Sentiment Analysis Of Gobis Suroboyo Bus Application Feature Using Support Vector Machinery (SVM) Algorithm

Avia Arista
Informatika, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur
Surabaya, Indonesia
19081010077@student.upn.jatim.ac.id

Abstract— **Congestion is something that often occurs in several big cities in Indonesia. According to a report from INRIX, a company that conducts data analysis and provides traffic data, in 2021 the city of Surabaya is listed as the 41st most congested city in the world, with a total wasted time of 62 hours in one year. This is of course very detrimental from various aspects, because of this the Surabaya Transportation Agency developed an application called GOBIS Suroboyo Bus to support digitizing the use of Surobyo Bus to break up traffic jams. The purpose of this study is to analyze public sentiment that has used the features in the GOBIS Suroboyo Bus application by using the application comment column on the Google Playstore to obtain data. The method used in this study is Complement Naïve Bayes which produces an accuracy value of 76.00% and Support Vector Machines which obtains an accuracy value of 73.00%.**

Keywords— kemacetan, GOBIS, analisis sentimen, naïve bayes, support vector machines

I. Introduction

Congestion is a problem that is often found in several big cities in Indonesia, including in the city of Surabaya. Congestion occurs because the number of private vehicles exceeds road capacity and people prefer private vehicles to public transportation. The impact of congestion includes increased travel time and wasted fuel. The Global Traffic Scorecard survey in 2021 shows that the City of Surabaya is ranked 41st with 62 hours of user time being wasted in a year due to being stuck in traffic jams [1]. To overcome the problem of congestion, the Surabaya City Government is developing the Suroboyo Bus public transportation in the hope of reducing congestion and attracting people to use public transportation. Suroboyo Bus has three main routes and uses a unique payment system by exchanging plastic bottles or cups into points to use as a payment method. The Surabaya City Government has also developed the GOBIS Suroboyo Bus mobile application to make it easier for users to find out the position of the bus and its arrival time. Although there are some complaints from users. [2]

A mobile application called GOBIS Suroboyo Bus (Golek Bis) has been developed by the city government to provide information about Suroboyo Bus to users via an internet connection such as 2G/3G/4G/5G/WiFi. One of the advantages of this application is that it makes it easier for users to track the position of the bus and find out the arrival time of the bus at the selected stop [3] However, since its launch, this application has received several complaints from users, including limited routes which are only available on main roads, difficulties exchange bottles for points, and a few other issues. [4]

This study aims to analyze public sentiment/opinion towards the GOBIS Suroboyo Bus application service by taking the example of the Halodoc application, besides that it also aims to determine the level of accuracy of the Support Vector Machines (SVM) algorithm.

II. Methodology

The method used in this study is the sentiment analysis method with the Support Vector Machines classification algorithm. SVM is a probabilistic and statistical technique used for classification. SVM has proven to be very fast when applied to large amounts of data. The stages of the research process carried out in the research will be explained in the following stages:

1. Data Collection

2. Data Labeling

3. Preprocessing Stage

3.1. Case Folding

3.2. cleansing

3.3. tokenization

3.4. Normalization

3.5. Stopwords

3.6. Stemming

4. Word Weighting

5. Distribution of training data and test data

6. Classification of Sentiment Analysis, and

7. Evaluation of Sentiment Analysis Results

1. Data Collection

Data collection was taken directly from the application comments column on the Google Play Store by utilizing the Google Play Scrapper module in the Python programming language. The user comment data that will be used in this study has a range of years from 2018 to April 2023.

2. Data Labeling

After all the data has been pre-processed, the next process is to label all tweet data into 2 classes, namely the positive class and the negative class. Where the positive class means that the tweet contains positive user reviews after using the GOBIS Suroboyo Bus application features. As for the negative class, it means user dissatisfaction with the GOBIS Suroboyo Bus application features.

3. Pre Processing Data

Pre-processing data is a process of cleaning data that has been labeled with sentiment class previously. What is meant by cleaning data here is the process of removing some attributes that do not affect the results of sentiment analysis later such as emoticons, hashtags, and also capital letters.

3.1. Case Folding

In writing tweets, users often use sentences that are not standard or use various types of words. So that the tweet data needs to be equated with the case folding process which later the tweet sentence will be converted into lowercase as a whole [5].

3.2. cleansing

In the tweet data cleaning process, the process of removing some attributes or elements that are not needed such as URLs, tags (#), emoticons and other characters will be carried out [6].

3.3. tokenization

Tokenaziation is data pre-processing which will produce tweet data text snippets in words. Where each word will later be separated using punctuation marks such as commas (,), or periods (.) [7].

3.4. Normalization

Is a process to normalize words that are not standard or commonly known as slang / slang into a more standard language by using a dictionary that has been prepared. [8]

3.5. Stopwords

It is a process of eliminating or deleting words that are not needed or in other words will not have an impact on the results of the analysis [9].

3.6. Stemming

It is pre-processing data by changing the words in the tweet data that have word affixes to their basic word forms, this aims to facilitate the analysis process later [10].

4. Word Weighting

Word weighting or the Word Weighting process is the process of assigning feature weights to each word within the weight range of each term, namely 0-1, with the aim of avoiding numerical difficulties during the calculation process. This procedure is known as TF-IDF, which will later calculate the Term Frequency (TF) and Inverse Document Frequency (IDF) values for each token (word) [11].

5. Distribution of training data and test data

After the data has passed the pre-processing, the next process is to divide the data into training data and test data. Where the training data itself is data that will later be used in the algorithm training process that has been prepared and the test data itself is data that will be used to test algorithms that have carried out the previous training process and also to measure the level of accuracy of the algorithm used.

6. Classification of Sentiment Analysis

The algorithm used in this study is Support Vector Machine (SVM) which is a classification method in data mining that uses accurate non-linear mapping to higher dimensions, data from two categories can be distinguished using the hyperplane [12].

7. Evaluation of Sentiment Analysis Results

The last process after the classification is done using Multinominal Naïve Bayes, then the next calculation is done using the Confusion Matrix. Confusion Matrix is one of the methods for testing an object and distinguishing which calculations are correct and which are wrong. In the calculation process using the Confusion Matrix, four calculations are carried out, namely Accuracy, Precision, Recall, and Specificity (f1-score)

III. Result and discussion

In this study, the data collection process was carried out using a crawling technique to review the GOBIS Suroboyo Bus application, which was further processed in pre-processing data. In the crawling process, data were collected from 600 user reviews of the GOBIS Suroboyo Bus application, which ranged from 2018 to 2023..

1. Data Collection

In the process of collecting research data, the authors take advantage of public comments after using the features in the GOBIS Suroboyo Bus application. The data collection process uses the Google Play Scrapper module with a total of 600 data. Meanwhile, an overview of the results of data collection can be seen in the image below:

no.	Content
1	Parah sih wirawiri suroboyo dengan nomor ww01, barusan ditinggal udah dikasih tanda stop udah nunggu dihalte tandes dari tadi malah gk berhenti malah langsung ditinggal padahal masih kosong didalam parah parah sumpah
2	Ini apk buat penipuan bro dengan modus baru mana ada gobiz suroboyo bus
3	Makin hari makin goodd beda kaya operator bus semangka
4	Ada penumpang di halte malah lanjut terus ga berhenti, niat apa ngga?
5	Sangat membantu orang yg tdk punya kendaraan
	…
596	Sipp banget.udah berkali kali naik bus ini.mungkin bus nya harus d tambah lg 😘
597	Pls untuk ditambahkan posisi bus biar kita gk banyak buang waktu nunggu bus
598	Hari sabtu 23 juni 2018, pukul 18.10 , saya tiba di pemberangkatan bus kota di terminal purabaya. Lalu saya tanya ke pegawai dishub yang ada di pos, nunggu di sebelah mana kalau mau naik suroboyo bus? Jawabnya : kalau jam segini, suroboyo bus udah habis. Akhirnya saya naik bus kota. Yang sayamau tanyakan, suroboyo bus beroperasi sampai jam berapa? Trus di terminal purabaya, nunggu suroboyo bus di sebelah mana? Makasih.
599	Cucok deh hmm 😂 😂
600	Sesuai dengan harapan 💛 maju terus suroboyo

2. Data Labeling

After the data is obtained, then the next step is the process of determining the label of each tweet sentence into a positive class and also a negative class. Where later this will be useful when the training process and testing the algorithm used in this study. Where is the sentiment class manually divided into two sentiment classes namely positive which is represented in number 0 and negative which is represented in number 1.

no.	content_Stemmed	sentiment
1	['parah', 'sih', 'bolak balik', 'surabaya surabaya', 'nomor', 'saya', 'barusan', 'tinggal', 'sudah', 'kasih', 'tanda', 'stop', 'sudah', 'nunggu', 'halte', 'tandes', 'tidak', 'henti', 'langsung', 'tinggal', 'kosong', 'dalam', 'parah', 'parah', 'sumpah']	1
2	['aplikasi', 'tipu', 'teman', 'modus', 'gobiz', 'surabaya surabaya', 'bus']	1
	…	…
598	['sabtu', 'juni', 'berangkat', 'bus', 'kota', 'terminal', 'purabaya', 'pegawai', 'dinas hubung', 'pos', 'nunggu', 'belah', 'surabaya surabaya', 'bus', 'jam', 'gin', 'surabaya surabaya', 'bus', 'sudah', 'habis', 'bus', 'kota', 'surabaya surabaya', 'bus', 'operasi', 'jam', 'terminal', 'purabaya', 'nunggu', 'surabaya surabaya', 'bus', 'belah', 'makasih']	1
599	['cocok', 'iya', 'keluh']	0
600	['sesuai', 'harap', 'maju', 'surabaya surabaya']	0

3. Pre Processing Data

After the data needed in the research has been obtained, the next process is cleaning the data to make it easier in the classification process later. Where for pre-processing there are several steps, including:

3.1. Case Folding

Pre-processing case folding data in this study uses the lower module in the python programming language, so that all letters in the tweet sentence will turn into lowercase. The results can be seen in the following image:

no.	Content_folding
1	parah sih wirawiri suroboyo dengan nomor ww01, barusan ditinggal udah dikasih tanda stop udah nunggu dihalte tandes dari tadi malah gk berhenti malah langsung ditinggal padahal masih

	kosong didalam parah parah sumpah
2	ini apk buat penipuan bro dengan modus baru mana ada gobiz suroboyo bus
	…
598	hari sabtu 23 juni 2018, pukul 18.10 , saya tiba di pemberangkatan bus kota
600	di terminal purabaya. lalu saya tanya ke pegawai dishub yang ada di pos,

3.2. Cleaning

The pre-processing of the cleaning stage will produce core tweet sentences after other unneeded attributes are removed. This process removes emoticons, punctuation marks, as well as like attributes, among other things. The end result of cleansing is as follows:

no.	**content_cleaned**
1	parah sih wirawiri suroboyo dengan nomor ww barusan ditinggal udah dikasih tanda stop udah nunggu dihalte tandes dari tadi malah gk berhenti malah langsung ditinggal padahal masih kosong didalam parah parah sumpah
2	ini apk buat penipuan bro dengan modus baru mana ada gobiz suroboyo bus
	…
598	hari sabtu juni pukul saya tiba di pemberangkatan bus kota di terminal purabaya lalu saya tanya ke pegawai dishub yang ada di pos nunggu di sebelah mana kalau mau naik suroboyo bus jawabnya kalau jam segini suroboyo bus udah habis akhirnya saya naik bus kota yang saya mau tanyakan suroboyo bus beroperasi sampai jam berapa trus di terminal purabaya nunggu suroboyo bus di sebelah mana makasih
599	cucok deh hmm
600	sesuai dengan harapan maju terus suroboyo

3.3. tokenization

In this process, the tweet data that has been cleaned will be separated per word by using commas. The results of the tokenization process can be seen in the table below:

no.	**content_tokenize**
1	['parah', 'sih', 'wirawiri', 'suroboyo', 'dengan', 'nomor', 'ww', 'barusan', 'ditinggal', 'udah', 'dikasih', 'tanda', 'stop', 'udah', 'nunggu', 'dihalte', 'tandes', 'dari', 'tadi', 'malah', 'gk', 'berhenti', 'malah', 'langsung', 'ditinggal', 'padahal', 'masih', 'kosong', 'didalam', 'parah', 'parah', 'sumpah']
2	['ini', 'apk', 'buat', 'penipuan', 'bro', 'dengan', 'modus', 'baru', 'mana', 'ada', 'gobiz', 'suroboyo', 'bus']
	…
598	['hari', 'sabtu', 'juni', 'pukul', 'saya', 'tiba', 'di', 'pemberangkatan', 'bus', 'kota', 'di', 'terminal', 'purabaya', 'lalu', 'saya', 'tanya', 'ke', 'pegawai', 'dishub', 'yang', 'ada', 'di', 'pos', 'nunggu', 'di', 'sebelah', 'mana', 'kalau', 'mau', 'naik', 'suroboyo', 'bus', 'jawabnya', 'kalau', 'jam', 'segini', 'suroboyo', 'bus', 'udah', 'habis', 'akhirnya', 'saya', 'naik', 'bus', 'kota', 'yang', 'saya', 'mau', 'tanyakan', 'suroboyo', 'bus', 'beroperasi', 'sampai', 'jam', 'berapa', 'trus', 'di', 'terminal', 'purabaya', 'nunggu', 'suroboyo', 'bus', 'di', 'sebelah', 'mana', 'makasih']
599	['cucok', 'deh', 'hmm']
600	['sesuai', 'dengan', 'harapan', 'maju', 'terus', 'suroboyo']

3.4. Normalization

In this process, each word in the tweet sentence will be normalized because some words are not standard, so they need to be converted into standard form. The results of normalization are as follows:

no.	**content_normalized**
1	['parah', 'sih', 'bolak balik;', 'surabaya;surabaya', 'dengan', 'nomor', 'saya;', 'barusan', 'ditinggal', 'sudah;', 'dikasih', 'tanda', 'stop', 'sudah;', 'nunggu', 'dihalte', 'tandes', 'dari', 'tadi',

	'malah', 'tidak;', 'berhenti', 'malah', 'langsung', 'ditinggal', 'padahal', 'masih', 'kosong', 'didalam', 'parah', 'parah', 'sumpah']
2	['ini', 'aplikasi;', 'buat', 'penipuan', 'teman;', 'dengan', 'modus', 'baru', 'mana', 'ada', 'gobiz', 'surabaya;surabaya', 'bus']
	…
598	['hari', 'sabtu', 'juni', 'pukul', 'saya', 'tiba', 'di', 'pemberangkatan', 'bus', 'kota', 'di', 'terminal', 'purabaya', 'lalu', 'saya', 'tanya', 'ke', 'pegawai', 'dinas perhubungan;', 'yang', 'ada', 'di', 'pos', 'nunggu', 'di', 'sebelah', 'mana', 'kalau', 'mau', 'naik', 'surabaya;surabaya', 'bus', 'jawabnya', 'kalau', 'jam', 'segini', 'surabaya;surabaya', 'bus', 'sudah;', 'habis', 'akhirnya', 'saya', 'naik', 'bus', 'kota', 'yang', 'saya', 'mau', 'tanyakan', 'surabaya;surabaya', 'bus', 'beroperasi', 'sampai', 'jam', 'berapa', 'terus', 'di', 'terminal', 'purabaya', 'nunggu', 'surabaya;surabaya', 'bus', 'di', 'sebelah', 'mana', 'makasih']
599	['cocok;', 'iya;', 'mengeluh']
600	['sesuai', 'dengan', 'harapan', 'maju', 'terus', 'surabaya;surabaya']

3.5. Stopwords

In the stopword process, the process of removing existing prepositions will be carried out because it will not affect the classification results later. The results of this process are described in the image below:

no.	**content_Stemmed**
1	['parah', 'sih', 'bolak balik', 'surabaya surabaya', 'nomor', 'saya', 'barusan', 'tinggal', 'sudah', 'kasih', 'tanda', 'stop', 'sudah', 'nunggu', 'halte', 'tandes', 'tidak', 'henti', 'langsung', 'tinggal', 'kosong', 'dalam', 'parah', 'parah', 'sumpah']
2	['aplikasi', 'tipu', 'teman', 'modus', 'gobiz', 'surabaya surabaya', 'bus']
	…
598	['sabtu', 'juni', 'berangkat', 'bus', 'kota', 'terminal', 'purabaya', 'pegawai', 'dinas hubung', 'pos', 'nunggu', 'belah', 'surabaya surabaya', 'bus', 'jam', 'gin', 'surabaya surabaya', 'bus', 'sudah', 'habis', 'bus', 'kota', 'surabaya surabaya', 'bus', 'operasi', 'jam', 'terminal', 'purabaya', 'nunggu', 'surabaya surabaya', 'bus', 'belah', 'makasih']
599	['cocok', 'iya', 'keluh']
600	['sesuai', 'harap', 'maju', 'surabaya surabaya']

3.6. Stemming

In the stemming process, which is the final stage of pre-processing data mining in this study, words that have affixes, either pre, last, or middle, become basic words. So the final result of pre-processing data is as follows

no.	**content_stop_removed**
1	['parah', 'sih', 'bolak balik;', 'surabaya;surabaya', 'nomor', 'saya;', 'barusan', 'ditinggal', 'sudah;', 'dikasih', 'tanda', 'stop', 'sudah;', 'nunggu', 'dihalte', 'tandes', 'tidak;', 'berhenti', 'langsung', 'ditinggal', 'kosong', 'didalam', 'parah', 'parah', 'sumpah']
2	['aplikasi;', 'penipuan', 'teman;', 'modus', 'gobiz', 'surabaya;surabaya', 'bus']
3	['bagus;', 'beda', 'kaya', 'operator', 'bus', 'semangka']
4	['penumpang', 'halte', 'tidak;', 'berhenti', 'niat', 'tidak;']
5	['membantu', 'orang', 'yang;', 'tidak;', 'kendaraan']
	…
596	['sipp', 'bangetudah', 'berkali', 'kali', 'bus', 'inimungkin', 'bus', 'nya', 'lg']
597	['pls', 'posisi', 'bus', 'biar', 'tidak;', 'buang', 'nunggu', 'bus']
598	['sabtu', 'juni', 'pemberangkatan', 'bus', 'kota', 'terminal', 'purabaya', 'pegawai', 'dinas perhubungan;', 'pos', 'nunggu',

	'sebelah', 'surabaya;surabaya', 'bus', 'jam', 'segini', 'surabaya;surabaya', 'bus', 'sudah;', 'habis', 'bus', 'kota', 'surabaya;surabaya', 'bus', 'beroperasi', 'jam', 'terminal', 'purabaya', 'nunggu', 'surabaya;surabaya', 'bus', 'sebelah', 'makasih']
599	['cocok;', 'iya;', 'mengeluh']
600	['sesuai', 'harapan', 'maju', 'surabaya;surabaya']

4. Word Weighting
The results of the weighting process for each word in the dataset are in the float data type format to facilitate the analysis process later.

5. Distribution of training data and test data
In this study the authors divided the 600 datasets that had been pre-processed into 80% training data and also 20% test data, or in other words the training data used was 480 data and the test data used was 120 data.

6. Classification of Sentiment Analysis
After classification using the Naïve Bayes algorithm, several results were obtained including 48 data True Positive (TP), 39 data True Negative (TN), 10 data False Positive (FP) and also False Negative (FN) values 23 data. The results mentioned above can later be used as a reference for evaluating the results of the analysis using accuracy, recall, precision, and also the F1-Score.

Confusion Matrix	
48	10
23	39

7. Evaluation of Sentiment Analysis Results
In the process of evaluating the results of sentiment analysis using the Support Vector Machines (SVM) algorithm, an accuracy value of 73.00% is obtained, which is explained in the table below in more detail for each class of sentiment on the value of precision, recall, and also the F1 score. With the results that have been obtained, it can be said that the program has a fairly good accuracy value with several influencing things such as the keywords used, the dataset and the algorithm used.

	Precision	***Recall***	***F1-Score***	***support***
Negative	68%	83%	74%	58
Positive	80%	63%	70%	62

The table above explains the precision value of the positive sentiment class is 68% because only 38 data were predicted correctly and the other 20 were predicted incorrectly. Whereas the negative sentiment class gets a value of 80% which means that 49 data are predicted as True Negative and 13 others are wrong in the prediction.
Second, the positive sentiment class recall is worth 83% because only 48 data were predicted correctly and the other 10 were predicted incorrectly. Whereas the negative sentiment class gets a value of 63%, which means that 39 data are predicted as True Negative and 23 are wrong in predictions.
Third, the positive sentiment class F1-Score is worth 74% because only 42 data were predicted correctly and 16 were predicted incorrectly. Whereas the negative sentiment class gets a value of 70%, which means that 43 data are predicted as True Negative and 19 others are wrong in the prediction.

ACKNOWLEDGMENT *(Heading 5)*

After doing sentiment analysis, it can be seen that people are still quite satisfied with the features in the GOBIS Suroboyo Bus application. The advantages of the application based on the wordcloud diagram are that the application is enough to help Suroboyo Bus passengers in knowing the position of the bus. As for the deficiencies that exist, including the use of applications which are still quite difficult for new users, especially in the account section.
Some of the things suggested by the authors in this study to increase the value of accuracy with the same research topic include:

1. Adding the amount of data in research so that the data used is more diverse so as to produce different conclusions in the end.
2. In the data labeling process, program code is used to facilitate further researchers.

REFERENCES

[1] Alhaqq, R. I., Putra, I. K., & Ruldeviyani, Y. (2022). Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store. Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 11(2), 105 - 113.

[2] Amrullah, A. Z., Anas, A. S., & Hidayat, M. A. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. Jurnal BITe : Jurnal Bumigora Information Technology, 2(1), 40-44.

[3] Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. (2020). SENTIMENT ANALYSIS ON E-SPORTS FOR EDUCATION CURRICULUM USING NAIVE BAYES AND SUPPORT VECTOR MACHINE. Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information), 109 - 121.

[4] Firmansyah, R. A., & Putra, K. H. (2019).

ANALISIS TINGKAT KEPUASAN PENGGUNA TRANSPORTASI UMUM “SUROBOYO BUS” RUTE HALTE RAJAWALI – TERMINAL PURABAYA DENGAN METODE IMPORTANCE PERFORMANCE ANALYSIS (IPA). Jurnal Teknologi Perencanaan, Perancangan, Lingkungan, dan Infrastruktur.

[5] Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. JEPIN (Jurnal Edukasi dan Penelitian Informatika), 4(2), 113 - 118.

[6] Hakim, S. N., Putra, A. J., & Khasanah, A. U. (2021). Sentiment analysis on myindihome user reviews using support vector machine and naïve bayes classifier method. International Journal of Industrial Optimization, 151-164.

[7] Handoko, M. R., & Neneng. (2021). SISTEM PAKAR DIAGNOSA PENYAKIT SELAMA KEHAMILAN MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB. Jurnal Teknologi dan Sistem Informasi (JTSI), 2(1), 50 - 58.

[8] Haqie, Z. A., Nadiah, R. E., & Ariyani, O. P. (2020). INOVASI PELAYANAN PUBLIK SUROBOYO BIS DI KOTA SURABAYA. Journal of Public Sector Innovations, 23-30.

[9] Hasri, C. F., & Alita, D. (2022). PENERAPAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORTVECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP DAMPAK VIRUS CORONA DI TWITTER. Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA), 3(2), 145 - 160.

[10] Herwinsyah, & Witanti, A. (2022). ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM). Jurnal Sistem Informasi dan Informatika (Simika), 5(2), 59 - 67.

[11] Laia, M. L., & Setyawan, Y. (2020). PERBANDINGAN HASIL KLASIFIKASI CURAH HUJAN MENGGUNAKAN ETODE SVM DAN NBC. Jurnal Statistika Industri dan Komputasi, 5(2), 51 - 61.

[12] Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 6(9), 4305-4313