JAIIT (Journal of Advances in Information and Industrial Technology)
Vol. 6, No. 1 May 2024, ISSN 2716-1935, E-ISSN 2716-1927

Application of Naive Bayes Algorithm for Sentiment Analysis of Service and Facility Satisfaction (Case Study: PKU Muhammadiyah Sukoharjo General Hospital)

Hasda Surya Putra[*1)], Vihi Atina[2)], Dwi Hartanti[3)]
[1,2,3)] Department of Informatics, University of Duta Bangsa Surakarta, Surakarta, Indonesia
[1*]202020238@mhs.udb.ac.id, [2]vihi_atina@udb.ac.id, [3]dwihartanti@udb.ac.id

ARTICLEINFO

Article history:
Received 20 June 2024
Revised 8 July 2024
Accepted 12 July 2024
Available online 16 July 2024

Keywords:
Hospital
Naive Bayes
Sentiment Analysis
Services And Facilities
Visitor Reviews

ABSTRACT

PKU Muhammadiyah Sukoharjo Hospital is a health facility located in Sukoharjo, Central Java. The role of the hospital in the Sukoharjo community is very important in terms of the level of community satisfaction with its services and facilities. Opinions and assessments from the community also affect the assessment of the quality of service provided by the hospital. A large assessment can be seen on Google Maps review, which will take time for the community. Public complaints on Google Maps review of PKU Muhammadiyah Sukoharjo Hospital became the basic object of this research. This research uses the Naive Bayes algorithm to categorize the assessment based on positive and negative sentiments. Two hundred nine review data from Google Maps were processed with text processing and classified using Naïve Bayes. Test results using 10% test data show a dominant positive sentiment, with 90% accuracy, 94% precision, and 83% recall.

1. Introduction

Google Maps is an official platform owned by Google that is used on Android or Web systems. Android users can see reviews of visitors who have visited a place, whether an object is good or not, through reviews on Google Maps. Users provide various reviews on Google Maps, which can be seen in the review column [1]. Google Reviews from Google Maps, which makes it easier for businesses, institutions, and visitors to see the assessment of an assessment in the form of text and scores to measure visitor satisfaction [2]. Community ratings on Google Maps Reviews can be analyzed based on data collected through web scraping. In the scraping process, data crawling will occur as much as the data needed. The database that has been collected facilitates the process of analyzing an object. [3].

The services provided by the hospital mainly have a composition of two views, namely positive and negative [4]. This research was conducted by looking at visitor complaints at the PKU Muhammadiyah Sukoharjo General Hospital, which became the core of the object of this research. It took visitor review data from Google Maps' review of PKU Muhammadiyah Sukoharjo hospitals. Besides being easy to get review data, the location of the hospital is close to the author.

There are various algorithms that can be used in data classification and sentiment analysis, such as the Naive Bayes algorithm, Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbor (KNN). Based on several reference articles that researchers have read, the Naïve Bayes algorithm has fairly high accuracy results and, in terms of time, is fairly fast in conducting sentiment analysis compared to other classification algorithms. [5]. In previous research from [6] with the research title Aspect-based Sentiment Analysis of Yogyakarta Tentrem Hotel Reviews using the Random Forest Classifier Algorithm based on the results of his research obtained the results of 90% accuracy value and 90% fl score, this value uses the best tree number and tree depth parameters that have an important influence on the prediction results.

In the research [7] with the research title Family Contribution in Predicting Students Graduating on Time Using the Support Vector Machine Model, the accuracy value is 86%, the f1 score is 93%, the precision score is 86%, and the recall score is 99%. Research from [8] with the research title Twitter User Sentiment Analysis on PPKM Extension Using the K-Nearest Neighbor Method, the accuracy score is 69.5%, recall is 69.5%, and precision is 68.7%. In the research [9] with the research title Application of Data Mining to Predict the Number of Best-Selling Products Using the Naive Bayes Algorithm Case Study (Toko Prapti), the accuracy value is 83.3%, precision is 84.2%, and the recall is 88.9%.

Although there are some weaknesses in the Naive Bayes algorithm, such as being vulnerable to outlier data and probability estimation is difficult, based on previous studies mentioned above, the Naïve Bayes algorithm has the advantage that the amount of training data is relatively small to predict the parameters needed in the classification process. Comparison of the Naive Bayes algorithm with other alternative algorithms such as decision trees, namely decision trees divide data based on rules that are easy to understand but tend to overfit on training data and are less efficient in classifying complex text data, comparison with SVM itself lies in text processing time, SVM requires more complicated parameter tuning which is time-consuming in processing text data, for comparison with random forests itself is almost the same as SVM, namely in higher computation time than Naive Bayes but random forests tend to give good results, and for comparison with Gardien Boosting lies in more complicated parameter tuning and more computational resources. If high model complexity and accuracy are required, alternatives such as SVM, Random Forests, or gradient boosting can be explored. The Naïve Bayes algorithm has also proven to be quite fast and accurate when applied to large datasets [10]. This research was conducted by looking at visitor feedback and complaints, which became the core of the object of research by taking data from a Google Maps review of PKU Muhammadiyah Sukoharjo General Hospital. The research explains the processes that produce sentiment analysis. The Naïve Bayes algorithm was chosen and used because it has been widely used for text mining processes because it has a high accuracy value and also has a relatively fast classification time in sentiment analysis in previous studies [11].

Based on the above background, the researcher will use the Naive Bayes algorithm in the research of visitor sentiment analysis on services and facilities at PKU Muhammadiyah Sukoharjo General Hospital. It is hoped that the results of this study can help the hospital monitor and manage the services and facilities of PKU Muhammadiyah Sukoharjo General Hospital better.

2. Methods

In this research, there are several processes or stages that must be carried out to get the best results [12]. The process consists of collecting data through scraping and then preprocessing the text, which consists of case folding, tokenizing, filtering, and stemming. After that, the data will be categorized using the Naive Bayes algorithm. The data evaluation stage involves the use of a Confusion Matrix, and for visualization, WordCloud is used, which can be seen in Figure 1 below.

Figure 1. Research Flow

2.1. Data collection/Data scraping

At this stage, the dataset is manually clustered using scraping techniques from the website and a Chrome extension called Instant Data Scraper. The data source comes from the Google Reviews page about the Sukoharjo General Hospital. The data that is successfully obtained is then stored in .csv format to facilitate the next stage.

 DOI: https://doi.org/10.52435/jaiit.v6i1.566

2.2. Prepocesing Text

The next stage is text preprocessing; not all record data and attributes can be used, so it is necessary to preprocess text or efforts to process initial data in order to obtain better data [13]. This text preprocessing stage consists of 4 stages, namely case folding, tokenizing, filtering, and stemming, which can be seen in Figure 2 below [5].

Figure 2: Text preprocessing stage

a. Case Folding

This case folding stage is the process of leveling letters that previously contained capital letters to be converted into all lowercase letters. This helps reduce variation and ensure consistency in text analysis. For example, "Data" and "data" are considered the same after case folding. The impact of case folding is that it can facilitate word comparison and reduce ambiguity.

b. Tokenizing

The next stage is the tokenizing stage, which is the process of cutting or separating sentences into words per word. The impact of tokenizing itself is to facilitate text analysis, such as calculating word frequencies or building models.

c. Filtering

The next stage after tokenizing is the filtering stage, which eliminates or removes words that are not needed in the next process, such as "and", "or", and "in". The impact of filtering can reduce the dimensionality of data and increase the efficiency of analysis.

d. Stemming

The next stage is the stemming stage, which converts words that have affixes into basic words. The impact of stemming is to speed up search and ensure consistency in text analysis.

2.3. Naive Bayes

The stage after preprocessing the text is the calcification stage of the model using the Naive Bayes algorithm. Naïve Bayes is a simple classification method that can calculate all possibilities by combining the number of combinations and density of values from a dataset that has been obtained [14]. The advantage of using this method is that it requires little training data to determine the parameter estimates needed in the classification process because it is assumed to be an independent variable [15].

2.4. Confusion Matrix

The next stage is the evaluation stage using the confusion matrix, which is an evaluation method used to measure the performance of classification methods. During the efficiency measurement, there are four terms that represent the results of the classification process. The four terms include True Positive (TF), True Negative (TN), False Positive (FP), and False Negative (FN) [16]. After obtaining the values of the four terms, the values for accuracy, precision, and recall can be obtained in the following formula:

$$accuracy = \frac{TP+TN}{\text{Total}} \quad (1)$$

Where TP = data results with positive predicted values and positive actual values and TN = data results with negative predicted values and negative actual values.

$$precision = \frac{TP}{\text{TP+FP}} \quad (2)$$

Where TP = data results with positive predicted values and positive actual values and FP = data results with positive predicted values and negative actual values.

$$recall = \frac{TP}{\text{TP+FN}} \quad (3)$$

where TP = data results with positive predicted values and positive actual values and FN = data results with negative predicted values and positive actual values [17].

2.5. WordCloud

The next stage is the evaluation process, namely the visualization stage using wordcloud. This stage is characterized by a wordcloud that summarizes the overall sentiment of positive and negative reviews described in the positive and negative review documents of PKU Muhammadiyah Sukoharjo General Hospital on Google Maps.

3. Results and Discussions

3.1. Data Scraping

Data retrieval or data scraping using the Chrome extension by installing an instant data scraper to get Google Maps review data for PKU Muhammadiyah Sukoharjo Hospital. The data obtained in the form of review ID, account name, review, and time for the sentiment itself is done manually using Microsoft Excel as much as 314 data, and after cleaning, it becomes 209 data, which is divided from 154 positive data and 55 negative data. Examples of data that has been scraped and given sentiment in Table 1 below:

Table 1. Scraped review data

No	Review Id	Account Name	Time	Reviews	Sentiment
1	Br-101	Cinta Sejati	4 days ago	Of the several hospitals I visited in Sukoharjo. Here the hospital service is the longest, less satset. It's a pity about the management.	Negative
2	Br-111	Nurafanti Afanti	1 month ago	Slow	Negative
3	Br-120	Rizky Ahmad	2 month ago	Pharmacy service is very poor	Negative
4	Br-147	Diah Retno Hermastuti	4 month ago	good service and adequate facilities	Positive
5	Br-181	Sony Cahtono	4 month ago	Good service	Positive
6	Br-190	Safira Dila	4 month ago	Fast service	Positive
7	Br-216	Eva Yohana	7 month ago	Limited space for queuing in the clinic. Not worth it for visitors	Negative
8	Br-283	Freddick Listern	1 year ago	Slow administration	Negative
9	Br-317	Muhammad Faeyza	2 year ago	Good service satisfying	Positive
10	Br-323	Rayi Rayendaria	2 year ago	A comfortable hospital in my opinion. Friendly staff	Positive
11	Br-325	Alena Valency	2 year ago	Vaccines at PKU Muhammadiyah Sukoharjo can comfortably get a gift	Positive
12	Br-328	Awan Dwi Kurniawan	2 year ago	Friendly hospital staff	Positive
13	Br-329	Ryan Slametraharjo	2 year ago	Good hospital, very good service	Positive
14	Br-333	Suratmi	2 year ago	The service was satisfactory and pleasant	Positive
15	Br-346	Anang Sang Ship's Cook	2 year ago	Good hospital services	Positive
16	Br-352	Ririn Murtiah	2 year ago	Good service	Positive
17	Br-360	Agus Triutomo	2 year ago	The service was very satisfying. Thank you	Positive
18	Br-396	Rafidha Adelina Siregar	3 year ago	Satisfactory service	Positive
19	Br-400	Yussuf Faisal	3 year ago	Very good service	Positive
20	Br-402	Gilang Erlanda	3 year ago	The nurses are not friendly and the service is not satisfactory.	Negative

DOI: https://doi.org/10.52435/jaiit.v6i1.566

3.2. Preprocesing Text Result

After data scraping, the next stage is text preprocessing; text preprocessing itself is a stage where the data will be equalized in all its forms and formats so that at the next stage, the data can be processed and processed. The text preprocessing stage consists of 4 stages, namely case folding, tokenizing, filtering, and stemming. Table 2 shows the data that has not been processed case folding; table 3 shows the data after case folding and figure 1 shows the source code of the case folding process.

a. Case Folding

The case folding stage is the process of converting all characters in a document into uniform letters, be it uppercase or lowercase. The goal is to facilitate comparison when indexing. Thus, words that have variations of capital letters can be considered equivalent.

Table 2. Data before case folding process

No	Account Name	Time	Reviews	Sentiment
1	Cinta Sejati	4 days ago	Of the several hospitals I visited in Sukoharjo. Here the hospital service is the longest, less satset. It's a pity about the management.	Negative
2	Nurafanti Afanti	1 month ago	slow	Negative
3	Rizky Ahmad	2 month ago	Pharmacy service is very poor	Negative
4	Diah Retno Hermastuti	4 month ago	good service and adequate facilities	Positive
5	Sony Cahtono	4 month ago	Good service	Positive
6	Safira Dila	4 month ago	Fast service	Positive
7	Eva Yohana	7 month ago	Limited space for queuing in the clinic... Not worth it for visitors	Negative
8	Freddick Listern	1 year ago	Slow administration	Negative
9	Muhammad Faeyza	2 year ago	Good service satisfying	Positive
10	Rayi Rayendaria	2 year ago	A comfortable hospital in my opinion. Friendly staff	Positive
11	Alena Valency	2 year ago	Vaccines at PKU Muhammadiyah Sukoharjo can comfortably get a gift	Positive
12	Awan Dwi Kurniawan	2 year ago	Friendly hospital staff	Positive
13	Ryan Slametraharjo	2 year ago	Good hospital, very good service	Positive
14	Suratmi	2 year ago	The service was satisfactory and pleasant	Positive
15	Anang Sang Ship's Cook	2 year ago	Good hospital services	Positive
16	Ririn Murtiah	2 year ago	Good service	Positive
17	Agus Triutomo	2 year ago	The service was very satisfying. Thank you	Positive
18	Rafidha Adelina Siregar	3 year ago	Satisfactory service	Positive
19	Yussuf Faisal	3 year ago	Very good service	Positive
20	Gilang Erlanda	3 year ago	The nurses are not friendly and the service is not satisfactory.	Negative

In Table 2 above, before the case folding process, the letters in the review data are still partially capitalized, and there are still numbers and punctuation marks that are not needed.

Table 3. Case folding results

No	Account Name	Time	Reviews	Sentiment
1	Cinta Sejati	4 days ago	of the several hospitals I visited in Sukoharjo. here the hospital service is the longest, less *satset*. It's a pity about the management.	Negative

No	Account Name	Time	Reviews	Sentiment
2	Nurafanti Afanti	1 month ago	slow	Negative
3	Rizky Ahmad	2 month ago	pharmacy service is very poor	Negative
4	Diah Retno Hermastuti	4 month ago	good service and adequate facilities	Positive
5	Sony Cahtono	4 month ago	good service	Positive
6	Safira Dila	4 month ago	fast service	Positive
7	Eva Yohana	7 month ago	limited space for queuing in the clinic not worth it for visitors	Negative
8	Freddick Listern	1 year ago	slow administration	Negative
9	Muhammad Faeyza	2 year ago	good service satisfying	Positive
10	Rayi Rayendaria	2 year ago	a comfortable hospital in my opinion. friendly staff	Positive
11	Alena Valency	2 year ago	vaccines at PKU Muhammadiyah Sukoharjo can comfortably get a gift	Positive
12	Awan Dwi Kurniawan	2 year ago	friendly hospital staff	Positive
13	Ryan Slametraharjo	2 year ago	good hospital, very good service	Positive
14	Suratmi	2 year ago	the service was satisfactory and pleasant	Positive
15	Anang Sang Ship's Cook	2 year ago	good hospital services	Positive
16	Ririn Murtiah	2 year ago	good service	Positive
17	Agus Triutomo	2 year ago	the service was very satisfying. thank you	Positive
18	Rafidha Adelina Siregar	3 year ago	satisfactory service	Positive
19	Yussuf Faisal	3 year ago	very good service	Positive
20	Gilang Erlanda	3 year ago	the nurses are not friendly and the service is not satisfactory	Negative

In Table 3 above, after the case folding process, letters are generalized into lowercase letters, and the removal of punctuation marks, numbers, and characters other than alphabets is involved. Thus, case folding can help ensure consistency in the use of letters and facilitate text analysis. The source code image of the case folding process can be seen in Figure 3 below.

```
# ubah teks menjadi lower
data['text_cleaning'] = data['text_cleaning'].str.lower()
data.head()
```

Figure 3. Source code case folding

b. Tokenizing

After case folding, the data then enters the tokenizing stage, which is the breakdown of words in the data into words, as in Table 4 below.

Table 4. Tokenizing result

No	Reviews
1	[of, several, hospitals, which, I, visited, in, Sukoharjo, here, service, hospital, the, longest, less, fast, regret, once, management]
2	[slow]
3	[service, pharmacy, very, bad]
4	[service, good, and, facilities, adequate]
5	[service, good]
6	[service, fast]
7	[space, limited, to, queue, at, clinic, not, worth, it, for, visitors]

 DOI: https://doi.org/10.52435/jaiit.v6i1.566

No	Reviews
8	[administration, slow]
9	[service, which, good, satisfactory]
10	[home, hospital, which, comfortable, according, to, me, staff, which, friendly]
11	[vaccine, in, pku, muhammadiyah, sukoharjo, can, with, comfort, get, gift]
12	[staff, hospital, who, friendly]
13	[hospital, hospital, which, good, service, which, very, good]
14	[service, satisfactory, and, pleasant]
15	[service, hospital, good]
16	[service, good]
17	[service, very, satisfying, thank you]
18	[service, satisfactory]
19	[service, very, good]
20	[nurse, not, friendly, and, the service, not, satisfactory]

c. Filtering

After the data tokenizing process enters the next stage, namely the filtering stage, the filtering process itself is useful for filtering or eliminating words that are not needed. The results of the filtering process can be seen in Table 5 below.

Table 5. Filtering result

Before Filtering	After Filtering
Pharmacy service is very poor	Poor pharmacy service

d. Stemming

The next stage is stemming, stemming functions to convert the affixed words in the data into basic words. To do so, the Sastarawi library is needed. The results of the stemming process can be seen in Table 6, and the literary import process in Figure 4 below.

Table 6. Stemming result

Before stemming	Afer stemming
Pharmacy service is very poor	poor pharmacy serv

```
# Stopword
import Sastrawi
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory, StopWordRemover, ArrayDictionary
more_stop_word = []

stop_words = StopWordRemoverFactory().get_stop_words()
new_array = ArrayDictionary(stop_words)
stop_words_remover_new = StopWordRemover(new_array)

def stopword(str_text):
  str_text = stop_words_remover_new.remove(str_text)
  return str_text

data['text_cleaning'] = data['text_cleaning'].apply(lambda x: stopword(x))
data.head()
```

Figure 4. Import Sastrawi

3.3. Data Split

The next process after preprocessing is to divide the data set into two, namely train data or train data set and test data or test dataset.

```
[ ]  #memsihkan data menjadi data latih dan data training
     X = data_clean['text_cleaning']
     y = data_clean['sentimen']
     X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.10, random_state=42)
```

Figure 5. Data Split

3.4. Naive Bayes Classification Results

After the preprocessing process is complete, the data that will be processed for testing the naïve Bayes model is obtained. Testing the amount of training and test data is done by experimenting with each comparison until you get the highest accuracy. The results of the comparison of accuracy values in testing the amount of training data and test data can be seen in Table 7, and the testing process to determine the accuracy value of each comparison can be seen in Figure 5.

Table 7. Comparison of training and test data

Comparison of training and test data	*Accuracy*
90%:10%	90%
80%:20%	79%
70%:30%	78%
60%:40%	79%
50%:50%	78%

The results from the above figure show that the accuracy for the positive class is higher than the accuracy for the negative class. As a result, the overall accuracy reaches 90%.

```
# Evaluasi Model
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred, target_names=['negatif', 'positif'])

print(confusion_matrix(y_test, y_pred))
print("Akurasi Model Naive Bayes : ", accuracy)
print("\nLaporan Klasifikasi :\n", classification_rep)
```

Figure 6. Testing the evaluation of the naive bayes model

Figure 5 above shows the source code for calculating the accuracy value of each comparison of training data and test data, and from Table 7 above, the highest accuracy value is 90%, with a ratio of training data and test data of 90%: 10%.

3.5. Confusion Matrix Result

The next stage, namely the confusion matrix stage above, shows that the four negative data tested are divided into 4 True negative data and 0 data as a false negative. While 15 positive data were tested, there were 15 True Positive data and 2 False Negative. The results of the confusion matrix can be seen in Table 8 below.

Table 8. Confusion Matrix Results

	True	False
Negative	4	0
Positive	15	2

Overall, the values of accuracy, precision-recall, and F1-measure, as well as the data above, have the highest accuracy values. The highest ratio is 90%: 10%. This is due to differences in the number and distribution of data in each aspect. The results themselves included a 90% accuracy value, 94% precision, 83% recall, and 87% F1 score.

 DOI: https://doi.org/10.52435/jaiit.v6i1.566

```
# Evaluasi Model
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred, target_names=['negatif', 'positif'])

print(confusion_matrix(y_test, y_pred))
print("Akurasi Model Naive Bayes : ", accuracy)
print("\nLaporan Klasifikasi :\n", classification_rep)
```

Figure 7. Model evaluation

3.6. K-Fold Cross Validation Testing

K-Fold Cross Validation testing on Naïve Bayes Classifier algorithm with the amount of training and test data 90%; 10%. Through this test, data validation will be carried out using K-Fold Cross-Validation. The best results on 6 th fold. K-Fold cross-validation test results test results of the Naïve Bayes algorithm are shown in Figure 8.

```
from sklearn.model_selection import cross_val_score

# Definisikan model
naive_bayes = MultinomialNB()

# Lakukan cross-validation (misalnya, 10-fold)
scores = cross_val_score(naive_bayes, X_train_resampled, y_train_resampled, cv=10)

# Cetak skor untuk setiap fold
print("Skor Akurasi untuk Setiap Fold:", scores)
```

```
Skor Akurasi untuk Setiap Fold: [0.67857143 0.82142857 0.85714286 0.75       0.82142857 0.92857143
 0.78571429 0.85714286 0.85185185 0.62962963]
```

Figure 8. K-fold cross validation

Based on Figure 8, it is known that iterations with the highest accuracy results are at the sixth iteration with an accuracy of 0.92857143 and can be rounded to 0.93 or 93%. The average value of the K-Fold cross-validation test is 75%.

3.7. WordCloud Result

In the last stage, we visualize the data using wordcloud to identify words that appear frequently in each sentiment. Below is the wordcloud result for positive sentiment, which can be seen in Figure 9.

Figure 9. Positive Sentiment WordCloud

In Figure 9, it can be seen that the words that often appear in positive sentiments are as follows: "good", "friendly", "service", "fast" and others. For negative sentiment, wordcloud can be seen in Figure 10.

Figure 10. Negative sentiment WordCloud

In Figure 10, it can be seen that the words that often appear in negative sentiments are as follows: "service", "long", "queue", "doctor", "less", and others.

4. Conclusion

From the results of the above research, namely regarding the analysis of visitor sentiment on the services and facilities of PKU Muhammadiyah Sukoharjo Hospital on Google Maps, which amounted to 209 data that had passed the text preprocessing stage consisting of transformation, tokenization, filtering, stemming and classification using the Naïve Bayes algorithm using the Chrome extension instant data scraper, it can be concluded that the reviews of PKU Muhammadiyah Sukoharjo Hospital visitors are classified as positive with the results of 90% accuracy presentation for 94% precision value and for 83% recall value. This value shows that the Naive Bayes algorithm classification is quite good in processing Google Maps review data because the accuracy percentage results reach 90%. This value proves that public sentiment regarding the services and facilities of PKU Muhammadiyah Sukoharjo Hospital is positive.

Acknowledgement

I would like to thank the general hospital PKU Muhammadiyah Sukoharjo for permission to access visitor review data and the supervisors who have provided direction, input, and guidance.

References

[1] T. A. Q. Putri, A. Triayudi, and R. T. Aldisa, "Implementasi Algoritma Decision Tree dan Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks," *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 641–649, 2023, doi: 10.47065/josh.v4i2.2949.

[2] F. U. Haq, "Penggunaan Google Review Sebagai Penilaian Kepuasan Pengunjung Dalam Pariwisata," *Tornare*, vol. 2, no. 1, p. 10, 2020, doi: 10.24198/tornare.v2i1.25826.

[3] B. R. Tsania Dzulkarnain, Dian Eka Ratnawati, "PENGGUNAAN METODE NAÏVE BAYES CLASSIFIER PADA ANALISIS THE USE OF THE NAÏVE BAYES CLASSIFIER METHOD IN SENTIMENT ANALYSIS OF THE COMMUNITY ' S ASSESSMENT OF HOSPITAL SERVICES IN," vol. 10, no. 7, pp. 1453–1460, 2023, doi: 10.25126/jtiik.2023107979.

[4] F. F. Mailoa, "Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia," *J. Inf. Syst. Public Heal.*, vol. 6, no. 1, p. 44, 2021, doi: 10.22146/jisph.44455.

[5] Gilbert, Syariful Alam, and M. Imam 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: 10.55123/storage.v2i3.2333.

[6] H. Chyntia Morama, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen berbasis Aspek terhadap Ulasan Hotel Tentrem Yogyakarta menggunakan Algoritma Random Forest Classifier," 2022. [Online]. Available: http://j-ptiik.ub.ac.id

[7] W. Wijiyanto and S. Sopingi, "Kontribusi Keluarga Dalam Prediksi Mahasiswa Lulus Tepat Waktu Menggunakan Model Support Vector Machine," *DutaCom*, vol. 17, no. 1, pp. 25–36, 2023, doi: 10.47701/dutacom.v17i1.3784.

[8] H. F. Arief Asro, "Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor," *KHATULISTIWA Inform.*, vol. 10, pp. 17–24, 2022.

[9] R. W. Abdullah, D. Hartanti, H. Permatasari, A. W. Septyanto, and Y. A. Bagaskara, "Penerapan Data Mining untuk Memprediksi Jumlah Produk Terlaris Menggunakan Algoritma Naive Bayes Studi Kasus (Toko Prapti)," *J. Ilm. Inform. Glob.*, vol. 13, no. 1, pp. 20–27, 2022, doi: 10.36982/jiig.v13i1.2060.

[10] A. C. Fauzan and K. Hikmah, "Implementasi Algoritma Naive Bayes Dalam Analisis Polarisasi Opini Masyarakat Terkait Vaksin Covid-19," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 2, pp. 122–128, 2022, doi: 10.36341/rabit.v7i2.2403.

[11] A. U. T. Ama, D. N. Mulya, Y. P. D. Astuti, and I. B. G. Prasadhya, "Analisis Sentimen Customer Feedback Tokopedia Menggunakan Algoritma Naïve Bayes," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, p. 50, 2022, doi: 10.30865/json.v4i1.4783.

[12] 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.

[13] N. M. A. Mahar, Vihi Atina, and Nugroho Arif Sudibyo, "Pemodelan Prediksi Kelulusan Mahasiswa Dengan Metode Naïve Bayes Di Uniba," *J. Manaj. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 148–158, 2023, doi: 10.36595/misi.v6i2.875.

[14] R. Rachman and R. N. Handayani, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.

[15] S. Alam, M. G. Resmi, and N. Masripah, "Classification of Covid-19 vaccine data screening with Naive Bayes algorithm using Knowledge Discovery in database method," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 4, no. 2, pp. 177–185, 2022, doi: 10.47709/cnahpc.v4i2.1584.

[16] Karsito and S. Santi, "Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah Dengan Algoritma Naïve Bayes Di Perumahan Azzura Residencia," *J. Teknol. Pelita Bangsa*, vol. 9, pp. 43–48, 2019.

[17] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 697–711, 2021.

This page intentionally left blank

 DOI: https://doi.org/10.52435/jaiit.v6i1.566