Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sentiment Analysis of Beauty Product Applications using the Naïve Bayes Method

Tiara Syavitri Rambe[1)*], Mila Nirmala Sari Hasibuan[2)], Muhammad Halmi Dar[3)]
[1,2,3)]Universitas Labuhanbatu, Indonesia
[1)] tiarasyavitri.rambe@gmail.com, [2)]milanirmalasari7@gmail.com, [3)]mhd.halmidar@gmail.com

Submitted : Mar 31, 2023 | **Accepted** : Apr 8, 2023 | **Published** : Apr 9, 2023

Abstract: The number of beauty products that appear on the market makes every producer compete in attracting consumers. One of the facilities provided by manufacturers to make it easier for consumers to shop is an online shopping application that can be accessed via gadgets. Where the feature of the application is the availability of user review services User reviews are often used as a recommendation for the product to be purchased. The more positive the reviews that appear, the greater the consumer's confidence to buy the product; conversely, the more negative the reviews that appear, the more reluctant consumers are to buy. This study aims to find out how much accuracy the Naïve Bayes algorithm has in conducting sentiment analysis on user reviews of beauty product applications with different combinations of training and test data. Furthermore, it is also important to know the frequency of words that often appear in the review. The sentiment class used is divided into three, namely, positive, negative, and neutral. This research method includes a number of stages, namely: data collection, data labeling, text pre-processing, data visualization, TF-IDF, sentiment analysis, etc., until the results are obtained. This research has produced the highest accuracy rate of 90.08% in the Naïve Bayes algorithm, with a composition of 90% training data and 10% test data. While the word that often appears in user reviews is "application," with a frequency of 446 occurrences, it is followed by the word "product," 444 times, and the word "price," 312 times. The greater the amount of training data used, the higher the level of accuracy resulting from the Naïve Bayes algorithm. Meanwhile, the greater the amount of test data used, the lower the resulting accuracy value.

Keywords: Beauty; Naïve Bayes; Review; Sentiment Analysis; Wordcloud.

INTRODUCTION

Today's beauty products have become a necessity for many people, especially women. Beauty products rank third most in demand by consumers when shopping online, with a percentage of 39% (Annur, 2022). Many beauty entrepreneurs review beauty products through social media and are able to attract the interest of many people to buy them. However, people are too lazy to come to the store in person, especially if the distance is far and the product prices are expensive. Regarding this problem, many beauty product stores make it easy for their buyers to shop online through gadgets without having to come directly to the store. One store application that sells various beauty products online is the Raena beauty product application. The Raena application is the largest reseller and dropship application for beauty products in Indonesia, with more than 100,000 users and selling more than 2,000 beauty products (Indonesia, n.d.).

Raena, the largest beauty product application in Indonesia, provides users with an application service review feature. These reviews can be used as recommendations by consumers and are useful for

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

companies to improve their quality. The review feature is very useful for viewing feedback from users (Rhohmawati, Slamet, & Pratiwi, 2019). Application reviews and ratings are usually represented by an asterisk, which aims to measure the quality of the application. To carry out the process of identifying a very large number of user review data, it is very inefficient if done manually (Wiratama & Rusli, 2019), because it will take a very long time. One effort that can be made is to do sentiment analysis. Sentiment analysis is able to identify reviews from users in depth by visualizing the value of reviews in positive, negative, or neutral classes (Kabiru & Sari, 2019).

Sentiment analysis of application reviews has been carried out by a number of researchers using various classification algorithm methods. By implementing the Naïve Bayes Classifier algorithm, a sentiment analysis was carried out on reviews of Shopee application users who bought smartphone products with an accuracy rate of 85% (Sihombing, Hannie, & Dermawan, 2021). In another study, sentiment analysis was also carried out, but by applying the K-Nearest Neighbor algorithm to instant hijab products, the resulting accuracy rate was 76.92% (Muktafin, Kusrini, & Luthfi, 2020). By combining the K-means algorithm and the Naïve Bayes Classifier, the results of the sentiment analysis show that the K-means algorithm is not considered optimal (Hariguna, Baihaqi, & Nurwanti, 2019). For this reason, selecting the right algorithm for the appropriate problem is very important when conducting sentiment analysis.

This study applies the Naïve Bayes algorithm. There are several advantages to using the Naïve Bayes algorithm, including that it is very efficient to implement and has fairly good accuracy and performance in classifying data (Astuti & Astuti, 2022). When compared to the Support Vector Machine algorithm, the Naïve Bayes method is the best method for classifying sentiments on user reviews (Negara, Muhardi, & Putri, 2020). The Naïve Bayes algorithm is also very often used in conducting sentiment analysis to review a product with high accuracy results (Azhar, Adikara, & Adinugroho, 2021).

In this study, sentiment analysis was carried out on user reviews of the Raena beauty product application using the Naïve Bayes algorithm. The class of emotions is divided into three types, namely: positive, negative, and neutral. Word weighting was carried out using the Term Frequency-Inverse Document Frequency (TF-IDF) method. In addition, the composition of the training data and test data is also different by applying nine combinations, ranging from 10% to 90% of the training data and test data. This combination is needed to see the effect of the amount of training and test data on the accuracy of the resulting Naïve Bayes algorithm. The formulation of the problem in this study is to see how much the highest accuracy value is produced by the Naïve Bayes algorithm in the combination of the percentage of training data and test data, then what words appear frequently in user reviews of the application.

LITERATURE REVIEW

Sentiment analysis has the same meaning as opinion mining, opinion extraction, sentiment mining, and review mining (Rolliawati, Khalid, & Rozas, 2020). The sentiment analysis algorithm can classify text into positive, neutral, or negative opinion classes automatically (Saputra, Nurhadryani, Wijaya, & Defina, 2021). The purpose of using sentiment analysis is to map user opinions based on predetermined topics (Negara et al., 2020). Sentiment analysis is an attempt to classify the polarity of sentiment associated with a series of aspects (Ma, Peng, & Cambria, 2018). The purpose of sentiment analysis is to get more accurate and detailed information from user reviews from various perspectives and aspects, such as game design, plot, graphics, sound effects, consumer experience, and technical level (Jiaxin, 2020).

Naïve Bayes is a classification algorithm that applies the probability and statistical methods discovered by Thomas Bayes by predicting future opportunities based on previous history (Afdhaluzzikri, Mawengkang, & Sitompul, 2022). The Naïve Bayes method combines previous experience with new knowledge with the basic concept of combining word and category probabilities to predict document category probabilities (Azhar et al., 2021). In carrying out the classification, the Naïve Bayes method has two stages, namely, the training stage to gain past experience and the classification stage by calculating the probability values of all the data (Negara et al., 2020).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Term Frequency and Inverse Document Frequency (TF-IDF) is a method of weighting words in documents to find keywords in the available categories by extracting features from the text (Normawati & Prayogi, 2021). The purpose of the word weighting process is to increase the machine's ability to carry out sentiment analysis (Que, Iriani, & Purnomo, 2020). Term Frequency has the role of finding the value of the word occurrence in one document, while Inverse Document Frequency has the function of finding the value of the word occurrence in all documents (Astuti & Astuti, 2022). The TF-IDF method shows the score from the frequency of occurrence of words for interesting words; this process is done by calculating the weight of each word in the training data using the sklearn library in Python with two schemes, namely, CountVectorizer for word counts and TfidfVectorizer for word frequencies (Negara et al., 2020).

METHOD

This research requires a research methodology that includes structured and systematic steps related to the activities and procedures carried out during the research (Afdhaluzzikri et al., 2022). The research stages describe the flow of research carried out from the beginning until the research ends. For this reason, the stages of this research use the framework shown in Figure 1.

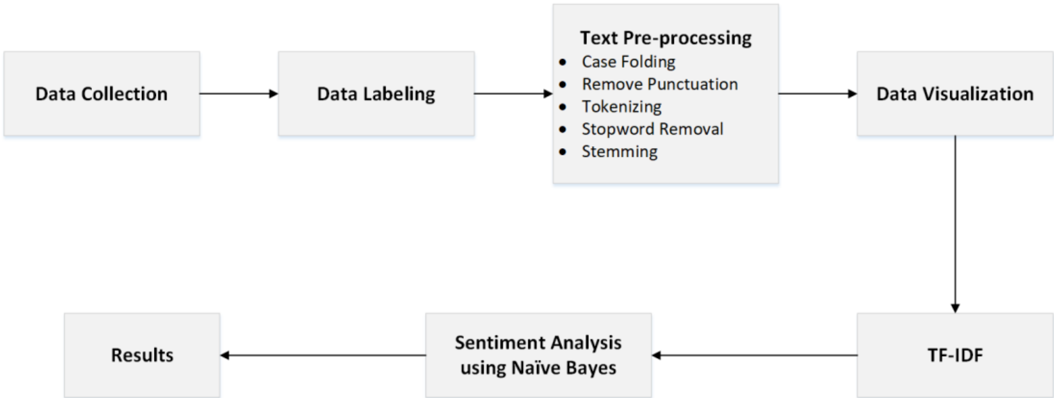

Fig 1. Research Framework

Data Collection

The data used in this study is user review data for the Raena beauty product application obtained from Google Play. The data collection process was carried out using a scraping technique in Python from reviews of application users who speak Indonesian and are located in Indonesia.

Data Labeling

Data labeling is the process of assigning sentiment values to data based on sentiment classes, which are divided into three types: positive, negative, and neutral. Whereas a score of 1-2 is expressed as a negative sentiment with a code of -1, a score of 3 is a neutral sentiment with a code of 0, and a score of 4-5 is a positive sentiment with a code of 1. At this stage, the distribution of sentiment will also be shown by year and number of reviews.

Text Pre-processing

At this stage, pre-processing of the text is carried out to clean the data of things that are not needed. At this stage, five processes are carried out, namely: case folding, removing punctuation, tokenizing, stopword removal, and stemming. In the case-folding process, all text in the document will be converted to lowercase. Taking out punctuation will clean the text of things that are not needed, such as punctuation, hashtags, emoticons, and so on. Tokenizing will cut each word into chunks of word units. Stopword removal will remove words that have nothing to do with the topic or sentiment value. Meanwhile, stemming is needed to remove affixes, both at the beginning and at the end of a word, so that each word returns to its base word. In the stemming process, the Python literature library is used. In this study, the resulting stemming will bring up user reviews that only consist of more than three words in one sentence.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Visualization

At this stage, the words in the user review will appear visually using a Wordcloud. The visual display in wordcloud form is divided into three categories based on sentiment classes: positive, negative, and neutral sentiment wordclouds. The size of the letters that appear on the Wordcloud is an interpretation of the frequency of occurrence of the word. The bigger the word size, the more often the word appears in the review, and vice versa, if these words rarely appear in the review, the smaller the word size will be.

TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) will convert text to a vector. TF will show the number of occurrences of the term in a document. Meanwhile, IDF will show the number of documents containing the term. This method is used to determine how relevant a word is in a data document that has been pre-processed.

Sentiment Analysis

At this stage of sentiment analysis, three processes will be carried out, namely: distribution of training data and test data; modeling using the Naïve Bayes algorithm; and accuracy testing based on a combination of sharing training data and test data. The distribution of training data and test data used in this study uses the nine combinations shown in Table 1 below.

Table 1. Combination of Training Data and Test Data

Combination	Training Data	Test Data
1	90%	10%
2	80%	20%
3	70%	30%
4	60%	40%
5	50%	50%
6	40%	60%
7	30%	70%
8	20%	80%
9	10%	90%

RESULT

Overall, this study uses the Python programming language on the online Google Collaboratory, starting from the process of scraping Raena application user data to visualizing results. The use of Google Colab with Python is intended to make it easier for researchers to carry out work on implementing the Naïve Bayes algorithm, because apart from this application being free to use, the library in Python is also very easy to install in this application. The rest is some data visualization using the help of the Microsoft Excel application. The results of this study will systematically describe the research stages described in the previous Methods section.

content	score	Year	Month	Day
Saya baru pakai aplikasi ini. Saat saya coba m...	4	2020	9	2
Kalo misalnya bisa dicek mana ready stock dan ...	4	2020	9	6
Aplikasi susah mengirim kode OTP.. Mungkin bis...	3	2020	9	6
Susah masuk aplikasi gara2 kode otp tidak muncul2	2	2020	9	14
Nggak ngerti gmn, sy masukkan nama toko Ol say...	1	2020	9	18
...	...	...	...	...
Aplikasinya bagus, dan skrg smakin berkembang....	5	2023	2	23
Untuk merek ngak lengkap ngakk ada marcks dan ...	2	2023	3	6
Apk nya lemot banget, padahal membantu bgt bua...	1	2023	3	10
Trjadi lg alotnya koordinasi dgn cs. Barang pe...	1	2023	3	17
Kenapa sekarang aplikasi lemot banget padahal ...	1	2023	3	23

Fig 2. Scraping Datasets

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data that was successfully obtained through scraping techniques on Google Play is shown in Figure 2. The data shown is a scraping result dataset that has been framed into five columns: content column, score, year, month, and day. The data obtained is 1919 rows of reviews from 2020 to 2023 after sorting the most relevant data. Content is a review of each user comment, while the score is a rating represented by an asterisk from 1 to 5. The number 1 on the score is a very dissatisfied opinion, and it continues to increase up to 5, which represents a very satisfied opinion.

Of the 1919 review lines generated through the scraping process, data recapitulation is carried out based on the score. Important recapitulation is done to see the number of user reviews based on the level of score given. The results of data recapitulation based on scores are visualized as shown in Figure 3. A score of 5, or the highest rating, ranks highest in user reviews with a total of 1184 reviews, while a score of 1, or the lowest rating, ranks second with a total of 329 reviews. Reviews with a rating of 2 are the reviews with the smallest number, namely 121 reviews.

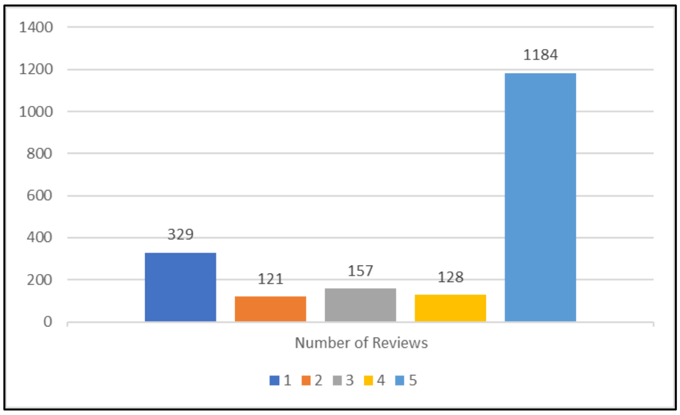

Fig 3. Scraping Data recapitulation

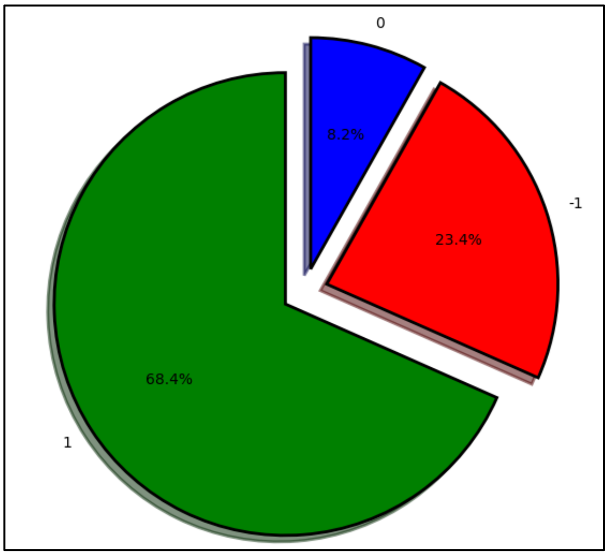

Fig 4. Distribution of Sentiment

Figure 4 shows the sentiment distribution resulting from labeling the scraped dataset. In the percentage-based distribution of sentiment, it can be seen that positive sentiment is represented by number 1 with a percentage of 68.4% and 1312 reviews, negative sentiment is represented by number -1 with 23.4% and 450 reviews, and neutral sentiment is represented by number 0 with a percentage of 8.2% and a total of 157 reviews. It can be seen that positive sentiment predominates in user reviews of the Raena application.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 2. Sentiment Distribution by Year

Year	Sentiment	Number of Reviews
2020	Negative	61
	Neutral	19
	Positive	46
2021	Negative	236
	Neutral	105
	Positive	616
2022	Negative	141
	Neutral	32
	Positive	638
2023	Negative	12
	Neutral	1
	Positive	12

Table 2 is the number of user reviews based on sentiment class per year. Table 2 shows the trend of increasing positive sentiment from 2020 to 2022. Reviews with the highest positive sentiment values are in 2022, with a total of 638. Reviews with the highest negative sentiment values are in 2021 with a total score of 236. Meanwhile, reviews with the most neutral sentiment will be available in 2021 with a total score of 105. In 2023, there will be a decrease because the review is only conducted for 3 months (January to March), in contrast to previous years, which amounted to 12 months.

Table 3. Pre-processing results

Pre-processing stage	Results
Dataset	Saya baru pakai aplikasi ini. Saat saya coba mendownload foto produk terdapat keterangan berhasil, namun sudah saya cari dan tidak tersimpan di hp saya. Apakah hanya saya yg begitu?
Case Folding	saya baru pakai aplikasi ini. saat saya coba mendownload foto produk terdapat keterangan berhasil, namun sudah saya cari dan tidak tersimpan di hp saya. apakah hanya saya yg begitu?
Remove Punctuation	saya baru pakai aplikasi ini saat saya coba mendownload foto produk terdapat keterangan berhasil namun sudah saya cari dan tidak tersimpan di hp saya apakah hanya saya yg begitu
Tokenizing	['saya', 'baru', 'pakai', 'aplikasi', 'ini', 'saat', 'saya', 'coba', 'mendownload', 'foto', 'produk', 'terdapat', 'keterangan', 'berhasil', 'namun', 'sudah', 'saya', 'cari', 'dan', 'tidak', 'tersimpan', 'di', 'hp', 'saya', 'apakah', 'hanya', 'saya', 'yg', 'begitu']
Stopword Removal	['pakai', 'aplikasi', 'coba', 'mendownload', 'foto', 'produk', 'keterangan', 'berhasil', 'cari', 'tersimpan', 'hp', 'yg']
Stemming	['pakai', 'aplikasi', 'coba', 'mendownload', 'foto', 'produk', 'terang', 'hasil', 'cari', 'simpan', 'hp', 'yg']

Table 3 shows an example of a text pre-processing result data which is considered representative because it appears on the top line. In the case of the dataset folding phase, all text has been converted to lowercase. The removing punctuation phase has succeeded in removing all unnecessary punctuation and characters. In the tokenizing phase, words have been formed into word units. The stopword removal phase has succeeded in eliminating words that are irrelevant to the topic of sentiment. In the stemming phase, each word has been returned to its base word.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Positive *Negative* *Neutral*

Fig 5. Wordcloud Sentiment

To bring up a visual image of the frequency of word occurrences, the visualization of text into images is performed using Wordcloud. Figure 5 shows the wordcloud of each positive, negative, and neutral sentiment. Each word that appears on the Wordcloud is a representation of the frequency of occurrence of the word in each sentiment. The more often the word appears, the larger the word size appears on the Wordcloud.

The results of the accuracy of sentiment analysis using the Naïve Bayes algorithm with different combinations of training data and test data are shown in Table 4.

Table 4. Naive Bayes Test Results

Combination	**Training Data**	**Test Data**	**Naïve Bayes**
1	90%	10%	90.08 %
2	80%	20%	89.06 %
3	70%	30%	88.30 %
4	60%	40%	85.93 %
5	50%	50%	83.76 %
6	40%	60%	84.22 %
7	30%	70%	82.25 %
8	20%	80%	80.08 %
9	10%	90%	72.33 %

Based on the data, it appears that the highest accuracy value in Naïve Bayes is generated from training data with a percentage of 90% and test data of 10% with an accuracy of 90.08% While the lowest accuracy value of 72.33% is obtained from the results of a combination of training data of 10% and test data of 90%, The difference between the highest accuracy value and the lowest accuracy value is 17.75%. Next, visualize the results of the accuracy data using the line plot and bar plot shown in Figure 6.

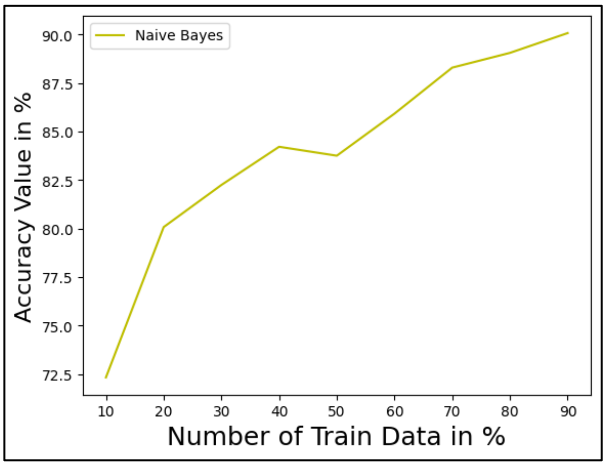

Fig 6. Sentiment Analysis Accuracy

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

From Figure 6, it can be seen that the trend of increasing accuracy has continued up to the 90% phase of the training data. However, there was also a 50% decrease in the condition of the training data, although there was again a 60% increase in the percentage of the training data.

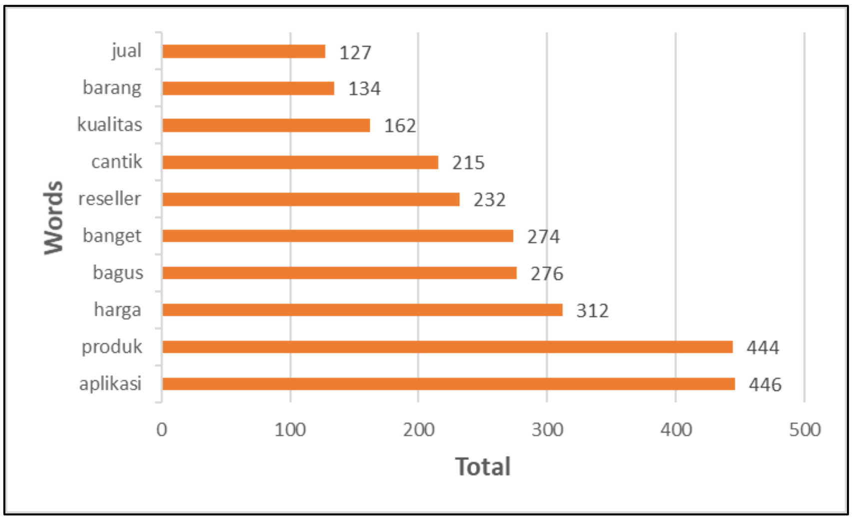

Fig 7. Frequency Distributions

Figure 7 shows the frequency distribution of word occurrences in Raena's application review. The figure shows the top ten words that appear most frequently in reviews. The word that appears the most for each sentiment class is the word application," with a total of 446 occurrences. This was followed by the word “product” 444 times and the word “price” appearing 312 times.

DISCUSSIONS

This research has produced a sentiment analysis model by applying the Naïve Bayes algorithm to nine different training data combinations with different accuracy levels. The influence of the amount of training data greatly affects the level of accuracy produced by the Naïve Bayes algorithm. The results of the study show that the greater the percentage of the use of training data, the greater the level of accuracy produced; conversely, the smaller the percentage of the use of training data, the lower the accuracy. However, there is an anomaly that arises when the training data is at 40%. There is an increase in accuracy of 0.46% from the previous position of the training data and test data, both at 50%.

The word "application" is the one with the greatest frequency, followed by the words "product" and "price". These results are very logical, as we can see that these words are very commonly used in user application review columns. However, it is necessary to investigate further whether the appearance of these words is more in the positive, negative, or neutral sentiment class.

The process of scraping Raena application user data that is carried out on the Google Play page uses a count variable of 50000 with the aim of getting 50,000 rows of data, but the scraping results obtained are only 1919 data reviews. Does the number of reviews from users of this application not reach 50,000, or is it true that the number of user reviews in Indonesian and located in Indonesia is only 1919 data reviews? For this reason, it is necessary to carry out further research with various "lang", "country", and "count" variables.

CONCLUSION

This research has been successfully carried out with the highest accuracy rate of the Naïve Bayes algorithm, namely 90.08% in the condition that the training data is 90% and the test data is 10%. While the lowest accuracy is at 72.33% with a composition of 10% training data and 10% test data, It can be concluded that the greater the amount of training data used, the higher the level of accuracy resulting from the Naïve Bayes algorithm. Meanwhile, the greater the amount of test data used, the lower the resulting accuracy value. The highest frequency of occurrence of the word in Raena's beauty application

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

user reviews is the word "application" with 446 occurrences of the word, followed by the word "product" 444 times, and the word "price" 312 times. Hopefully, with sentiment analysis of user reviews of the Raena beauty product application, it will be taken into consideration for consumers to buy and will become suggestions and input for the company to improve the quality of its services.

REFERENCES

Afdhaluzzikri, A., Mawengkang, H., & Sitompul, O. S. (2022). Perfomance of Naive Bayes method with data weighting. *SinkrOn*, *7*(3), 817–821. https://doi.org/10.33395/sinkron.v7i3.11516

Annur, C. M. (2022). 8 Produk yang Paling Diminati Konsumen Saat Belanja Online, Apa Saja? *Katadata*, 1. Retrieved from https://databoks.katadata.co.id/datapublish/2022/03/28/8-produk-yang-paling-diminati-konsumen-saat-belanja-online-apa-saja

Astuti, T., & Astuti, Y. (2022). Analisis Sentimen Review Produk Skincare Dengan Naïve Bayes Classifier Berbasis Particle Swarm Optimization (PSO). *Jurnal Media Informatika Budidarma*, *6*(4), 1806. https://doi.org/10.30865/mib.v6i4.4119

Azhar, N., Adikara, P. P., & Adinugroho, S. (2021). Analisis Sentimen Review Kedai Kopi Menggunakan Metode Naive Bayes dengan Seleksi Fitur Algoritme Genetika. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(3), 609. https://doi.org/10.25126/jtiik.2021834436

Hariguna, T., Baihaqi, W. M., & Nurwanti, A. (2019). Sentiment Analysis of Product Reviews as A Customer Recommendation Using the Naive Bayes Classifier Algorithm. *International Journal of Informatics and Information Systems*, *2*(2), 48–55. https://doi.org/10.47738/ijiis.v2i2.13

Indonesia, R. (n.d.). RAENA Reseller & Dropship App. Retrieved from Google Play website: https://play.google.com/store/apps/details?id=com.raenaapp&hl=id&gl=US

Jiaxin, S. (2020). Aspect-Based Sentiment Analysis on Mobile Game Reviews Using Deep Learning. *Graduate School of Computer and Information Sciences, Hosei University*, *15*, 1–6. https://doi.org/10.15002/00022720

Kabiru, I. N., & Sari, P. K. (2019). Analisa Konten Media Sosial E-Commerce Pada Instagram Menggunakan Metode Sentimen Analysis Dan LDA-Based Topic Modeling (Studi Kasus : Shopee Indonesia). *E-Proceeding of Management*, *6*(1), 12–19.

Ma, Y., Peng, H., & Cambria, E. (2018). Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM. *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, 5876–5883. https://doi.org/10.1609/aaai.v32i1.12048

Muktafin, E. H., Kusrini, K., & Luthfi, E. T. (2020). Analisis Sentimen pada Review Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing. *Jurnal Eksplora Informatika*, *10*(1), 32–42. https://doi.org/10.30864/eksplora.v10i1.390

Negara, A. B. P., Muhardi, H., & Putri, I. M. (2020). Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(3), 599. https://doi.org/10.25126/jtiik.2020711947

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *5*(2), 697–711. Retrieved from http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369

Que, V. K. S., Iriani, A., & Purnomo, H. D. (2020). Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, *9*(2), 162–170. https://doi.org/10.22146/jnteti.v9i2.102

Rhohmawati, U., Slamet, I., & Pratiwi, H. (2019). Sentiment Analysis Using Maximum Entropy on Application Reviews (Study Case: Shopee on Google Play). *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, *5*(1), 44–49. https://doi.org/10.26555/jiteki.v5i1.13087

Rolliawati, D., Khalid, K., & Rozas, I. S. (2020). Teknologi Opinion Mining untuk Mendukung Strategic Planning. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(2), 293. https://doi.org/10.25126/jtiik.2020721685

Saputra, F. T., Nurhadryani, Y., Wijaya, S. H., & Defina, D. (2021). Analisis Sentimen Bahasa Indonesia pada Twitter Menggunakan Struktur Tree Berbasis Leksikon. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(1), 135. https://doi.org/10.25126/jtiik.0814133

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12303
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sihombing, L. O., Hannie, H., & Dermawan, B. A. (2021). Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algortima Naïve Bayes Classifier. *Edumatic: Jurnal Pendidikan Informatika*, *5*(2), 233–242. https://doi.org/10.29408/edumatic.v5i2.4089

Wiratama, G. P., & Rusli, A. (2019). Sentiment Analysis of Application User Feedback in Bahasa Indonesia Using Multinomial Naive Bayes. *2019 5th International Conference on New Media Studies (CONMEDIA)*, 223–227. https://doi.org/10.1109/CONMEDIA46929.2019.8981850

*name of corresponding author