International Journal of Information Technology and Computer Science Applications (IJITCSA)
p-ISSN: 2964-3139 e-ISSN: 2985-5330
Vol. 01, No. 03, page 129 - 137
Submitted 02/07/2023; Accepted 08/09/2023; Published 10//09/2023

Sentiment Analysis of the Use of Digital Banking Service Applications in Google Play Store Reviews Using Naïve Bayes Method

[1]Amalia Nur Soliha, [2]Tb Ai Munandar, [3]Muhammad Yasir
[1,2,3]Informatics, Bhayangkara Jakarta Raya University, Bekasi, INDONESIA
e-mail : [1]amalianrs23@gmail.com, [2]tb.aimunandar@dsn.ubharajaya.ac.id, [3]muhammad.yasir@dsn.ubharajaya.ac.id

Corresponding Autor: Amalia Nur Soliha

Abstract
The development of the financial system is characterized by the emergence of digital banking service applications that are widely circulated and can be accessed for free. With so many applications, users often feel confused in choosing which applications are safe to use. Before downloading an application on the Google Play Store, users will usually look at ratings and reviews first. However, the title of the best application cannot be pinned if only seen from the rating and number of downloads. This research was conducted to analyze sentiment on user reviews of digital banking service applications on the Google Play Store using the NBC (Naïve Bayes Classifier) method. Research using the NBC algorithm produced an accuracy value of 81% on the classification of Allo Bank reviews and 78% on the classification of Line Bank reviews.

Keywords—Sentiment Analysis, Digital Banking Service Applications, Reviews, Naïve Bayes, Google Play Store

1 Introduction

Economic growth in Indonesia can be seen from the increasingly sophisticated and stable financial system. One of them is the banking world which has an increasing number of products offered so that people can easily transact from saving to investing quickly and precisely. Digital banking is a banking service where all activities are carried out electronically or digitally [1]. User reviews and ratings about services are found in every application on the Google Play Store. These reviews can be in the form of praise, suggestions, criticism, and complaints that are useful for prospective users of the application of one of the banks that also provide digital services, namely Allo Bank and Line Bank. From these criticisms or negative comments, it can have an impact on the bank's performance because if potential users read, they are likely to move to another better digital banking service application. Many negative comments on the Allo Bank and Line Bank applications that appear on the Google Play Store have yet to receive a significant response from the bank.

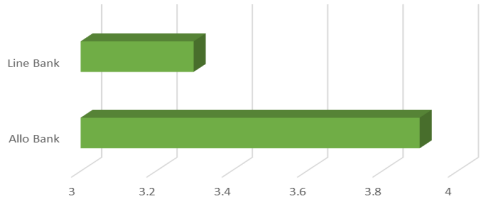

Figure 1. Digital Banking Service App Star Rating Chart

Figure 1 is a graph of the number of star ratings of two digital banking apps on the Google Play Store. Most people will look at the star rating before downloading an app and speculate that the best app can be seen from the star rating, if the star rating is above 3 then the app is declared as the best app and vice versa if the star rating is below 3 then it is declared as a bad app. However, the title of the best application is certainly less relevant if it is only based

on the rating and the number of downloads. Therefore, this study aims to analyze user review comments of digital banking service applications Allo Bank and Line Bank using text mining and classified into two classes, namely positive and negative as additional variables to determine which application is the best and safe in the Google Play Store and can also be obtained sentiment results from each application. There are many approaches that can be used for sentiment analysis, one of which is NBC (Naïve Bayes Classifier). Naïve Bayes Classifier is a classifier with the Bayesian Theorem which has simple probabilities and has the assumption that each variable X is independent [2], [3]. Based on its function, Naïve Bayes algorithm has three types, namely Multinominal Naïve Bayes [4], Bernoulli Naïve Bayes [5], and Gaussian Naïve Bayes [6]. Many studies use the Naïve Bayes algorithm, for example, to analyze sentiment on news or reviews on applications, such as sentiment analysis of the Ruang Guru application on Twitter [7], sentiment analysis of BMKG National twitter data review [8], sentiment analysis of JD.ID online store customers based on emotion icon conversion [9], sentiment analysis of halodoc apps [10] and there are many other research studies such as those conducted [11]–[13]. Although there are many other approaches to sentiment analysis, this research uses the Naïve Bayes algorithm because it is a simple model that can compete with other algorithmic models, while working well for text classification [14].

2 Literature Study

There have been many studies related to sentiment analysis of product reviews on social media. Some of them utilize social media reviews to analyze sentiment using the Naïve Bayes algorithm and TF-IDF weighting, the study states that the Naïve Bayes algorithm assumes very strong independence when there is little training data and TF-IDF weighting can be done without normalization [15]–[18]. Other studies also mention that the naïve bayes algorithm is a reliable algorithm in text mining classification [19].

In addition, there are also several studies that compare several classification algorithms to get the most superior. Some of them are in research [20], [21] compared the Support Vector Machine and Naïve Bayes algorithms to conclude that the Naïve Bayes algorithm is superior because it produces a higher accuracy value. Another study compared the NBC, KNN and Decision Tree algorithms, the analysis resulted in the NBC algorithm producing the highest accuracy value compared to other methods, namely 100% [22]. There are also studies that mention that the NBC algorithm produces F1 measure, precision value, and recall value better than the KNN method [23], [24].

By looking at previous research, this research will conduct sentiment analysis on Digital Banking Service applications, namely Allo Bank and Line Bank. This research uses python programming language and Naïve Bayes Classifier method with TF-IDF weighting. The tools used in the analysis process are Google Colab and RapidMiner. The Naïve Bayes Classifier method will be utilized as a method to generate predictions for reviews with positive and negative categories. Python programming language was chosen because it is easy to use and apply [15].

3 Research methods

3.1 Research Stages

This research begins with collecting data on reviews and star ratings of the Allo Bank and Line Bank applications on the Google Play Store using google colab. The data is then labeled positive and negative reviews based on star ratings. Then pre-processing is used to get data that is more ready to be analyzed. After that, TF-IDF weighting is done to measure and assess how important words are in a document using statistical measures [25]. Then the data is divided into two types, the first is 70% training data that is entered into the Naïve Bayes algorithm for the needs of the classification model and 30% test data for accuracy testing. In the last stage, analysis is carried out using the Naïve Bayes Classifier algorithm to obtain accuracy, precision, recall and f1 score values based on the performance of the Machine Learning (ML) algorithm using the Confusion Matrix reference. After the analysis results are obtained, then a prediction test is carried out on a dataset that does not have a label using RapidMiner tools to determine whether the analysis results are accurate.

3.2 TF-IDF Weighting

TF-IDF weighting is a process carried out to calculate the frequency of occurrence of a word in a document [26]. The formula for calculating TF-IDF word weighting is as follows:

$$TF.IDF_{std}\,(t) = tf_d^t \times \log\frac{n}{df^t} \text{ .....(1)}$$

Description (1) :

tf_d^t = Number of word frequencies
n = Number of documents
df_t = Number of document frequencies per word.

3.3 Naïve Bayes Algorithm

Naïve Bayes algorithm is an algorithm that performs clustering using prospects and data [12]. The following is the basic formula for the Bayes theorem equation in the Naïve Bayes algorithm.

$$P(H|X) = \frac{P(X|H)\times P(H)}{P(X)} \text{.....(2)}$$

Description (2) :

P(H|X) = Probability of H based on condition X
P(X|H) = Probability of X based on hypothesis H
P(H) = Probability on hypothesis H (prior)
P(X) = Probability of X (observed sample data)
X = Data with unknown class
H = Hypothesis X is a specific class

After the Naïve Bayes calculation, the next step is to evaluate the performance using confusion matrix [27]. *Confusion Matrix is used to measure the performance of classification algorithms* [28].

Table 1. *Confusion Matrix*

		Actual	
Predicted		True Positive (TP)	False Positive (FP)
		False Negative (FN)	True Negative (TN)

So accuracy, precision, and recall can be formulated as follows [29]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{.....(3)}$$

$$Precision = \frac{TP}{TP+FP} \text{.....(4)}$$

$$Recall = \frac{TP}{TP+FN} \text{.....(5)}$$

4 Results and Discussion

Based on the results of data collection and labeling reviews with a star rating ≤ 3 given a negative label and a star rating ≥ 3 given a positive label, then a cleaning process is carried out in data pre-processing with several stages, namely case folding to convert text from uppercase to lowercase letters; filtering to remove words that are not important; tokenizing to separate sentences into words based on spaces [26]; and stemming to convert affix words into basic words [30]. The following Table 1 to Table 4 are some of the pre-processing results.

Table 1. Case Folding Process Result

Review	Case Folding
Tidak berfungsi dengan baik. Login di hp ini selalu salah pasword, padahal di hp lain bisa2 aja	tidak berfungsi dengan baik login di hp ini selalu salah pasword padahal di hp lain bisa aja
Ko akun saya tidak bisa dibuka padahal sandi nya benar nomer juga benar	ko akun saya tidak bisa dibuka padahal sandi nya benar nomer juga benar

Table 2. Filtering Process Result

Review	Filtering
Tidak berfungsi dengan baik. Login di hp ini selalu salah pasword, padahal di hp lain bisa2 aja	tidak berfungsi login hp salah pasword
Ko akun saya tidak bisa dibuka padahal sandi nya benar nomer juga benar	akun tidak bisa dibuka

Table 3. Tokenizing Process Result

Review	Tokenizing
Tidak berfungsi dengan baik. Login di hp ini selalu salah pasword, padahal di hp lain bisa2 aja	[‘tidak’, ‘berfungsi’, ‘login’, ‘hp’, ‘salah’, ‘pasword’]
Ko akun saya tidak bisa dibuka padahal sandi nya benar nomer juga benar	[‘akun’, ‘tidak’, ‘bisa’, ‘dibuka’]

Table 4. Stemming Process Result

Review	Tokenizing
[‘tidak’, ‘berfungsi’, ‘login’, ‘hp’, ‘salah’, ‘pasword’]	[‘tidak’, ‘fungsi’, ‘login’, ‘hp’, ‘salah’, ‘pasword’]
[‘akun’, ‘tidak’, ‘bisa’, ‘dibuka’]	[‘akun’, ‘tidak’, ‘bisa’, ‘buka’]

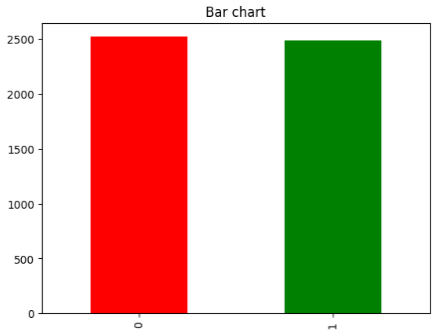

Figure 5. Allo Bank Review Comparison Bar Chart

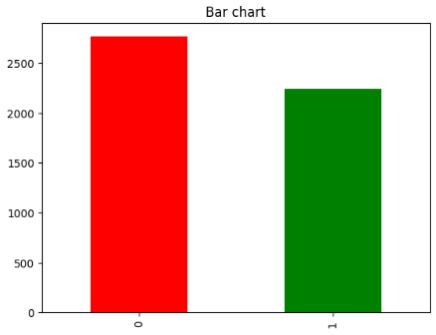

Figure 6. Line Bank Review Comparison Bar Chart

From the results of data preprocessing, visualization is obtained to see words that are often used in reviews by users of the Allo Bank and Line Bank applications. This data visualization is done separately between positive reviews and negative reviews. This data visualization is presented in the form of a wordcloud.

Figure 7. Wordcloud Allo Bank Positive Reviews

Figure 8. Wordcloud Allo Bank Negative Reviews

Figure 7 shows a collection of words with positive meanings that are often used by Allo Bank application users, such as smooth login, fast transactions, easy application, free admin fees, and good applications. While Figure 8 shows a collection of words with negative meanings that are often used by Allo Bank application users, such as failed registration, ordered to re-login due to error, cannot verify, complicated application.

Figure 9. Wordcloud Line Bank Positive Reviews

Figure 10. Wordcloud Line Bank Negative Reviews

Figure 9 shows a collection of words with positive meanings that are often used by Line Bank app users, such as good features, free admin fees, and easy transfers. While Figure 10 shows a collection of words with negative

meanings that are often used by Line Bank application users, such as failed updates, rejected transactions, incorrect verification, and difficult applications.

In the last stage of analysis, the modeling process is carried out using the Naïve Bayes Clasifier (NBC) method to obtain accuracy, precision, recall and f1 score values based on the performance of the Machine Learning (ML) algorithm using the Confusion Matrix reference. The following are the results of the accuracy value of the Allo Bank and Line Bank applications from the application of the NBC method.

```
---------- confusion matrix ----------
[[182  70]
 [ 27 221]]
---------- classification report ----------
              precision    recall  f1-score   support

           0       0.87      0.72      0.79       252
           1       0.76      0.89      0.82       248

    accuracy                           0.81       500
   macro avg       0.82      0.81      0.80       500
weighted avg       0.82      0.81      0.80       500
```

Figure 11. Allo Bank App Accuracy Score

```
---------- confusion matrix ----------
[[198  78]
 [ 34 190]]
---------- classification report ----------
              precision    recall  f1-score   support

           0       0.85      0.72      0.78       276
           1       0.71      0.85      0.77       224

    accuracy                           0.78       500
   macro avg       0.78      0.78      0.78       500
weighted avg       0.79      0.78      0.78       500
```

Figure 12. Line Bank App Accuracy Score

From the results of the calculation of the accuracy value above, the resulting accuracy value for digital banking services. It can be seen at 81% for the Allo Bank application and 78% for the Line Bank application. So from the results of the overall analysis that has been done, it can be concluded that based on reviews on the Allo Bank and Line Bank applications, the highest accuracy value is the Allo Bank application.

With the results of the above analysis, a prediction test is carried out on 100 new data that does not have a label to find out whether the analysis results are accurate or not. From the results of analysis and prediction tests on new data, it produces a total of 28 positive reviews and 72 negative reviews for the Allo Bank application. While the Line Bank application produces a total of 40 positive reviews and 60 negative reviews.

Row No.	prediction(label)	confidence(Negatif)	confidence(Positif)	text
1	Positif	0	1	transaksi aman

Figure 13. Allo Bank App Positive Prediction Test

Row No.	prediction(label)	confidence(Negatif)	confidence(Positif)	text
1	Negatif	1	0	aplikasi jelek

Figure 14. Allo Bank App Negative Prediction Test

Row No.	prediction(label)	confidence(Negatif)	confidence(Positif)	text
1	Positif	0	1	bank terbaik

Figure 15. Line Bank App Positive Prediction Test

Row No.	prediction(label)	confidence(Negatif)	confidence(Positif)	text
1	Negatif	1	0	bank ribet aneh jelek

Figure 16. Line Bank App Negative Prediction Test

Figure 13 and Figure 15 show the positive review accuracy test by entering the sentence "transaksi aman" and "bank terbaik". While Figure 14 and Figure 15 show the negative review accuracy test by entering the sentence "aplikasi jelek" and "bank ribet aneh jelek".

5 Conclusion

Based on the testing and analysis that has been done, it can be concluded that:

1. The results of the implementation of the Naïve Bayes Classifier method show that the highest level of accuracy is generated in the Allo Bank application which is 81%, while the Line Bank application has an accuracy rate of 78%.
2. By paying attention to the class label user reviews on the Allo Bank and Line Bank applications tend to give negative responses compared to positive responses. Because in the Allo Bank application there are 2517 negative reviews out of 5000 total data, while in the Line Bank application there are 2763 negative reviews out of 5000 total data.

6 Suggestion

Suggestions for further research are to use more review data so that the resulting accuracy value is better. And do research with other classification methods for comparison and find classification methods that are superior in producing accuracy values in sentiment analysis.

BIBLIOGRAPHY

[1] R. I. T. Linggadjaya, B. Sitio, and P. Situmorang, "Transformasi Digital Pt Bank Jago Tbk dari Bank Konvensional menjadi Bank Digital," *Int. J. Digit. Entrep. Bus.*, vol. 3, no. 1, pp. 9–22, Feb. 2022, doi: 10.52238/ideb.v3i1.76.

[2] Y. Yuliana, P. Paradise, and K. Kusrini, "Sistem Pakar Diagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Classifier Berbasis Web," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 10, no. 3, p. 127, Mar. 2021, doi: 10.22303/csrid.10.3.2018.127-138.

[3] M. M. Saritas and A. Yasar, "International Journal of Intelligent Systems and Applications in Engineering Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE*, vol. 7, no. 2, pp. 88–91, 2019, doi: 10.1039/b000000x.

[4] M. Abbas, K. Ali, M. Saleem, and A. Jamali, "Multinomial Naive Bayes Classification Model for Sentiment Analysis Study about symmetric and Asymmetric algorithms View project Overview of China View project Multinomial Naive Bayes Classification Model for Sentiment Analysis," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, p. 62, 2019, doi: 10.13140/RG.2.2.30021.40169.

[5] N. S. Wardani, A. Prahutama, and P. Kartikasari, "ANALISIS SENTIMEN PEMINDAHAN IBU KOTA NEGARA DENGAN KLASIFIKASI NAÏVE BAYES UNTUK MODEL BERNOULLI DAN MULTINOMIAL", [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[6] L. Cataldi, L. Tiberi, and G. Costa, "Estimation of MCS intensity for Italy from high quality accelerometric data, using GMICEs and Gaussian Naïve Bayes Classifiers," *Bull. Earthq. Eng.*, vol. 19, no. 6, pp. 2325–2342, Apr. 2021, doi: 10.1007/s10518-021-01064-6.

[7] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," *J. Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.

[8] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," vol. 15, no. 1.

[9] F. V. Sari and A. Wibowo, "ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI," *J.*

SIMETRIS, vol. 10, no. 2, Jan. 2019.

[10] E. Indrayuni, A. Nurhadi, and D. A. Kristiyanti, "Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc," *Fakt. Exacta*, vol. 14, no. 2, p. 64, Aug. 2021, doi: 10.30998/faktorexacta.v14i2.9697.

[11] V. A. Permadi, "Analisis Sentimen Menggunakan Algritma Naïve Bayes Terhadap Review Restoran di Singapura 141." [Online]. Available: https://www.kaggle.com/hj5992/restaurantreviews

[12] D. S. Utami and A. Erfina, "Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes," 2022.

[13] A. Sri Widagdo, B. W. Soedijono A, and A. Nasiri, "Analisis Tingkat Kepopuleran E-Commerce Di Indonesia Berdasarkan Sentimen Sosial Media Menggunakan Metode Naïve Bayes," 2020.

[14] E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," vol. VII, no. 1, 2019.

[15] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes," vol. 9, no. 2, pp. 785–795, 2022.

[16] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 2, p. 422, Apr. 2021, doi: 10.30865/mib.v5i2.2845.

[17] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and Samudi, "ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE," 2020.

[18] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 1, pp. 157–163, Jan. 2021, Accessed: May 25, 2023. [Online]. Available: http://stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/2580/1870

[19] E. Y. Hidayat and M. A. Rizqi, "Klasifikasi Dokumen Berita Menggunakan Algoritma Enhanced Confix Stripping Stemmer dan Naïve Bayes Classifier," *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 90–99, Aug. 2020, doi: 10.25077/teknosi.v6i2.2020.90-99.

[20] C. F. Hasri and D. Alita, "PENERAPAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP DAMPAK VIRUS CORONA DI TWITTER," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022, [Online]. Available: http://jim.teknokrat.ac.id/index.php/informatika

[21] E. D. N. Sari and Irhamah, "Analisis Sentimen Nasabah Pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM)," 2019.

[22] M. K. Anam, B. N. Pikir, and M. B. Firdaus, "Penerapan Na ̈ive Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen danPemeritah," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 1, pp. 139–150, Nov. 2021, doi: 10.30812/matrik.v21i1.1092.

[23] U. Kiftiyani, N. Yudistira, and N. Sunan Kalijaga Yogyakarta, "Perangkingan Dokumen Berbahasa Arab Berdasarkan Query dengan Metode Klasifikasi Naïve Bayes dan K-Nearest Neighbor Arabic Document Ranking based on Query using Naïve Bayes and K-Nearest Neighbor Classification Method," 2020.

[24] H. Putri, A. I. Purnamasari, A. R. Dikananda, O. Nurdiawan, and S. Anwar, "Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN," *Build. Informatics, Technol. Sci.*, vol. 3, no. 3, pp. 331–337, Dec. 2021, doi: 10.47065/bits.v3i3.1093.

[25] J. A. Septian, T. M. Fahrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor." [Online]. Available: https://t.co/9WloaWpfD5

[26] A. Riyani, M. N. Zidny, and A. Burhanuddin, "Penerapan Cosine Similarity dan Pembobotan TF-IDF untuk Mendeteksi Kemiripan Dokumen," 2019.

[27] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," 2021.

[28] N. Salmi and Z. Rustam, "Naïve Bayes Classifier Models for Predicting the Colon Cancer," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jul. 2019. doi: 10.1088/1757-899X/546/5/052068.

[29] N. Amalina, "Uji Akurasi Aplikasi Augmented Reality Pembelajaran Huruf Alfabet Bahasa Isyarat Indonesia (BISINDO) pada Vuforia Menggunakan Confussion Matrix," Universitas Islam Negeri Maulana Malik

Ibrahim, Malang, 2019.

[30] M. U. Albab, Y. Karuniawati P, and M. N. Fawaiq, "Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic," vol. 20, no. 2, pp. 1–10, 2023, doi: 10.26623/transformatika.v20i2.5374.