

d. The fourth stage is the stopwords filter. The purpose of this operator is to remove irrelevant or unimportant words that are referenced from the Indonesian stopwords dictionary, such as personal pronouns and conjunctions, which have no sentiment value, as in table 4 of the stopwords filter results [23].

 Table 4. Stopwords Filter Results

Before Stopwords Filter	After Stopwords Filter
good, for, job, seekers, continue	good, job, seekers, continue
very, helpful, in, developing, my, business	very, helpful, in, developing, business
apk, really, helps, the young, generation, who, don't, have, a job	apk, really, helps, the young, generation, have, a job
the application, is quite, good, and, really, helps, me, as, a company, owner, however, it, is a shame, that, you, cannot, use, digital, payments, such, as gopay, shopee, pay, dana, etc, to, pay, for, advertising, or, paid, posts	the application, is quite, good, really, helps, a company, owner, however, it, is a shame, that, cannot, use, digital, payments, such, as gopay, shopee, pay, dana, etc, pay, advertising, or, paid, posts
login, image, verification, error, occurs, several, times, please, fix, it	login, image, verification, error, occurs, please, fix, it
sorry, in, advance, to, linkedIn, why, now, if, you, add, an attachment, it often, doesn't, work, Is this, a bug, from, linkedIn, please, fix it, thank you	sorry, in, advance, linkedIn, now, if, add, an attachment, it often, doesn't, work, Is this, a bug, linkedIn, please, fix it, thank you

3.3 Data Labeling

After completing the Preprocessing process, the next stage is the labeling process on the dataset. Manual labeling is done by dividing the data into training data and test data. From the scrapping results, 1000 data were obtained, then the number of datasets changed to 999 review data after going through the Preprocessing stage. Researchers divided the data in a ratio of 60:40 for training data and test data [24]. A total of 599 data were used for training data and test data obtained, 599 user comments were manually labeled positive or negative. The remaining 400 user comments were not labeled for testing in Naïve Bayes classification. The purpose of this labeling is so that each dataset has its own positive and negative sentiments to be used as a reference, where positive labels indicate comments containing praise or user satisfaction with the application [25]. Meanwhile, negative labels are comments containing complaints or user dissatisfaction with the application, as in table 5 of sentiment labeling results.

Table 5.	Sentiment Labeling Results
----------	----------------------------

Sentiment	Review
Positive	It is very helpful for finding work and being able to share it with others in the network
Negative	Very difficult to enter login Always returns to the main menu
Positive	Very helpful in professional networking
Negative	It's still difficult to log in and ask me to choose an image that isn't clear
Positive	I joined this application but I think this application is very good and easy to use for looking for relationships and business opportunities as well as for advertising promotions too
Negative	The clearest application. I wanted to register for the application but instead I was told to verify the puzzle several times for up to a year ~? Please, if you want to make an application, try it first so that it will be easier for users who want to use it
Positive	I hope this application is more flexible in searching for job vacancies, especially more relevant keywords and relevance of search results that are more in line with the skills mastered. I hope there is a job vacancy filter feature that does not require mastering English

Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib

Volume 8, Nomor 3, Juli 2024, Page 1647-1655 ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)



DOI: 10.30865/mib.v8i3.7879

3.4 Algorithm Implementation

After going through the preprocessing and labeling stages, the next stage is algorithm implementation, using the Naïve Bayes method to develop an algorithm model based on a previously designed model. By using the Naïve Bayes method, you can estimate the probability of similarity between previous experiences and future probabilities [26]. The use of the Naïve Bayes method has shown significant accuracy in handling large amounts of data. Figure 6 training data creation mechanism, showing the training data creation process. Starting by using the read CSV operator to read the CSV file that has been prepared, then connected with the example filter operator. This operator is set to the option (is not missing) to filter data that is not missing so that later sentiments that are still empty or have not been labeled in the training data will be retrieved and processed [27]. Next, to process the data you need a nominal to text operator so that the document process operator can read and process the data. The process document operator contains the tokenize operator, transform cases, and stopwords filter. Data analysis is then carried out using the Naïve Bayes operator, and the data is stored in stores called the model store and training data store.



Figure 6. Training Data Creation Mechanism

After creating the training data, the next steps for implementing Naïve Bayes are shown in Figure 7 the mechanism for implementing Naïve Bayes. The stages of implementing Naïve Bayes begin with the CSV read operator and then connected with the examples filter operator to eliminate unnamed or missing data. The nominal to text operator and the process document operator have the same explanation as in Figure 6. Then the results of the data analysis in Figure 6, which is called the training data store, are combined together with the process document operator. Next, it is connected with the replace mising values operator, while the data obtained from the analysis which is called the store model is connected with the apply model operator together with the data from the previous analysis or what is called the training data store.



Figure 7. Naïve Bayes Implementation Mechanism

After going through the stages of preprocessing, labeling, and implementing Naïve Bayes, prediction results are obtained. In figure 8 the pie diagram, displays a visualization of the results of implementing Naïve Bayes in the form of a pie diagram. From a total of 999 data processed, 460 comments were categorized as positive sentiment and 539 comments were categorized as negative sentiment. The prediction results are based on the research conducted, it can be concluded that user opinions regarding the LinkedIn application are generally negative. This is caused by many users who regret the repeated login verification on the LinkedIn application.

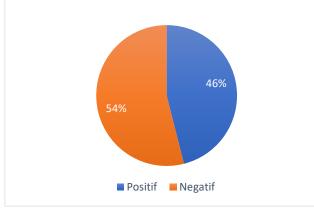


Figure 8. Pie chart



Volume 8, Nomor 3, Juli 2024, Page 1647-1655 ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

Not only that, the results of implementing Naïve Bayes can also be visualized in wordcloud form to find out the words that appear most often and those that are most often used in user reviews. There were 999 review data used in this research, but the researchers only showed 20 words with the highest frequency. The most frequent occurrence is the word 'application' with a total occurrence of 407 words. Due to the many user reviews that regret the repeated login verification on the LinkedIn application. Figure 9 the results of the wordcloud visualization, below shows a visualization in wordcloud form where the greater the frequency of words appearing, the larger the size of the words in the display format.

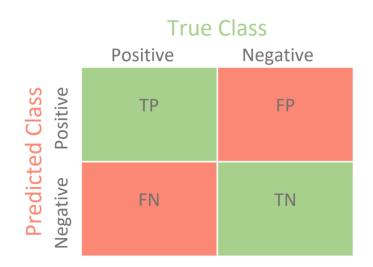
ExampleSet

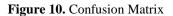


Figure 9. Wordcloud visualization results

3.5 Evaluation

The final stage in this research process is evaluation. Evaluation is an important step that aims to measure and understand the performance of the model that has been built [28]. The main purpose of the evaluation is to find out the extent to which the model can provide accurate and relevant results according to the data analyzed. By conducting evaluations, researchers can identify the strengths and weaknesses of the model, and determine whether the model is suitable for use in real applications or requires further improvement [29]. This research uses the confusion matrix equation to determine the results of the model evaluation. Confusion matrix is a very useful tool in classification model performance analysis. The confusion matrix displays the number of correct and incorrect predictions made by the model compared to the actual values in the dataset. Figure 10 shows the confusion matrix, which is used in this research to calculate the levels of accuracy, precision and recall. This evaluation provides a comprehensive picture of the model performance, helping in understanding whether the Naïve Bayes model has classified the data correctly or whether there are many prediction errors.





The meaning of the terms used in the confusion matrix is True Positive (TP) if the actual and predicted class are positive. False Positive (FP) if the actual class is negative and the predicted class is positive. False Negative (FN) if the actual class is positive and the predicted class is negative. And True Negative (TN) if the

Volume 8, Nomor 3, Juli 2024, Page 1647-1655

ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online) Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

actual and predicted class are negative. In Figure 11 the confusion matrix mechanism, shows the various operators used in the confusion matrix calculation process in Rapidminer.

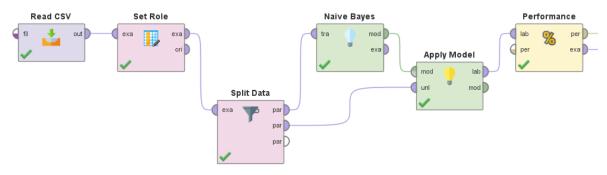


Figure 11. Confusion Matrix Mechanism

The confusion matrix mechanism begins with the CSV read operator to enter a file containing prediction data, then the set role operator is connected to read the prediction (sentiment) column as a label. The split data operator is used to split the data with a ratio of 60% to 40% and then connected to the Naïve Bayes operator and the apply model operator. Next, the apply model operator is connected to the performance operator to find out the accuracy results. The results of the confusion matrix process stage produce an accuracy value of 95.74% as shown in Figure 12 of the confusion matrix results.

accuracy: 95.74%

	true Positif	true Negatif	class precision
pred. Positif	182	0	100.00%
pred. Negatif	17	200	92.17%
class recall	91.46%	100.00%	

Figure 12. Confusion Matrix results

By using the confusion matrix equation, the results of calculating accuracy, precision and recall from sentiment analysis obtained the following results.

Accuracy =
$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{182 + 200}{182 + 200 + 0 + 17} = \frac{382}{399} = 0.9573 = 95.74\%$$

Accuracy calculations produce an accuracy of 95.74%, meaning that the model is able to classify the data as a whole at a very good level.

$$Precision = \frac{TP}{TP + FP} = \frac{182}{182 + 0} = \frac{182}{182} = 1 = 100\%$$

The precision calculation produces a precision of 100%, which means it shows the model is able to identify prediction results correctly.

Recall =
$$\frac{TP}{TP + FN} = \frac{182}{182 + 17} = \frac{182}{199} = 0.9145 = 91.46\%$$

The recall calculation produces a recall of 91.46%, which indicates that the model was successful in finding most of the true positive instances. With high recall, the model tends to be successful in identifying the majority of positive cases.

4. CONCLUSION

After conducting research using data from the Play Store using the Naïve Bayes method, we obtained 999 data to be researched by dividing the data, including 599 data used for training data and 400 data used for test data. Of the 999 data processed, 460 reviews were categorized as positive sentiment and 539 reviews were categorized as negative sentiment. Conclusions can be drawn in this research by using the Naïve Bayes method and confusion matrix at the evaluation stage to obtain quite high accuracy values, namely reaching 95.74%, precision of 100%, and recall of 91.46%. The prediction results are based on research conducted that the Naïve Bayes and confusion matrix methods in classifying data have high accuracy. It can be concluded that user sentiment about the LinkedIn application is generally negative. This research shows that the sentiment analysis of LinkedIn application users regarding job vacancies tends to be poor, where negative labels are reviews that contain user dissatisfaction. This is because many users regret the existence of repeated login verification on the LinkedIn application. Suggestions for future researchers are the need to compare several classification methods to find out which method is the best.

Volume 8, Nomor 3, Juli 2024, Page 1647-1655 ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)

Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib DOI: 10.30865/mib.v8i3.7879

REFERENCES

- A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," J. Inform. dan Tek. Elektro Terap., vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.
- [2] K. D. Indarwati and H. Februariyanti, "Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier," JATISI (Jurnal Tek. Inform. dan Sist. Informasi), vol. 10, no. 1, 2023, doi: 10.35957/jatisi.v10i1.2643.
- [3] N. Benanda et al., "Analisis Pengaruh Platform LinkedIn Terhadap Fenomena Fear of Missing Out (FOMO) di Kalangan Mahasiswa ITS," J. Sist. Inf. dan Ilmu Komput., vol. 1, no. 4, 2023, doi: https://doi.org/10.59581/jusiik-widyakarya.v1i4.1965.
- [4] U. Kulsum, M. Jajuli, and N. Sulistiyowati, "Analisis Sentimen Aplikasi LinkedIn di Google Play Store Menggunakan Algoritma Support Vector Machine," J. Appl. Informatics Comput., vol. 6, no. 2, pp. 205–212, 2022, doi: 10.30871/jaic.v6i2.4802.
- [5] S. Indriyana, T. N. Muthmainah, and A. A. Ridha, "Analisis User Experience Pada Platform LinkedIn Menggunakan Metode User Experience Questionnaire," J. Manaj. Inform. Komputerisasi Akunt., vol. 7, no. 2, pp. 208–216, 2023, doi: https://doi.org/10.46880/jmika.Vol7No2.pp208-216.
- [6] R. K. Pratama and P. W. Setyaningsih, "Analisis Komentar Pada Twitter Terhadap Lapangan Kerja Dengan Metode Naïve Bayes," J. Inf. Syst. Artif. Intell., vol. 3, no. 2, pp. 217–227, 2023, doi: https://doi.org/10.26486/jisai.v3i2.129.
- [7] Y. W. S. Putra et al., Pengantar Aplikasi Mobile. Haura Utama, 2023.
- [8] R. Saputri and R. H. Utami, "Pengaruh Information Overload Terhadap Prokrastinasi Akademik Pada Mahasiswa," Nusant. J. Ilmu Pengetah. Sos., vol. 10, no. 2, pp. 710–720, 2023, doi: 10.31604/jips.v10i11.2023. 4928-4933.
- [9] F. A. Tanjung, A. P. Windarto, and M. Fauzan, "Penerapan Metode K-Means Pada Pengelompokkan Pengangguran Di Indonesia," Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform., vol. 6, no. 1, p. 61, 2021, doi: 10.30645/jurasik.v6i1.271.
- [10] D. A. Nugroho and F. N. Hasan, "Analisis Sentimen Kegiatan Pembersihan Sampah Pada Media Sosial X Menggunakan SVM dan Naïve Bayes," J. Media Inform. Budidarma, vol. 8, no. April, pp. 843–853, 2024, doi: 10.30865/mib.v8i2.7562.
- [11] M. K. Anam, B. N. Pikir, and M. B. Firdaus, "Penerapan Na ive Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen danPemeritah," MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput., vol. 21, no. 1, pp. 139–150, 2021, doi: 10.30812/matrik.v21i1.1092.
- [12]K. Widodo, Bobby, H. Matondang, Nur, S. Prasvita, and Desta, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Penggunaan Aplikasi Jobstreet Implementation of Naive Bayes Algorithm For Sentiment Analysis Of The Use Of Jobstreet Application," 2022. doi: https://doi.org/10.33633/tc.v21i3.6361.
- [13] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes," JATISI (Jurnal Tek. Inform. dan Sist. Informasi), vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [14] O. P. Zusrotun, A. C. Murti, and R. Fiati, "Analisis Sentimen Terhadap Belajar Online pada Media Sosial Twitter Menggunakan Algoritma Naive Bayes," J. Nas. Pendidik. Tek. Inform., vol. 11, no. 3, pp. 310–319, 2022, doi: 10.23887/janapati.v11i3.49160.
- [15] H. Z. Muflih, A. R. Abdillah, and F. N. Hasan, "Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes," KLIK Kaji. Ilm. Inform. dan Komput., vol. 4, no. 3, pp. 1613–1621, 2023, doi: 10.30865/klik.v4i3.1303.
- [16] R. I. A. P. H. Dr. Uky Yudatama, S.Si., M. Kom., M.M., Nur Syamsiyah, Irmawati, Ade Davy Wiranata, Rahmi Imanda S.Kom., M.Kom, Hadiansyah Ma'sum, Aris Wahyu Murdiyanto, Yahya, R. Arri Widyanto, Dhita Diana Dewi, Samuel Aleksander Mandowen, Firman Noor Hasan, Halomoan Ed, Memahami Teknologi Informasi: Prinsip, Pengembangan, dan Penerapan. Kaizen Media Publishing, 2023.
- [17] Meliyawati and F. N. Hasan, "Analisis Sentimen Pengguna Aplikasi CapCut Pada Ulasan di Play Store Menggunakan Metode Naïve Bayes," Media Online), vol. 4, no. 4, pp. 2272–2280, 2024, doi: 10.30865/klik.v4i4.1555.
- [18] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier," J. Minfo Polgan, vol. 11, no. 2, pp. 30–35, 2022, doi: 10.33395/jmp.v11i2.11640.
- [19] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine," TEKNOSAINS J. Sains, Teknol. dan Inform., vol. 10, no. 2, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.
- [20] E. Yuniar, D. S. Utsalinah, and D. Wahyuningsih, "Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompet Digital dengan Menggunakan Algoritma Machine Learning," J. Janitra Inform. dan Sist. Inf., vol. 2, no. 1, pp. 35–42, 2022, doi: 10.25008/janitra.v2i1.145.
- [21]M. K. Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di," J. Mhs. Tek. Inform., vol. 7, no. 1, pp. 478–483, 2023, doi: https://doi.org/10.36040/jati.v7i1.6373.
- [22] D. P. Santoso and W. Wibowo, "Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Metode Naïve Bayes Classifier pada Situs Google Play Store," J. Sains dan Seni ITS, vol. 11, no. 2, 2022, doi: 10.12962/j23373520.v11i2.72534.
- [23] A. Erfina and M. F. Al-shufi, "Analisis Sentimen Aplikasi Jasa Kurir Di Play Store Menggunakan Algoritma Naive Bayes," J. Sist. Inf. dan Inform., vol. 5, no. 2, pp. 103–110, 2022, doi: 10.47080/simika.v5i2.1789.
- [24] A. T. Mukti and F. N. Hasan, "Analisis Sentimen Warganet Terhadap Keberadaan Juru Parkir Liar Menggunakan Metode Naive Bayes Classifier," J. Media Inform. Budidarma, vol. 8, no. 1, pp. 644–653, 2024, doi: 10.30865/mib.v8i1.6982.
- [25]G. Darmawan, S. Alam, and M. I. Sulistyo, "Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi Mypertamina Pada Google Playstore Menggunakan Metode Naïve Bayes," STORAGE – J. Ilm. Tek. dan Ilmu Komput., vol. 2, no. 3, pp. 100–108, 2023, doi: https://doi.org/10.55123/storage.v2i3.2333.
- [26] R. Yunita and M. Kamayani, "Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Penghapusan Kewajiban Skripsi," Indones. J. Comput. Sci., vol. 12, no. 5, pp. 2879–2890, 2023, doi: 10.33022/ijcs.v12i5.3415.
- [27]S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," STORAGE J. Ilm. Tek. dan Ilmu Komput., vol. 2, no. 3, pp. 109–117, 2023, doi: 10.55123/storage.v2i3.2334.
- [28] Alfandi Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," J. Sist. Inf., vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.
- [29] N. Herlinawati, Y. Yuliani, S. Faizah, Windu Gata, and Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," CESS (Journal Comput. Eng. Syst. Sci., vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186.