Sentiment Analysis of Alight Motion Reviews in Indonesian Using Random Forest

Wahyu Andikaa), Khoirunnisa Alfita Hikmahb), Putri Nayla Sabric), Herdiana Dwi Maharanid), and Hardianto Wibowoe)

*Department of Engineering*, *Universitas Muhammadiyah Malang, Malang, Indonesia*

e) Corresponding author: [ardi@umm.ac.id](mailto:ardi@umm.ac.id)

a) [idikach30@gmail.com](mailto:idikach30@gmail.com)

b) [khoirunnisa18@gmail.com](mailto:khoirunnisa18@gmail.com)

c) [putrinaylasabrii@gmail.com](mailto:putrinaylasabrii@gmail.com)

d) [rhrdianarani@gmail.com](mailto:rhrdianarani@gmail.com)

**Abstract.** Digital technology in the creative industry has brought many conveniences and innovations, one of which is in video editing which can now be done through smartphones with applications such as Alight Motion. This app features animation, motion graphics, and visual effects that make it easy for users. However, although Alight Motion has provided various features, some problems are still experienced by users, as seen from the 3.6 rating from 997,000 users on the Google Play Store. To better understand user feedback, this study conducted sentiment analysis using the Random Forest algorithm, classifying user reviews into positive, negative, and neutral. The process involves collecting review data, pre-processing such as data cleaning, case folding, normalisation, tokenization, and stemming, and data labeling before analysis. Model validation with K-Fold Cross Validation showed high accuracy results, with an average accuracy of 93.54% and the best accuracy of 97%. The data visualization with a word cloud shows frequent words such as "bagus", "aplikasi", "motion", "video", and "perbaiki". The results of this analysis can help in product development and more effective marketing strategies based on user sentiment.

**Keywords:** Alight Motion, User Reviews, Random Forest Algorithm, Sentiment

# INTRODUCTION

Digital technology in the creative industry has brought many important conveniences and innovations. Initially, video editing could only be done on a computer, but with the development of technology, many smartphone applications now offer applications to perform the video editing process. With this, video creation is now much more flexible. This allows more people to be creative. Alight Motion is one of the first professional Motion graphics apps that bring animation, motion graphics, visual effects, video editing, and more.

Although Alight Motion has provided various features to help users, some problems are still experienced by some users. From the results of the assessment of the Alight motion application on the Google Play Store, this application gets a rating of around 3.6 from approximately 997,000 users compared to other video editing applications which have an average rating above 4 in PlayStore. Alight motion video editing app has the lowest rating. To conclude more positive or negative responses from the rating. So, sentiment analysis is needed to understand the user response to the Alight Motion application.

Sentiment analysis is a process of evaluating a digital text to find out whether the sentiment categorization is classified as positive, negative or neutral with the aim of knowing the expression of each phrase or word in it [1]. In analyzing a sentiment, a method is needed in this research, Random Forest is a machine learning algorithm used for data classification. This algorithm functions by combining several randomly generated decision trees to generate predictions [2].

The research related to sentiment analysis using random forest was conducted by Chandra Ayunda Apta Soemedhy et al [3] with the title “Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest” providing 84% model accuracy results. Further research was conducted by Evita Fitri et al [4]. With the title of his research “Analisis sentimen terhadap aplikasi Ruang guru menggunakan algoritma Naïve Bayes, Random Forest dan Support Vector Machine” with each of the results of Naïve bayes accuracy obtaining 94.16% accuracy, Random forest 97.16%, Support Vector Machine 96.01% of the three results random Forest has better accuracy than the other two algorithm models. Furthermore, the research conducted by Natalya Br sidauruk et al [5] with the title “Penggunaan metode SVM dan random forest untuk analisis sentimen ulasan pengguna terhadap KAI Access di google play store” with the results of 97% SVM accuracy and 93% random forest. Further research was conducted by Gilbert Jeffson Sagala et al [6] with the title "Sentiment Analysis on ChatGPT Application Reviews on Google Play Store Using Random Forest, Support Vector Machine and Naïve Bayes Algorithms" with each random forest accuracy producing the same accuracy between random forest and svm of 91%. for naive bayes as the lowest accuracy of 90%. Furthermore, for previous research from Khuswatun Hasanah [7] with the title “Comparison of Sentiment Analysis Models for Shopee Comments on Google Play Store” in this study has several models used with the results of SVM accuracy with 88% accuracy, Logistic regression with 85% accuracy, Random forest with 85% accuracy, K-NN with 83% accuracy, Naive bayes with 78% accuracy.

Through this research, sentiment analysis will help better understand what makes users give positive, negative or neutral reviews of the app. By using the Random Forest method, which has proven effective in previous studies, this research can provide valuable information for product development and more effective marketing strategies. A deeper understanding of user sentiment can also help Alight Motion's developers to improve the app's features, fix existing issues, and design a better user experience.

# RESEARCH METHODS

The research methodology is a flow of stages carried out during the research process to ensure the research objectives are achieved. The following **FIGURE 1** is the flow of methodology used in this research:



**FIGURE 1***. Research phases*

## Problem Identification

This research was conducted based on several problems, including the following:

1. To find out the sentiments of Alight Motion application users. The results will be three groups of positive, neutral, and negative sentiment categorization.
2. To find the highest accuracy that can be generated by random forest.
3. To find words that have the frequency of occurrence of the categorization results for each positive, neutral and negative.

## Data Collection

Data collection is the first step before performing sentiment analysis. In this context, it involves collecting Alight Motion app reviews from the Google Play Store. Tools such as Google Colab can be used for web scraping to automate the extraction of information from websites, eliminating the need for manual copying. This automated technique is essential for collecting sentiment data to be analyzed [8].

## Pre-Processing data

The data that has been collected still cannot be used because there needs to be filtering in it. The function of this data cleaning is to avoid duplicate data, emojis, mentions, hashtags... In this stage, a process is carried out including data cleaning, Case Folding, Word Normalization, Tokenizing, Stemming.

1. Data cleaning

Data cleaning is the process of cleaning raw data that has been obtained through the data collection process.  Elements that will later be removed duplicate data, special characters, hashtags, emojis, and other symbols [9]. This aims to facilitate the process in the next stage. This step helps in equalizing the text format so that there is no difference between uppercase and lowercase letters that can affect the analysis results.

1. Case Folding

Case Folding is the process of converting capital letters into lowercase letters in text [10]. This helps when later doing word normalization and Stopword.

1. Word Normalization

Word normalization is the process of correcting the writing of words that are wrong or have typos [11]. For example, the word “mksh” will be normalized to “terima kasih”. This is important to ensure that all words used in the analysis have a consistent form and can be recognized by the system.

1. Stopwords removal

Stopwords are the process of removing words that are not important in the sentence. Words such as “dan”, “yang”, and “di” often do not add value to the analysis and can be removed to focus on more important words [12]. Removing stopwords helps to reduce noise in the data and improve analysis efficiency.

1. Tokenization

Tokenization is the process of converting a sentence into a series of words stored in an array in the form of a string [13]. This process enables more detailed analysis at the word level, which is important for various applications such as text analysis and machine learning. With tokenization, we can separate words in a sentence and analyze them individually

1. Stemming

Stemming is the process of removing affixes from words so that they become the basic form of the word [14]. Thus, the analysis becomes simpler and more focused on the core meaning of the words. However, there are words that must be considered again because in the stemming process there are several words with affixes that cannot be cleaned.

## Labelling

In the data labeling stage, the data will be divided into three categories: positive, neutral, and negative. Before doing the labeling process, the data will be translated into English. This translation process is carried out using the deep translator library which supports various translation functions needed for further analysis. To perform the labeling process using the Lexicon-Based method in this method will measure based on the sentiment score if 1 then it will be considered positive, then 0 will be considered neutral and -1 will be considered as a negative sentiment [15].

## Random Forest Classification

Random Forest classification is one of the data classification processes. This algorithm uses the ensemble method, which combines several models to create one optimal prediction model. The Random Forest algorithm operates by dividing data into several individual decision trees [16].

(1)

## Evaluating Model

Evaluating model is essential to assess its performance in accurately predicting the right class or label for each data. The data will be divided into two stages: data splitting and cross-fold validation. During the data splitting stage, the review data is separated into training and testing sets. In the cross-fold validation stage, the data is randomly divided into 10 parts [17]. Common evaluation metrics for classification models include:

1. Accuracy: The percentage to test an algorithm between the true value and the predicted value made by the model [18].
2. Precision: an algorithm used to compare the correct data from the available data [19].
3. Recall: a process used to compare the correct data with all the data retrieved or not retrieved by the system [19].
4. F1-Score: A metric that combines precision and recall to measure the balance between them, providing a single score that reflects both aspects [20].

## Word Visualization

This process uses a word cloud to find out which words often appear, the larger the size of the word indicates that the word is often written by users [21].

# RESULT AND DISCUSSION

## Web Scraping

This research collects data from the latest Alight Motion application reviews, using web scraping techniques to generate 1000 data. in **FIGURE 2** This technique involves the help of Google Colab tools with the URL com.alightcreative.motion from the Google Play Store site. This research only targets the latest data and is based on reviews in Indonesian.

A screenshot of a computer

Description automatically generated

**FIGURE 2.**Alight motion user review dataset (data search is written in Indonesian)

## Pre-Processing

In this research, data cleaning on the raw data obtained through web scraping will be cleaned by removing irrelevant elements. This process involves filtering out duplicate data, hashtags, numeric values, emojis, and other extraneous characters or symbols. In **TABLE 1**, the yellow text that will be cleaned or deleted leaves 1000 review data that is successfully processed and forwarded to the next stage.

**TABLE 1.** Cleaning data (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | aplikasi ini memang bagus saat untuk mengedit tapi saya ingin protes karena saat mengekspor video sudah 95% malah keluar tanpa sebab tolong segera di perbaikiðŸ™ðŸ¼ðŸ™ðŸ¼ ya apk sialan |
| 2 | Tolong masalah ekspor di benerin dulu soalnya kalau mau ekspor tu susah banget kaya bakal lama harus di coba 4x baru bisa dan satu lagi biasanya kalau lagi ngedit tiba tiba force close |
| 3 | Aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sihðŸ˜¡. Malah kelamaan Mau ekspor. Jadi tolong update lagi. Gimana caranya biar GK ada iklannya ðŸ¥ºðŸ™ðŸ™ |

Case folding, a step in pre-processing involves converting all letters in the text into lowercase letters from capital letters. This process is illustrated in **TABLE 2**, where the initial text, “Aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sih. Malah kelamaan Mau ekspor. Jadi tolong update lagi. Gimana caranya biar GK ada iklannya” was modified to “aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sih. malah kelamaan mau ekspor. jadi tolong update lagi. gimana caranya biar gk ada iklannya.”

**TABLE 2.** Case folding data (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | aplikasi ini memang bagus saat untuk mengedit tapi saya ingin protes karena saat mengekspor video sudah malah keluar tanpa sebab tolong segera di perbaiki ya apk sialan |
| 2 | Tolong masalah ekspor di benerin dulu soalnya kalau mau ekspor tu susah banget kaya bakal lama harus di coba x baru bisa dan satu lagi biasanya kalau lagi ngedit tiba tiba force close |
| 3 | Aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sih. Malah kelamaan Mau ekspor. Jadi tolong update lagi. Gimana caranya biar GK ada iklannya |

Word normalization is the process of converting words contained in sentences into basic words. This process aims to correct words that still have errors in writing, so that the words return to their basic form. An example of this process can be seen in **TABLE 3** in the document, where words like “apk” are changed to “aplikasi”.

**TABLE 3.** Word normalization data (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | aplikasi ini memang bagus saat untuk mengedit tapi saya ingin protes karena saat mengekspor video sudah malah keluar tanpa sebab tolong segera di perbaiki ya apk sialan |
| 2 | tolong masalah ekspor di benerin dulu soalnya kalau mau ekspor tu susah banget kaya bakal lama harus di coba x baru bisa dan satu lagi biasanya kalau lagi ngedit tiba tiba force close |
| 3 | aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sih. malah kelamaan mau ekspor. jadi tolong update lagi. gimana caranya biar gk ada iklannya |

Stopwords in this context are used to remove certain words that have no significance, such as conjunctions, question words, and other unimportant terms. in **TABLE 4** Examples of words that will be removed include “gimana”, “dan”, and “kalau”. In addition, words such as “kaya”, “tu”, and “but” will also be removed as they do not provide meaningful information.

**TABLE 4.** Stopwords data (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | aplikasi ini memang bagus saat untuk mengedit tapi saya ingin protes karena saat mengekspor video sudah malah keluar tanpa sebab tolong segera di perbaiki ya aplikasi sialan. |
| 2 | tolong masalah ekspor di benerin dulu soalnya kalau mau ekspor tu susah banget kaya bakal lama harus di coba x baru bisa dan satu lagi biasanya kalau lagi ngedit tiba tiba force close |
| 3 | aplikasi nya bagus sih tapi pas lagi di ekspor kenapa ada iklannya juga sih. malah kelamaan mau ekspor. jadi tolong update lagi. gimana caranya biar tidak ada iklannya |

Tokenization, in this process, involves parsing the words in a review sentence. In addition, this process also removes certain characters such as punctuation marks and filters based on text length. In this context, tokenization results in the formation of a one-dimensional array, as illustrated in **TABLE 5**.

**TABLE 5.** Tokenization (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | [‘ bagus’  ,’mengedit’, ‘protes’, ‘mengekspor’, ‘video’ ,’keluar’, ‘tanpa’, ‘perbaiki’,‘aplikasi’, ‘sialan’] |
| 2 | [‘tolong’, ‘masalah’ ,‘ekspor’ ,‘benerin’ ,‘ekspor’ ,‘banget’  ,‘lama’ ,‘coba’ ,‘baru’ ,‘bisa’ ‘satu’ ,‘ngedit’ ,‘tiba’ ,‘tiba’ ,‘force’ ,‘close’] |
| 3 | [‘Aplikasi’,‘bagus’,‘ekspor’,‘iklannya’,‘kelamaan’,‘ekspor’,‘update’,‘tidak’,‘ada’, ‘iklannya’ ] |

Word stemming is the conversion of words in sentences into their basic form. This process aims at removing or eliminating affixes such as “se-”, “di-”, “per-”, “-an”, and others, thus returning the word to its basic form. An example of this procedure is illustrated in **TABLE 6** of the document, where words such as “mengedit” are simplified to “edit”.

**TABLE 6.** Stemming data (data search is written in Indonesian)

|  |  |
| --- | --- |
| **No** | **Text** |
| 1 | [‘ bagus’  ,’mengedit’, ‘protes’, ‘mengekspor’, ‘video’ ,’keluar’, ‘tanpa’, ‘perbaiki’,‘aplikasi’, ‘sialan’] |
| 2 | [‘tolong’, ‘masalah’ ,‘ekspor’ ,‘benerin’ ,‘ekspor’ ,‘banget’  ,‘lama’ ,‘coba’ ,‘baru’ ,‘bisa’ ‘satu’ ,‘ngedit’ ,‘tiba’ ,‘tiba’ ,‘force’ ,‘close’] |
| 3 | [‘Aplikasi’,‘bagus’,‘ekspor’,‘iklannya’,‘kelamaan’,‘ekspor’,‘update’,‘tidak’,‘ada’, ‘iklannya’ ] |

## Labeling data

Before labeling the data, it is necessary to translate the sentences into English. This translation is needed so that it can be used at the time of labeling. in **TABLE 7** the translation results using deep translator from Indonesian into English. divided into 3 parts positive, negative and neutral.

**TABLE 7.** Translation on data

|  |  |  |
| --- | --- | --- |
| **No** | **Text** | **Category** |
| 1 | The application is good, editing is really good, there are lots of export features even though I just installed the application, this is helpful | Positive |
| 2 | give stars, the application supports being an editor for several paid effects and then saving videos and paying for updates | Neutral |
| 3 | little problem exporting the project likes to lag | Negative |

In **TABLE 8** of the 1000 datasets, the detailed results of the number of each labeling category are explained. From the table it is found that the positive category has 666 sentiments. The neutral category has 109. The negative category has 225.

**TABLE 8.** Labelling on data

|  |  |
| --- | --- |
| **Category** | **Total** |
| positive | 666 |
| neutral | 109 |
| negative | 225 |

## Validation of Random forest

At this stage, described in **TABLE 9**, the validation process will be carried out using K-Fold Cross Validation to ensure accuracy. This process is repeated from K-Fold 1 to K-Fold 10. The results of this iteration will be presented in the table below.

**TABLE 9.** Validation results for each K-Fold

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **K-Fold** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| K-Fold 1 | 93% | 93% | 93% | 92% |
| K-Fold 2 | 96% | 96% | 96% | 95% |
| K-Fold 3 | 92% | 92% | 92% | 92% |
| K-Fold 4 | 95% | 95% | 95% | 94% |
| K-Fold 5 | 91% | 92% | 91% | 91% |
| K-Fold 6 | 92% | 92% | 92% | 91% |
| K-Fold 7 | 97% | 97% | 97% | 97% |
| K-Fold 8 | 94% | 94% | 94% | 94% |
| K-Fold 9 | 91% | 92% | 91% | 91% |
| K-Fold 10 | 91% | 91% | 91% | 91% |

Through the implementation of validation on the Random Forest Algorithm, it was found that the best results were achieved with a division of K-7, which resulted in an accuracy of 97%. The average values for Accuracy, Precision, Recall, and F1-Score were also evaluated. From the **TABLE 10**, the average accuracy is 93.54%, average precision is 93.76%, average recall is 93.54%, average F1-Score is 93.44%.

**TABLE 10.** Average Classification Report Random Forest

|  |  |
| --- | --- |
| Average Accuracy | 93.54% |
| Average Precision | 93.76% |
| Average Recall | 93.54% |
| Average F1-Score | 93.44% |

## Data Visualization

In the process of visualizing a word that appears from each positive, negative, and neutral sentiment using the word cloud tool. This tool helps to determine the keywords that are often mentioned by Alight Motion users. From the results of the overall sentiment visualization in **FIGURE 3** shown. words that often appear include “bagus”, “aplikasi”, “motion”, “video”, and “perbaiki”.

A close up of words

Description automatically generated

**FIGURE 3.**Visualization of frequently occurring words (data search is written in Indonesian)

# CONCLUSION

This research successfully conducted sentiment analysis of user reviews of the Alight Motion application on the Google Play Store using the Random Forest method. Review data was collected using web scraping techniques from the Google Play Store with the help of Google Colab, resulting in 1000 recent review data. The data pre-processing stage involved data cleaning, case folding, word normalisation, stopwords removal, tokenization, and stemming. The review data was translated into English using a deep translator to facilitate the labelling process and categorised into three sentiment types: positive (666 reviews), neutral (109 reviews), and negative (225 reviews). Random Forest the algorithm was used to classify user reviews and model validation was performed using K-Fold Cross Validation (K=10), showing the the best accuracy of 97% at split K-7. Word clouds were used for the Visualization of frequently occurring words in each sentiment category, with words such as "bagus", "aplikasi", "motion", "video", and "perbaiki" appearing frequently, indicating the aspects most talked about by users.

From the results of the analysis, it can be concluded that The Alight Motion app received a lot of positive reviews, but there were also several negative and neutral reviews indicating issues that need to be fixed by the app developers. This sentiment analysis can help with decision making for a more effective product development and marketing strategies, with the Random Forest method proves effective in classifying sentiment with high accuracy, providing reliable results for understanding users' views and feelings towards this app.

# ACKNOWLEDGMENT

First and foremost, I would like to express my deepest gratitude to Allah SWT for His endless blessings and guidance, without which this research would not have been possible. I am sincerely grateful to Mr. Hardianto Wibowo, Lecturer in the Department of Informatics Engineering, for his invaluable guidance, support, and dedication throughout the course of this research. His expertise and insightful feedback have significantly shaped the direction and outcome of this study. I also extend my thanks to the Department of Informatics Engineering for providing the essential resources, facilities, and a conducive academic environment that made this research possible. On a personal note, I am profoundly grateful to my parents for their endless support throughout this journey. I would also like to thank my friends, who provided constant motivation and understanding during challenging times.

# REFERENCES

1. Audrey, O., D.E. Ratnawati, and I. Arwani, *Analisis Sentimen Pengguna Twitter Terhadap Opini Non Fungible Token di Indonesia Menggunakan Algoritma Random Forest Classifier.* Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2022. **6**(12): p. 5889-5897.

2. Bahrawi, N., *Sentiment analysis using random forest algorithm-online social media based.* Journal of Information Technology and Its Utilization, 2019. **2**(2): p. 29-33.

3. Larasati, F.A., D.E. Ratnawati, and B.T. Hanggara, *Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest.* Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2022. **6**(9): p. 4305-4313.

4. Fitri, E., *Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine.* Jurnal Transformatika, 2020. **18**(1): p. 71-80.

5. Sidauruk, N., N. Riza, and R.N.S. Fatonah, *PENGGUNAAN METODE SVM DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA TERHADAP KAI ACCESS DI GOOGLE PLAYSTORE.* JATI (Jurnal Mahasiswa Teknik Informatika), 2023. **7**(3): p. 1901-1906.

6. Sagala, G.J. and Y.T. Samuel, *Sentiment Analysis on ChatGPT App Reviews on Google Play Store Using Random Forest Algorithm, Support Vector Machine and Naïve Bayes.* International Journal of Engineering Business and Social Science, 2024. **2**(04): p. 1194-1204.

7. Hasanah, K., *Comparison of Sentiment Analysis Model for Shopee Comments on Google Play Store.* Jurnal Sisfokom (Sistem Informasi dan Komputer), 2024. **13**(1): p. 21-30.

8. Rudini, D., D.G. Purnama, and A.A. Khan, *Penggunaan Teknik Web Scraping dalam Aplikasi Pengambilan Data dari Google Maps untuk Menunjang Digital Marketing.* Lentera: Multidisciplinary Studies, 2023. **2**(1): p. 10-19.

9. Muliawan, J. and E. Dazki, *Sentiment Analysis of Indonesia's Capital City Relocation Using Three Algorithms: Naïve Bayes, KNN, and Random Forest.* Jurnal Teknik Informatika (JUTIF), 2023. **4**(5): p. 1227-1236.

10. Fitriyana, V., et al., *Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine.* Jurnal Buana Informatika, 2023. **14**(01): p. 40-49.

11. Adhan, S.N., et al., *ANALISIS SENTIMEN ULASAN APLIKASI WATTPAD DI GOOGLE PLAY STORE DENGAN METODE RANDOM FOREST.* AnoaTIK: Jurnal Teknologi Informasi dan Komputer, 2024. **2**(1): p. 6-15.

12. Putra, A.D.A. and S. Juanita, *Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN.* JATISI (Jurnal Teknik Informatika dan Sistem Informasi), 2021. **8**(2): p. 636-646.

13. Naf'an, M.Z., et al., *Sentiment analysis of cyberbullying on instagram user comments.* Journal of Data Science and Its Applications, 2019. **2**(1): p. 38-48.

14. Aufa, M.J. and A. Qoiriah, *Analisis Sentimen Pengguna Platform Belajar Online Coursera menggunakan Random Forest dengan Metode Ekstraksi Fitur Word2vec.* Journal of Informatics and Computer Science (JINACS), 2022: p. 244-255.

15. Manullang, O., C. Prianto, and N.H. Harani, *Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based Dan Random Forest.* Jurnal Ilmiah Informatika, 2023. **11**(02): p. 159-169.

16. Warsito, B. and A. Prahutama. *Sentiment analysis on tokopedia product online reviews using random forest method*. in *E3S Web of Conferences*. 2020. EDP Sciences.

17. Doloksaribu, H.P. and Y.T. Samuel, *Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi.* Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika, 2022. **16**(1): p. 1-11.

18. Jihad, M.A.A., A. Adiwijaya, and W. Astut, *Analisis sentimen terhadap ulasan film menggunakan algoritma random forest.* eProceedings of Engineering, 2021. **8**(5).

19. Yuniar, E., D.S. Utsalinah, and D. Wahyuningsih, *Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompet Digital dengan Menggunakan Algoritma Machine Learning.* Jurnal Janitra Informatika dan Sistem Informasi, 2022. **2**(1): p. 35-42.

20. Herjanto, M.F.Y. and C. Carudin, *Analisis Sentimen Ulasan Pengguna Aplikasi Sirekap Pada Play Store Menggunakan Algoritma Random Forest Classifer.* Jurnal Informatika dan Teknik Elektro Terapan, 2024. **12**(2).

21. Soemedhy, C.A.A., et al., *Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube “Kekerasan Seksual”).* Jurnal Informatika: Jurnal Pengembangan IT, 2022. **7**(2): p. 80-84.