Sentiment Analysis On Tripadvisor Travel Agent Using Random Forest, Support Vector Machines, and Naïve Bayes Methods

Ariq Ammar Fauzi[1], Anik Vega Vitianingsih[2*], Slamet Kacung[3], Anastasia Lidya Maukar[4], Seftin Fiti Ana Wati[5]

[1,2,3]Informatics Department, Universitas Dr. Soetomo, Surabaya, East Java, Indonesia
[4]Industrial Engineering Department, President University, Bekasi, West Java, Indonesia
[5]Information System Department, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, East Java, Indonesia
E-mail: [1]ariqaf682@gmail.com, [2*]vega@unitomo.ac.id, [3]slamet@unitomo.ac.id [4]almaukar@president.ac.id, [5]seftin.fitri.si@upnjatim.ac.id

(Received: 10 Feb 2025, revised: 20 Feb 2025, accepted: 21 Feb 2025)

Abstract

TripAdvisor faces problems in improving the quality of service on its application, namely the presence of unexpected or non-functional features, which can affect the user experience and reduce trust in the application. This research aims to develop an application capable of performing sentiment analysis on TripAdvisor application user reviews on the Google Play Store with negative, positive, and neutral classifications using the Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB). The RF method was chosen in this study because of its ability to handle large and complex data very accurately, while SVM is able to classify data on a large scale and is resistant to overfitting, while NB is able to classify text with clear probabilities. The Lexicon-based method as data labelling. The results of sentiment analysis from 1500 reviews with web scrapping show the classification of positive, negative, and neutral sentiments of 48, 726, and 646 data, respectively. Model performance in RF, SVM, and NB testing gets an accuracy value of 94%, 93.6%, and 77.8%, respectively. The RF model produces the best accuracy compared to other methods. The RF model produces the best accuracy compared to other methods. The results of sentiment analysis from 1500 user reviews allow developers to identify features that are often criticized or do not function properly in their application services.

Keywords: TripAdvisor Sentiment Analysis, Random Forest, Support Vector Machines, Naïve Bayes.

I. INTRODUCTION

The COVID-19 pandemic has had a significant impact on various sectors of the economy, including the tourism industry in Indonesia. Travel restrictions and the closure of tourist destinations have led to a drastic decrease in the number of tourists, which has had an impact on the decline in income in the tourism sector and other related sectors, such as hospitality, transportation, and culinary [1]. As one of the popular online travel agents, TripAdvisor still faces several problems in improving the quality of service on the application to meet user needs. This includes unexpected features or features that do not function properly, which can affect the user experience and reduce user trust in the application. User reviews can indicate that quality review or development should be improved to make the app relevant and useful to users.

This study focuses on several key aspects that need to be examined, namely how to analyze user sentiment regarding the services provided by Tripadvisor. This evaluation was conducted to understand the level of user satisfaction with Tripadvisor's services and to find out about their experience using the platform. This step is expected to provide an overview of the extent to which the services provided are able to meet user expectations and identify areas that still need improvement. This study also explores how reviews provided by users can contribute to the management strategy of online travel agents (OTAs). Reviews published by users can be valuable input for OTAs in designing policies and strategies to improve service quality. Therefore, a systematic approach is needed to analyze and understand how these reviews can be used as a reference in a more effective decision-making process.

DOI: 10.34148/teknika.v14i1.1198

This research refers to previous studies by researchers [2] that implemented the long short-term memory (LSTM) algorithm with tourism, reviews, and TripAdvisor parameters, which obtained an accuracy value of 71.67%. Researchers [3] used the SVM, K-Nearest Neighbor (K-NN), and NB methods with the Google Play Store TripAdvisor dataset. They obtained the value with the best accuracy using the SVM algorithm at 89.8%. Another researcher [4] used the NB method with sentiment analysis parameters, pegipegi.com, with an accuracy rate of 89.5%. Based on literature studies in previous research, many researchers have used the NB and SVM methods in the context of this study. However, in this study, the RF method was developed as an alternative to be compared with NB and SVM. According to researchers [5] Random Forest is one of the classification algorithms that has a high level of accuracy that can reduce overfitting of individual models. In recent years, this algorithm has been increasingly used for classification tasks because its performance has proven superior to SVM, Naïve Bayes, and other machine learning algorithms. This development aims to evaluate the extent to which RF can improve performance in data analysis and determine classification models with a more optimal level of accuracy.

This research aims to develop an application capable of performing sentiment analysis on TripAdvisor application user reviews on the Play Store with negative, positive, and neutral classifications for OTA service improvement. The dataset used in this research is from user reviews of the TripAdvisor Google Play Store application. This research uses the Lexicon-based method for data labelling, while the RF, SVM, and NB methods identify reviews in the context of classification. RF can predict individual trees and their correlation [6], and SVM can separate sentiment classes and utilize information [7]. At the same time, NB is widely used by many researchers because it outperforms other algorithms in terms of classification [8]. Lexicon-based is used in this study because it does not require large training data [9]. RF, SVM and NB classification methods will be used to perform sentiment analysis on user reviews. An evaluation of the accuracy of each method will be generated to determine the most effective algorithm for classifying the sentiment of user reviews so as to produce more accurate insights for OTA management in improving the quality of the services they offer.

II. RESEARCH METHODOLOGY

This research analyzes user reviews of the TripAdvisor application using RF, SVM, and NB classification models to improve the application's quality of service. The flow of this research is illustrated in Figure 1.

In this research methodology, the initial stage is web scraping to collect Tripadvisor data. Then, the data is processed through a pre-processing stage, which involves cleaning, case folding, tokenizing, normalizing, filtering, and stemming. Furthermore, the data is labeled based on lexicons and grouped into three categories: positive, negative, and neutral. Then, TF-IDF is used for word weighting and classification using RF, SVM, and NB methods. In the last stage, the confusion matrix is evaluated and obtains accuracy, precision, and recall values.

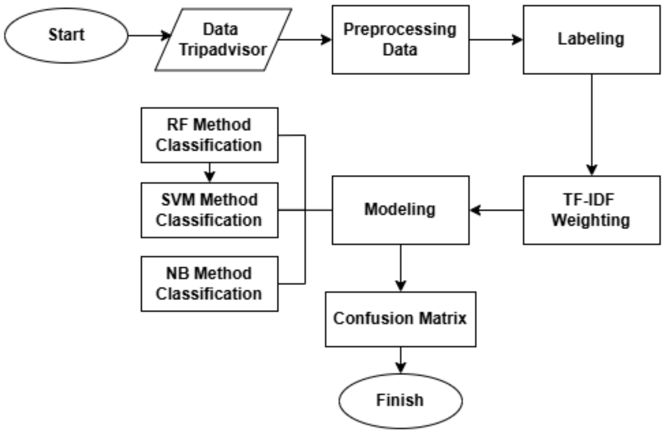

Figure 1. Research Flow

A. Data Retrieval

Data retrieval uses the google_play_scrapper API with an existing module in the Python library, which involves an API endpoint to retrieve data on reviews, ratings, and other information. The data covers the period from 2022 to 2024 with a total of 1,500 review data

B. Data Pre-processing

Pre-processing is a method to convert raw text into useful information that is performed in several phases, such as cleaning, tokenizing, case folding, normalized, stemming, and filtering [9]. After pre-processing decisions are made and applied on 1,500 dataset, features (variables) are extracted from the text using text mining with closed vocabulary, open vocabulary, or both. Figure 2 is the pre-processing flow.

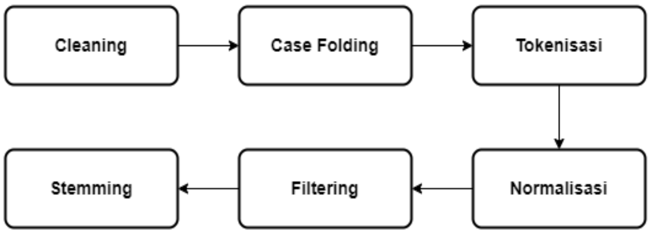

Figure 2. Pre-Processing Flow

1) Cleaning: This process cleans data such as unnecessary or irrelevant components, special characters, and numbers.
2) Case Folding: The case folding process converts text into lowercase letters.
3) Tokenization: The tokenization process separates the text into words or small units.
4) Normalization: Text normalization aims to make the format more consistent, including removing irrelevant punctuation, converting letters to lowercase, and making other necessary adjustments.

DOI: 10.34148/teknika.v14i1.1198

ISSN 2549-8037, EISSN 2549-8045

5) Filtering: The filtering stage is the process of removing texts that are irrelevant and have no significant meaning in the analysis.
6) Stemming: This stage removes affixed words to get a text's base word or basic form.

C. Data Labelling

This process can be done manually and automatically with the help of machine learning. Manual data labelling is not effective because it is very subjective. Automatic labelling can use sentiment lexicon datasets or emotion dictionaries, and this method is good for large amounts of text to minimize document labelling errors. The score calculation will be grouped as a negative class if the score is less than 0, a positive class if the score is more than 0, and a neutral class if the score equals 0. The sentiment score calculation uses Equation 1 [9].

$$Score = (\Sigma positif\ words - \Sigma negatif\ words\ ) \qquad (1)$$

The lexicon-based labelling approach has many significant advantages that make it especially important because it does not require training data when it is difficult to obtain manually labelled data or unavailable in sufficient quantities. Researchers can perform sentiment analysis directly by utilizing existing sentiment dictionaries without going through the time-consuming data labelling process. Furthermore, this approach offers a high degree of interpretability and openness, as any emotional value assigned to a word or sentence can be attributed to the value of the lexicon used. This helps researchers understand and describe their analysis findings more clearly [9].

D. Word Weighting

At this stage, the weighting method used in text analysis to determine how significant a word is compared to the rest of the document set is TF-IDF (*Term Frequency-Inverse Document Frequency*) [10]. This method combines two principles of weight calculation: how often a word appears in a particular document and how rarely it appears in other documents. Inverse document frequency (IDF) is the number of documents containing a term based on all documents in the dataset [11]. TF-IDF weighting consists of two components: tf and idf. Equation 2 shows the formula used to calculate TF-IDF weighting, $tftd$ shows the frequency of term t in document d. The variable N represents the total number of documents in the entire corpus, dft shows the number of documents that contain the term in *t* variable [9].

$$TF.IDF_{std}(t) = tf_d^t\ x\ log\frac{N}{df^t} \qquad (2)$$

E. Model Classification

The review data in this study will be classified using the RF, SVM, and NB methods.

1) Random Forest: RF is an algorithm for regression and classification tasks [12]. Random Forest is an ensemble learning method based on the Decision Tree (D-Tree) algorithm and consists of a number of decision trees as classifiers. Classification results are determined based on the most dominant decision from all decision trees used. This process of voting from various decision trees contributes to the increased accuracy of Random Forest Random forest operates in a different way compared to Decision Tree (D-Tree). D-Tree only builds a classification model using one large decision tree, whereas Random Forest builds many smaller decision trees. This approach makes Random Forest have better performance, especially in terms of accuracy [5].

2) Support Vector Machine: SVM is an algorithm in machine learning that focuses on data classification by finding the optimal hyperplane or dividing line. This algorithm maximizes the margin between data through a minimization method using Lagrange techniques. SVM is known to have a solid theoretical basis and can provide high classification accuracy compared to many other algorithms [9]. Equation 3 is used in the data classification process of SVM, where the relationship between input x and the resulting class is represented by the variable (x). The weight vector w serves to determine the hyperplane's direction, whereas the feature vector x illustrates the input data. Meanwhile, the bias value b is used to shift the position of the hyperplane from its starting point [9].

$$f(x) = w.x + b \qquad (3)$$

The choice of kernel affects the final classification result in the SVM algorithm. In SVM, the kernel function has a crucial role in improving the performance of the algorithm [13]. There are three types of kernels commonly used in SVM.

1) Kernel Linear (Equation 4)
$$K(x, x_i) = sum(x * x_i) \qquad (4)$$

2) Kernel Polynomial (Equation 5)
$$K(x, x_i) = 1 + sum(x * x_i) \qquad (5)$$

3) Kernel RBF (Equation 6)
$$K(x, x_i) = \exp\ (-gamma * sum((x * x_{i^2}) \qquad (6)$$

Where:
K(x,xi) = Kernel Function
x = Training data chunks

3) *Naïve Bayes:* The NB model is very easy to build. This model is suitable for large data and shows excellent classification results. This algorithm separates the data into two groups: data used for the model training process (training) and data used to test the performance of the model (testing) [14]. Naïve Bayes is a classification with probability and statistical methods proposed by British scientist Thomas Bayes, which predicts future opportunities based on previous experience, known as Bayes' Theorem [15]. Equation 7 is a method for calculating posterior probabilities.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{7}$$

Where:
P(c│x) = Posterior probability on class with known predictor attributes
P(c) = Prior probability of the class
P(x│c) = probability of attribute x with known class
P(x) = Probability prior of predictor

F. Evaluation Model

A crucial stage in classification is the assessment of model performance [16]. The Confusion Matrix mechanism will be applied to evaluate model performance as the final stage in this research. The confusion matrix produces important metrics such as accuracy, precision, and recall to assess classification performance [9]. The F1 score provides a balanced picture of precision and recall in one value. An in-depth description of the mechanism's performance is essential in this research; hence, a confusion matrix is used. Accuracy measures the proportion of correct positive and negative predictions to the overall data, calculated using Equation 8. Precision is calculated using the ratio of correct positive predictions to total positive predictions, described in Equation 9. In comparison, recall is the ratio between the number of correctly predicted positive events and the total positive events that occurred, calculated by Equation 10 [9].

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

$$Recall = \frac{TP}{TP+FN} \tag{10}$$

True Positive (TP) indicates the number of examples from the positive class that are correctly identified. Conversely, a False Positive (FP) occurs when an instance from a negative class is mistakenly classified as positive. False Negative (FN) occurs when examples from the positive class are mistakenly classified as negative, while True Negative (TN) refers to examples from the negative class that correctly identified [9].

III. RESULT AND DISCUSSION

This chapter presents the results of the data analysis that has been carried out, and the study results will be displayed in the form of tables, graphs, and narrative explanations to provide a clear picture of the performance of the model used.

A. Data Retrieval

Between 2022 and 2024, review data on the TripAdvisor app was collected using a scraping technique utilizing the Google Play Scrapper library, and 1,500 data points were collected. The data retrieved from the reviews is stored in a file with CSV format. Table 1 is the result of data scrapping.

Table 1. Data Scrapping Results

Star	Created At	Full Text
5	2023-06-28	*Wow aplikasi ini sangat bagus dalam mencari tempat wisata terdekat*
4	2022-12-03	*Dengan adanya TripAdvisor kita jadi mudah menemukan tempat yg bagus untuk dikunjungi*
1	2024-01-29	*Saya instal aplikasi ini kok nggak bisa dibuka?*
1	2024-12-24	*Aplikasi ini tipu2, sudah pesan dan sudah di bayar tapi d batalkan secara sepihak. Dan setelah saya refund selalu di tolak! Aplikasi penipuan!*
5	2023-07-06	*Lumayan*
5	2022-03-20	*Bermanfaat*

From the scrapping results in Table 1, there are a variety of reviews with varying ratings, ranging from positive to negative. Some reviews are written in detail and provide deep insights, while others are short, consisting of just one word or sentence. This process enabled the identification of aspects that needed to be improved in the app based on user feedback.

B. Text Pre-Processing

Table 2 is one of the results of text pre-processing from scraping 1,500 data. This raw data consists of review texts that vary in length and writing style, reflecting users' subjective opinions of the app. Text pre-processing aims to prepare the text data to make it easier to process and analyze. These stages include Cleaning, Case Folding, Tokenizing, Normalized, Filtering, and Stemming.

Table 2. Pre-processing Result

Raw Data	
Sangat bermanfaat dan banyak membantu usaha kami. Terimakasih TripAdvisor	
Text Pre-Processing	
Cleaning	*Sangat bermanfaat dan banyak membantu usaha kami Terimakasih TripAdvisor*
Case Folding	*sangat bermanfaat dan banyak membantu usaha kami terimakasih TripAdvisor*
Tokenization	*sangat,bermanfaat,dan,banyak,membantu, usaha,kami,terimakasih,TripAdvisor*
Normalization	*sangat,bermanfaat,dan,banyak,membantu, usaha,kami,terimakasih,TripAdvisor*
Filtering	*bermanfaat,membantu,usaha,terimakasih, TripAdvisor*
Stemming	*manfaat,bantu,usaha,terimakasih,TripAdvisor*

C. Labelling

At this stage, the text data that has been processed will be labelled using a Lexicon-Based dictionary, referring to Equation 1. The results of this labelling process obtained 726 positive sentiments, 48 negative sentiments, and 646 neutral sentiments. Table 3 is the result of the labelling phase.

Table 3. Labelling Results

Cleaned_text	Score	Sentiment
TripAdvisor bantu traveler temu indah senang dunia bantu bisnis sukses baik pokok top	4	Positive
instal posting foto susah buka atur hp sulit	-3	Negative
coba praktek	0	Neutral

Table 3 is the output of the data labeling process, which aims to determine the sentiment of each review before further classification. This table has three main columns, namely Cleaned_Text, Score, and Sentiment. Cleaned_Text contains the review text that has gone through the processing stage. Score shows the sentiment score obtained using the Lexicon-Based approach. Sentiment is a sentiment category based on the score obtained, where reviews are categorized as positive if the score is more than 0, negative if the score is less than 0, and neutral if the score is equal to 0. Reviews with positive sentiments generally contain words such as helpful, happy, successful, and good, which give a high score and are classified as positive. In contrast, reviews with negative sentiment tend to contain words such as difficult, open, and difficult, which reflect user dissatisfaction and result in a negative score. Meanwhile, neutral sentiment includes reviews that do not have any strong emotionally charged words or are merely informative, resulting in a score of 0 and categorized as neutral.

After the labelling process, the information from the word frequency will be grouped and visualized through a word cloud. Each review is associated with positive, negative, or neutral words, which represents a graphic of the words in a text. The size of each word reflects the frequency of its occurrence in the analyzed data. The larger a word is displayed, the more often it appears in the text. Figure 3 is a Word Cloud Sentiment Analysis Result that represents a graphic of the words contained in a text. The size of each word reflects the frequency of its occurrence in the analyzed data. The larger a word is displayed, the more often it appears in the text.

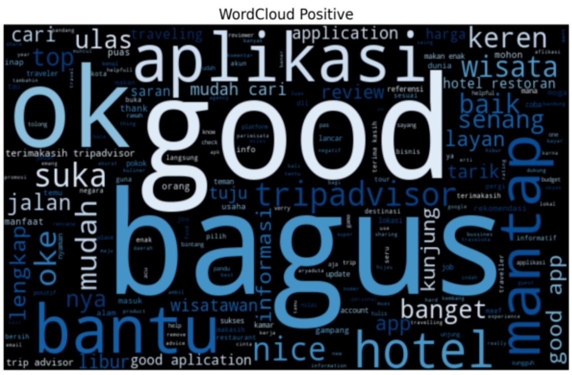

Figure 3. Word Cloud Sentiment Analysis Results Positive

Figure 4. Word Cloud Sentiment Analysis Results Negative

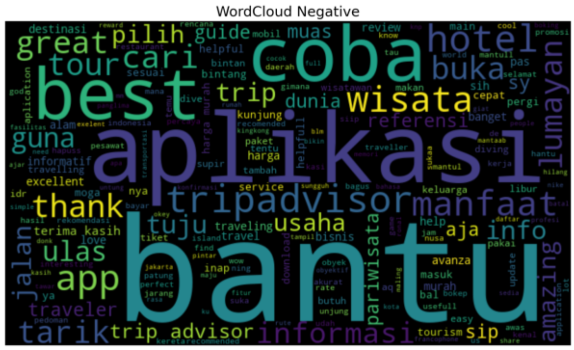

Figure 5. Word Cloud Sentiment Analysis Results Neutral

Based on Figure 3 illustrate that words such as *"bagus* (good)*", "bantu* (help)*", "mantap* (good)*", "ok* (ok)*", "aplikasi* (app)*"* are frequently used in the positive reviews, indicating user satisfaction with the app's features and performance. Figure 4 shows that *"jelek* (ugly)*", "ribet* (complicated)*", "sulit* (difficult)*", "mahal* (expensive)*", "payah* (lame)*"* are negative reviews that often appear on the TripAdvisor application. These words indicate dissatisfaction with the user experience, whether featured, prices, services, or ease of use. Figure 5 shows frequent neutral reviews such as *"coba* (try)*", "wisata* (travel)*", "bantu* (help)*"*. These words reflect opinions that do not have strong emotional sentiments and thus describe an informative and descriptive experience.

D. Evaluation

Evaluation is the final step. After the data is processed, an evaluation describes the analysis results. Currently, the system assesses the classification results in various ways. First, the weight or importance of each word in the comment is calculated using TF-IDF extraction using the Python library. Next, a confusion matrix presents the classification result by displaying the number of accurate and inaccurate predictions for each sentiment category. For ease of interpretation, the analysis results are visualized as diagrams. After pre-processing and labelling, this process produces clean data of 1500 TripAdvisor app user comments that have been transformed. It can be analyzed using Lexicon-based techniques and RF, SVM, and NB methods to classify the data set using text polarity, and these three methods generate negative, positive, and neutral values. This research uses data in Indonesian that follows the lexicon-based approach so that the categorization is more precise and accurate.

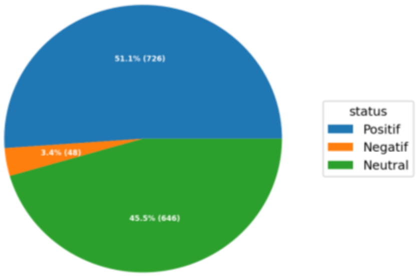

Figure 3. Sentiment Classification Results

Based on Figure 6, the polarity of positive sentiment reaches 51.1%, neutral polarity is 3.4%, and negative polarity is 45.5%. The next step is to test the performance of the applied classification model. Table 4 displays the accuracy, precision, and recall outcomes of data processing from the start to the assessment phase.

Table 4. Confusion Matrix Results

Confusion Matrix				
	Accuracy	Precision	Recall	F1-Score
RF	94%	68%	96%	69.7%
SVM	93.6%	74%	95.7%	78.5%
NB	77.8%	53%	54%	52%

Referring to Equations 8, 9, and 10, Table 4 is the result of the confusion matrix, the SVM algorithm provides the best overall performance, with the highest F1-score value of 78.5%. The RF algorithm showed excellence in the recall, with the highest value of 96%, making it suitable for use when the main priority is to capture all positive instances. In contrast, the NB algorithm has the lowest performance on all metrics, having values of 77.8%, 53%, 54%, and 52%, respectively, making it less recommended for the dataset used. Thus, SVM is the most optimal algorithm for this study. SVM algorithms are effective in overtraining, especially on small datasets, because the maximum likelihood approach used by SVM can produce a more stable and generalizable model [17]. Meanwhile, the accuracy value of RF is higher because the RF method can reduce overfitting in individual models.

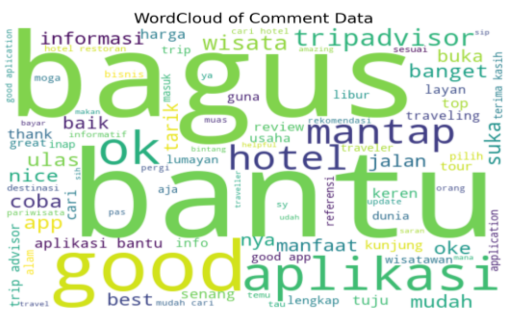

Figure 4. Word Cloud Data

Figure 7 illustrates the reviews given on the Tripadvisor application in a Word cloud format. This illustration provides an overall picture of user perceptions of the Tripadvisor application.

IV. CONCLUSION

Based on the analysis results, this study shows that the SVM algorithm provides the best performance in sentiment analysis on TripAdvisor app user reviews, with the highest F1-score of 78.5%. The RF algorithm excels in recall with a value of 96%, making it a good choice if the priority is to detect all positive instances. In contrast, the NB algorithm showed the lowest performance on all evaluation metrics, including 77.8% accuracy, 53% precision, 54% recall, and 52% F1-score, making it less recommended for this dataset. It can be concluded that SVM proved to be more effective in generating stable and generalized models, especially on small datasets. Meanwhile, the accuracy value of RF is higher because the RF method can reduce overfitting in individual models. This research shows that SVM is superior on the TripAdvisor review dataset, increasing our understanding of SVM's superiority in creating more stable models on small datasets.

For future research, it is recommended that further exploration of deep learning models, such as LSTM or BERT, be carried out to compare their performance with SVM. This aims to find out whether deep learning models can provide more optimal results in sentiment analysis, especially on the TripAdvisor app user review dataset.. For OTA developers, the application of RF in real-time sentiment analysis can help improve service quality, for example by providing more precise recommendations and responding more quickly to negative reviews.

REFERENCES

[1] P. U. Rukmana, O. N. Pratiwi, and H. Fakhrurroja, “Perbandingan Analisis Sentimen Aplikasi Traveloka dan Tiket.com pada Twitter dengan Metode Support Vector Machine,” *J. Sist. Cerdas*, vol. 6, no. 3, pp. 241–250, 2023, doi: 10.37396/jsc.v6i3.350.

[2] J. Nurvania, J. Jondri, and K. M. Lhaksamana, “Analisis Sentimen Pada Ulasan di TripAdvisor Menggunakan Metode Long Short-Term Memory (LSTM),” *eProceedings Eng.*, vol. 8, no. 4, 2021, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15240

[3] A. M. Ndapamuri, D. Manongga, and A. Iriani, “Analisis Sentimen Ulasan Aplikasi Tripadvisor Dengan Metode Support Vector Machine, K-Nearest Neighbor, Dan Naive Bayes,” *INOVTEK Polbeng - Seri Inform.*, vol. 8, no. 1, p. 127, 2023, doi: 10.35314/isi.v8i1.3260.

[4] M. Naufal, B. Balit, F. S. Utomo, I. S. Program, and C. Java, “Sentiment Analysis of pegipegi . com Review on

Google Play Store with Naïve Bayes," vol. 13, pp. 1044–1053, 2024, [Online]. Available: https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/3913

[5] T. S. Kasus, P. Pada, and S. Tripadvisor, "Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR)," vol. 8106, pp. 21–29, 2021, doi: 10.20895/INISTA.V3.

[6] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, "RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12343 LNCS, pp. 503–515, 2020, doi: 10.1007/978-3-030-62008-0_35.

[7] A. Supian, B. T. Revaldo, N. Marhadi, - Rahmaddeni, and L. Efrizoni, "Penerapan SVM dan Word2Vec untuk Analisis Sentimen Ulasan Pengguna Aplikasi DANA ," *J. Ilm. Komputasi*, vol. 23, no. 3 SE-Sistem Informasi, p. 4, Sep. 2024, doi: 10.32409/jikstik.23.3.3642.

[8] E. Damayanti, A. V. Vitianingsih, S. Kacung, H. Suhartoyo, and A. Lidya Maukar, "Sentiment Analysis of Alfagift Application User Reviews Using Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) Methods," *Decod. J. Pendidik. Teknol. Inf.*, vol. 4, no. 2, pp. 509–521, 2024, doi: 10.51454/decode.v4i2.478.

[9] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," *Comput. Linguist.*, vol. 37, pp. 267–307, Jun. 2011, doi: 10.1162/COLI_a_00049.

[10] G. R. Putri, M. A. Maulana, and S. Bahri, "Perbandingan Algoritma Naïve Bayes dan TextBlob Untuk Mendapatkan Analisis Sentimen Masyarakat Pada Sosial Media," *Teknika*, vol. 13, no. 2, pp. 213–218, 2024, doi: 10.34148/teknika.v13i2.815.

[11] N. Wiliani, N. L. Chusna, and P. B. Ramadhan, *Analisis Sentimen terhadap Pro Kontra Aksi Unjuk Rasa Mahasiswa dengan Naïve Bayes dan Information Gain*. Penerbit NEM, 2023. [Online]. Available: https://books.google.co.id/books?id=RvTUEAAAQBAJ

[12] D. Jollyta, A. Hajjah, E. Haerani, and M. Siddik, *Algoritma Klasifikasi untuk Pemula Solusi Python dan RapidMiner*. Deepublish, 2023. [Online]. Available: https://books.google.co.id/books?id=jVjcEAAAQBAJ

[13] C. M. Putri, M. Afdal, R. Novita, and M. Mustakim, "Perbandingan Evaluasi Kernel Support Vector Machine dalam Analisis Sentimen Chatbot AI pada Ulasan Google Play Store," *J. Teknol. Sist. Inf. dan Apl.*, vol. 7, no. 3, pp. 1236–1245, 2024, doi: 10.32493/jtsi.v7i3.41354.

[14] A. H. Nurdy and A. Rahim, "Analisis Sentimen Ulasan Game Stumble Guys Pada Playstore Menggunakan Algoritma Naïve Bayes Sentiment Analysis of Stumble Guys Game Reviews on Playstore Using the Naïve Bayes Algorithm," vol. 13, no. November, pp. 388–395, 2024, doi: 10.34148/teknika.v13i3.993.

[15] I. Tahyudin, *Pengenalan Machine Learning Menggunakan Jupyter Notebook*. in Mechine Learning. Zahira Media Publisher, 2020. [Online]. Available: https://books.google.co.id/books?id=_uMREAAAQBAJ

[16] W. S. Lestari and Y. M. Saragih, "Comparison of Deep Neural Networks and Random Forest Algorithms for Multiclass Stunting Prediction in Toddlers," vol. 13, no. November, pp. 412–417, 2024, doi: 10.34148/teknika.v13i3.1063.

[17] D. A. Kristiyanti and Sri Hardani, "Sentiment Analysis of Public Acceptance of Covid-19 Vaccines Types in Indonesia using Naïve Bayes, Support Vector Machine, and Long Short-Term Memory (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 3, pp. 722–732, 2023, doi: 10.29207/resti.v7i3.4737.