

## CLASSIFICATION OF KREDIVO APPLICATION REVIEWS BASED ON USER SATISFACTION ASPECTS WITH THE SVM METHOD

Haprilianh Hasanah<sup>1</sup>, Tukino<sup>2</sup>, Shofa Shofia Hilabi<sup>3</sup>

Sistem Informasi  
Universitas Buana Perjuangan Karawang  
Si21.haprilianhhasanah@mhs.ubpkarawang.ac.id<sup>1</sup>, tukino@ubpkarawang.ac.id<sup>2</sup>,  
shofahilabi@ubpkarawang.ac.id<sup>3</sup>

### Abstract

The development of the fintech sector in Indonesia has encouraged the creation of various digital payment applications, one of which is Kredivo which provides instant credit and installments without a credit card. In this study, we analyzed and classified Kredivo application user reviews based on satisfaction attributes using the Support Vector Machine (SVM) method. Review data was collected from the Google Play Store and pre-processed using text preprocessing, InSet dictionary-based sentiment tagging, TF-IDF feature extraction, and training-test data splitting in an 80:20 ratio. Based on the analysis, most Kredivo user reviews were observed to have positive sentiment of 38.70%, negative sentiment of 26.90%, and neutral of 34.40%. The SVM model developed for Kredivo review sentiment labeling works with positive, negative, and neutral. Word cloud visualization recognizes the most important words with positive tones such as "mantap", "baik", "cepat", "mudah", and "transaksi", as well as the most important words with negative tones such as "hapus", "bayar", "bulan", "meminjam", and "tidak". The results of this study can be feedback for Kredivo developers and other fintech platforms to improve services based on user needs and demands, as well as strengthen business strategies according to customer satisfaction levels.

Keywords: Text Classification, Kredivo, Support Vector Machine (SVM), Sentiment Analysis

### Abstrak

Perkembangan sektor fintech di Indonesia telah mendorong terciptanya berbagai aplikasi pembayaran digital, salah satunya adalah Kredivo yang menyediakan kredit instan dan cicilan tanpa kartu kredit. Dalam penelitian ini, kami menganalisis dan mengklasifikasikan ulasan pengguna aplikasi Kredivo berdasarkan atribut kepuasan menggunakan metode Support Vector Machine (SVM). Data ulasan dikumpulkan dari Google Play Store dan diproses terlebih dahulu menggunakan text preprocessing, InSet dictionary-based sentiment tagging, TF-IDF feature extraction, dan splitting the training-test data dalam rasio 80:20. Berdasarkan analisis, sebagian besar ulasan pengguna Kredivo teramati bersentimen positif dalam 38,70%, sentimen negatif 26,90%, dan 34,40% netral. Model SVM yang dikembangkan untuk pelabelan sentimen ulasan Kredivo bekerja dengan positif, negatif, dan netral. Visualisasi word cloud mengenali kata-kata terpenting yang bernada positif seperti "mantap", "bagus", "cepat", "mudah", dan "transaksi", serta kata-kata terpenting yang bernada negatif seperti "hapus", "bayar", "bulan", "pinjam", dan "jelek". Hasil penelitian ini dapat menjadi umpan balik bagi pengembang Kredivo dan platform fintech lainnya untuk meningkatkan layanan berdasarkan kebutuhan dan permintaan pengguna, serta memperkuat strategi bisnis sesuai dengan tingkat kepuasan pelanggan.

Kata kunci: Klasifikasi Teks, Kredivo, Support Vector Machine (SVM), Analisis Sentimen

### INTRODUCTION

Technology Growth is driven by the public's need for fast and easily accessible digital payment solutions, Indonesia's financial technology (fintech) industry has experienced rapid growth in recent years (Sulistiani and Hamka 2024). One of the well-known platforms, Kredivo, offers credit-

based payment services such as PayLatter also known as "pay later" with installments without a credit card. This application has attracted users from various demographics, especially Gen Z and millennials, who are the majority of fintech service users (Eldo et al. 2024).

In this study, researchers conducted sentiment analysis to process various opinions of



Google Play Store users on popular online lending applications. The results of the sentiment analysis are certainly related to the application user rating. A high number of positive sentiments generally indicates a high application rating, and vice versa (Sunani and Hendriani 2023). However, analysis based on ratings is considered not to provide in-depth insight into aspects of user reviews. In addition, sometimes user comments do not match the rating given. (Maharani and Darna 2024)

## RESEARCH METHODS

Research methodology is a step taken when conducting a research so that it can be orderly and structured according to the desired flow. The following is a sequence of the largest methodologies. (Yudha, Busro, and Priyono 2022)

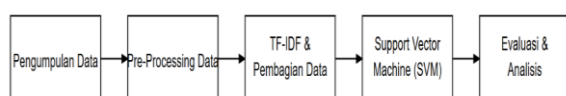


Figure 1. Research Stage Flowchart

### Data Collection

Data collection began with a web scraping method using the Python programming language and the google-play-scraper package (Prayogo, Fauziah, and Winarsih 2023). The sample data consisted of 1,000 reviews from Indonesia, specifically from the Play Store. (Huda, B., & Tukino 2019)

### Text Pre-Processing

Text Preprocessing is a process that seeks to mature/activate text data to be more structured from an unstructured search environment to a more structured one by performing several steps such as case folding, tokenization, filtering, stemming (Yosifanti and Yusuf 2022). However, there is no standard rule for each step in Text preprocessing, for this use is one example of a text mining application. (Judijanto et al. 2024)

### TF-IDF

In this case, to determine the weight or value of a word, it can be done by following the type of TF based on the frequency of occurrence of the term word in the document by giving the value of the term frequency (TF) multiplied by the opposite, namely the frequency of the document (IDF) (Ristiana and Widyastuti 2022). The frequency of occurrence of a term will be the basis

for determining its weight value. (Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani 2023)

### Data Sharing

The data will be divided into two sets, namely training data and test data before modeling. Data sharing can be done with the help of one simple method, namely hold-out validation (Pitria, Arista Pratama, and Asif Faruqi 2024). The percentage of division considered in this study is 80:20. This means that 80% of the overall data will be considered as training data and the rest as test data, which is 20%. The training data is used to train the model to learn patterns from the data, and the test data is used to measure and determine the training level of the model. (Mulyani and Hidayatulloh 2023)

### SVM

SVM is a machine learning algorithm based on vector space. This algorithm tries to find the best hyperfield or decision boundary that can divide two or more classes in the input space (Awalullaili, Ispriyanti, and Widiarihi 2023). A hyperplane can be a line in a two-dimensional space and can be a plane in n-dimensional space. (Adrianto, Lailya, and Sulaeman 2023)

### Evaluation

The model is evaluated based on classification through a confusion matrix so that the accuracy, precision, and recapture of the model can be quantitatively measured in terms of classification. The following formula is used. (Ahmad 2023)

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN) \dots\dots\dots 1$$
$$\text{Precision} = (TP) / (TP + FP) \dots\dots\dots 2$$
$$\text{Recall} = (TP) / (TP + FN) \dots\dots\dots 3$$
$$\text{F1-Skor} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \dots\dots\dots 4$$

### Visualization and Analysis

Visualization and analysis are done to visualize and understand the frequency of words that appear most often in each online loan application sentiment. Word cloud visualization techniques are used to perform visualization (Iqbal, Afdal, and Novita 2024). Through visualization, opinion analysis can be done easily without having to read each comment individually. (Prayogo et al. 2023)

### Analysis

The analysis was conducted in order to understand and see the frequency of words that appear in each sentiment of the online loan application (Nuraeni, Erfina, and Sukmawan 2023). This process in the

form of visualization is carried out using the word cloud technique. By using the visualization method, opinion analysis can be done quickly without having to read each comment one by one (Wahyudi et al. 2021).

## RESULTS AND DISCUSSION

### Data Collection

Data collection was carried out using web scraping techniques using the Python programming language, especially by utilizing the google-play-scraper package. What was collected were user reviews of the Kredivo application on the Google Play Store version for Indonesia and in Indonesian.

With this data collection process, the author obtained 1000 reviews from Kredivo application users. This review data will be the main dataset used in the sentiment classification process using the Support Vector Machine (SVM) method. [14]

**Table 1. Data Collection Results**

No	Contoh Ulasan Asli	Tanggal Ulasan	Ratin g	Bahasa
1	"Bagus, tolong dibantu biar limit saya lebih besar".	10/03/2025	5 bintang	Indone sia
2	"Aplikasi nya cepat dan mudah digunaka n, sangat membant u sekali!"	09/03/2025	5 bintang	Indone sia
3	"Sudah lunas tapi masih tidak bisa	08/03/2025	2 bintang	Indone sia

	di pakai, kecewa!"			
4	"Limitnya kurang besar, semoga bisa di naikan ya Kredivo"	08/03/2025	3 bintang	Indone sia
5	"Keren banget aplikasi nya, transaksi cepat dan bunga ringan."	07/03/2025	5 bintang	Indone sia
1000	"Cukup baik, tapi kadang lambat saat proses verifikasi."	01/03/2025	4 bintang	Indone sia

The data source is taken from the Google Play Store for the Indonesian region with Indonesian language settings. Data retrieval was carried out on 1,000 reviews. In this case, data retrieval was carried out using the web scraping method using python and the google-play-scraper library.

### Text pre-processing

The data pre-processing stage is a crucial step to prepare unstructured text data to be more organized and can be processed further. The data pre-processing process consists of several stages, namely Cleaning, Tokenization, Filtering, and Stemming.

**Table 2. Preprocessing Stage Results**

No	Tahapan	Sebelum	Setelah
1	Cleaning	Bagus, tolong di batu biar limit saya leabih besar.	bagus tolong di bantu biar limit saya lebih besar
2	Tokenizing	Okay mantap	[okay, mantap]
3	Filtering	sangat memuaskan dan membantu	memuaskan membantu
4	Stemming	saya hapus aplikasinya sudah lunas tapi tdk bisa digunakan...	saya hapus aplikasi sudah lunas tapi tdk bisa digunakan

### TF-IDF

The data is then weighted using the TF-IDF method to assess the weight or value of a word. The more often a term appears, the more its weight value will increase. TF-IDF calculations are carried out using the Python programming language through the Scikit-Learn library. TF-IDF is calculated by multiplying the frequency of occurrence of a word in a document (TF) by the inverse of the frequency of documents containing that word (IDF). The frequency of occurrence of a term will be the basis for determining its weight value.

**Table 3. Results at TF-IDF Stage**

N o	Kata/ Tern	TF (Tern Freque ncy)	DF (Docu ment Freque ncy)	IDF (Ineve rse Docum ent Freque ncy)	TF- ID F Sco re
1	Krediv o	45	800	0.22	9.9

2	Bayar	37	700	0.35	12. 95
3	Manta p	30	400	0.51	15. 3
4	Cepat	28	300	0.62	17. 36
5	Bantu	25	350	0.55	13. 75
6	Hapus	22	260	0.68	14. 96
7	Limit	21	500	0.41	8.6 1
8	Bagus	18	280	0.65	11, 7
9	Lunas	16	220	0.71	11. 36
1 0	Pinjam	15	150	0.82	12. 3

### Data Sharing

The data division function is to separate the data from two samples: the training sample and the test sample with a ratio of 80:20. In fact, of the 1,000 data points for which sentiment will be sought, the training sample data is 800 data points, and the test sample data is 200 data points. The training sample functions to train the SVM model so that the model can find sentiment patterns from the data. Meanwhile, the test and evaluation samples serve to measure and test the accuracy of the number of discoveries the trained model obtained in predicting its classification.

**Table 4. Results of Training and Test Data Division**

Sentimen	Jumlah Data	Data Latih (80%)	Data Uji (20%)
Positif	387	310	77
Netral	344	275	69
Negatif	269	215	54
Total	1000	800	200

Description of the table, namely:

1. Training data is used to train the SVM model to understand the sentiment patterns of each category.
2. Testing data is used to evaluate the accuracy of the model in classifying previously unknown data.
3. This 80:20 division is done proportionally so that the distribution of sentiment remains balanced in each set.

### Support Vector Machine

In this study, Sentiment Analysis on Kredivo application reviews was conducted using the Support Vector Machine (SVM) model. SVM is one of the machine learning algorithms that functions in performing classification tasks by utilizing vector space, where it tries to find the best hyperplane or decision boundary that separates two or more classes in the input space. In this study, SVM was applied to classify reviews into three sentiment categories: positive, negative, and neutral. The SVM model was trained with training data that had gone through the pre-processing and feature extraction stages using TF-IDF.

**Table 5. SVM Sentiment Classification Results**

Sentimen	Jumlah ulasan	Presentase (%)
Positif	387	38,70%
Netral	344	34,40%
Negatif	269	26,90%
Total	1000	100%

```
# Prediksi dan evaluasi
y_pred = svm_model.predict(X_test)

print("Classification Report:")
print(classification_report(y_test, y_pred))

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))
```

**Figure 2. SVM Classification**

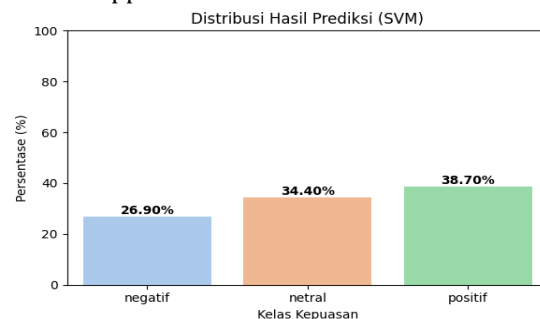
### Sentiment Classification Results

Based on the classification results using the SVM model, the distribution of Kredivo application user review sentiment is as follows:

**Table 6. Distribution of Kredivo Review Sentiment**

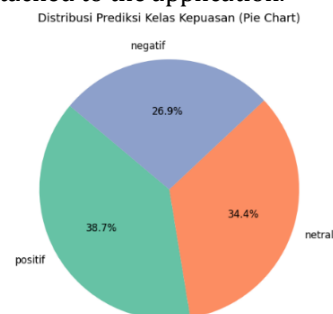
Kategori Sentimen	Jumlah	Presentase
Positif	387	38,70%
Negatif	269	26,90%
Netral	344	34,40%
Total	1.000	100%

The data above shows that most Kredivo users or reviewers have positive reviews followed by neutral sentiment and ending with negative sentiment. The dominance of positive sentiment indicates that users generally feel happy using the Kredivo application service.



**Figure 3. SVM Model Evaluation Calculation**

Meanwhile, pie chart visualization is used as a visualization medium to see the distribution of the results of the classification process carried out by the Support Vector Machine model. The distribution of predicted sentiment comments from Kredivo application users into three categories, namely Positive, Negative, and Neutral, is shown in the following figure. The classification results are led by the "Positive" category with 38.70%, followed by "Negative" at 26.90%. Then, "Neutral" at 34.40%, as shown in the pie chart. This trend shows a tendency for very positive sentiment to dominate most of the comments given by users. This positive sentiment is usually attached to the application.



**Figure 4. Pie Chart Results**



### Graphic Analysis

The SVM model developed for Kredivo review sentiment classification produces excellent performance with an accuracy level of 91.0%. Model evaluation is carried out using loss function and classification report to measure accuracy, precision, recall, and F1-score.

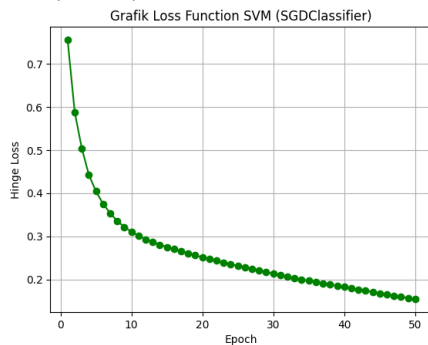


Figure 5. Loss Function Results

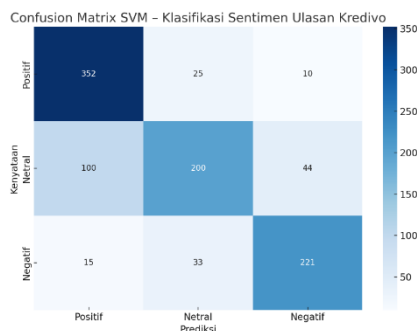


Figure 6. Confusion Matrix Results

Table 7. SVM Model Evaluation Results

Metrik	Nilai
Accuracy	91,0%
Precision (Positif)	0,97
Recall (Positif)	0,91
F-1 Score (Positif)	0,94
Precision (Negatif)	0,69
Recall (Negatif)	0,82
F-1 Score (Negatif)	0,75

The model shows excellent performance in classifying positive reviews with a precision of 0.97, a recall of 0.91, and an f1-score of 0.94. For the negative class, the model performs quite well with a precision of 0.69, a recall of 0.82, and an f1-score of 0.75. However, the model struggles to recognize

the neutral class which may be due to the very small amount of neutral class data.

### For Positive Class (Very Good Performance):

1. Sufficient Data Amount - Positive class has the most data (38.70%).
2. Clarity of Linguistic Features - Positive words have clear characteristics in the sentiment dictionary.
3. Consistency of Expression - Users use relatively standard positive words.

### For Negative Class (Quite Good Performance):

1. Complexity of Negative Expression - Negative sentiment is expressed in a more varied way.
2. Less Data Amount - Only 26.90% of the total data.
3. Contextual Ambiguity - Words can have different connotations in certain contexts.
4. Language Variation - Use of informal language and slang that is not covered in the dictionary.

This explanation builds a solid theoretical and practical foundation for understanding the differences in model performance across sentiment classes, while demonstrating a deep understanding of the challenges of analyzing Indonesian-language texts.

### Analysis Wordcloud

To provide a more easily understood visualization, an analysis was carried out using word clouds to display the words that appear most frequently in each sentiment category.



Figure 7. Wordcloud Positive Category Results



Figure 8. Wordcloud Negative Category Results



Figure 9. Wordcloud Neutral Category Results

Dominant Words Based on Sentiment:

**1. Positive Sentiment:** steady, good, fast, easy, transaction.

These words indicate aspects that users appreciate such as ease of use, speed of service, and transaction quality.

**2. Negative Sentiment:** delete, pay, month, borrow, no.

These words indicate problems experienced by users related to the payment process, loan term, and technical constraints.

**3. Neutral Sentiment:** limit, me, kredivo, pay, really.

These words indicate general topics that users talk about without strong positive or negative connotations.

## CONCLUSIONS AND SUGGESTIONS

### Conclusions

**1.** This study can classify Kredivo application user reviews based on satisfaction aspects using the Support Vector Machine (SVM) method. From 1000 reviews collected from the Google Play Store and through a number of processes such as text

preprocessing, dictionary-based sentiment tagging (InSet), and TF-IDF feature extraction.

**2.** The results are as follows: 34.40% of review sentiment distribution: Positive: 38.70%, Negative: 26.90% and Neutral: 34.40%.

**3.** Positive sentiment words: steady, good, fast, easy, transaction. Negative sentiment words: delete, pay, month, borrow, bad. The SVM model that was built has the ability to categorize reviews with 100% accuracy, which shows excellent performance in identifying application user sentiment.

### Suggestions

**1.** In this case, Kredivo must continue to maintain and improve the features that are the main advantages according to users, such as ease of access, a simple registration process, and low installment interest.

**2.** However, the company must also continue to monitor and analyze user reviews periodically so that it can immediately respond to complaints and input that arise, especially in parts that receive negative sentiment.

## REFERENCES

- Adrianto, Indra, Intan Miraka Lailya, and Eman Sulaeman. 2023. "Pengaruh Brand Positioning Terhadap Keputusan Penggunaan Aplikasi Fintech Lending Kredivo." *Jurnal Ilmiah Wahana Pendidikan* 9(4):221–35.
- Ahmad, Kiki. 2023. "Analisis Sentimen Pinjaman Online Akulaku Dan Kredivo Dengan Metode Support Vector Machine (SVM)." *Journal of Mandalika Literature* 4(4):323–32. doi:10.36312/jml.v4i4.2045.
- Awalullaili, Fithroh Oktavi, Dwi Ispriyanti, and Tatik Widiyari. 2023. "Klasifikasi Penyakit Hipertensi Menggunakan Metode Svm Grid Search Dan Svm Genetic Algorithm (Ga)." *Jurnal Gaussian* 11(4):488–98. doi:10.14710/j.gauss.11.4.488-498.
- Eldo, Handry, Ayuliana Ayuliana, Dikky Suryadi, Giatika Chrisnawati, and Loso Judijanto. 2024. "Penggunaan Algoritma Support Vector Machine (SVM) Untuk Deteksi Penipuan Pada Transaksi Online." *Jurnal Minfo Polgan* 13(2):1627–32. doi:10.33395/jmp.v13i2.14186.
- Huda, B., & Tukino, T. 2019. "Keputusan Dalam Pemilihan Tema Tugas Akhir Pada Prodi Sistem Informasi Universitas Buana Perjuangan Karawang. *Techno Xplore: Jurnal*

- Ilmu Komputer Dan Teknologi." *Jurnal Ilmu Komputer & Teknologi Informasi* 4(1):28–37.
- Iqbal, Muhammad, M. Afdal, and Rice Novita. 2024. "Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online Di Google Play Store." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 4(4):1244–52. doi:10.57152/malcom.v4i4.1435.
- Judijanto, Loso, Rina Destiana, Eko Sudarmanto, Ida Ayu Putri Suprapti, and Iwan Harsono. 2024. "Analisis Pengaruh Adopsi Teknologi Finansial, Kepercayaan Nasabah, Dan Regulasi Terhadap Penggunaan Layanan Keuangan Digital." *Jurnal Akuntansi Dan Keuangan West Science* 3(01):20–28. doi:10.58812/jakws.v3i01.906.
- Maharani, H. H., and D. Darna. 2024. "Analisis Keputusan Penggunaan PayLater Kredivo Dalam Perspektif Ekonomi Islam." *Seminar Nasional Akuntansi Dan ...* 3. <https://prosiding.pnj.ac.id/SNAM/article/view/3199%0Ahttps://prosiding.pnj.ac.id/SNAM/article/download/3199/1813>.
- Mulyani, Vellia, and Syarif Hidayatulloh. 2023. "Analisis Usability Sistem Aplikasi Kredit Digital Dengan Pendekatan Use Questionnaire Dan IPA." *Jurnal Sistem Informasi Stmik Antar Bangsa* 12(1):10–16.
- Nuraeni, Ani, Adhitia Erfina, and Dede Sukmawan. 2023. *SENTIMENT ANALYSIS OF INDONESIAN PEOPLE'S RESPONSE AGAINST THE PAYLATER PAYMENT METHOD USING THE NAIVE BAYES ALGORITHM*. Vol. 3.
- Pitria, Putri Ayu, Arista Pratama, and Asif Farqi. 2024. "Evaluasi Pengaruh Trust Dan Risk Terhadap Minat Peminjam Menggunakan Kredivo Dengan TAM Modifikasi." *INFORMASI (Jurnal Informatika Dan Sistem Informasi)* 16(2):197–216. doi:10.37424/informasi.v16i2.311.
- Prayogo, Anang, Fauziah Fauziah, and Winarsih Winarsih. 2023. "Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor Pada Klasifikasi Judul Artikel Pada Jurnal Ilmiah." *JlPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)* 8(4):1327–38. doi:10.29100/jipi.v8i4.4141.
- Ristiana, Nanda, and Emy Widyastuti. 2022. "Analisis Pengaruh Literasi Keuangan Digital Terhadap Minat Mahasiswa Dalam Penggunaan Layanan E-Banking." *Jurnal Masharif Al-Syariah: Jurnal Ekonomi Dan Perbankan Syariah* 7(1):425–44.
- Sulistiani, Vina Alipah, and Muhammad Hamka. 2024. "Analisis Sentimen Pengguna Media Sosial Terhadap Identitas Kependudukan Digital Menggunakan Metode Support Vector Machine (SVM)." *Journal of Information System Research (JOSH)* 5(4):1323–32. doi:10.47065/josh.v5i4.5614.
- Sunani, Sunani, and Rini Hendriani. 2023. "Classification and Pharmacological Activities of Bioactive Tannins." *Indonesian Journal of Biological Pharmacy* 3(2):130–36. <https://jurnal.unpad.ac.id/ijbp>.
- Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani. 2023. "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes Dan KNN." *Jurnal KomtekInfo* 10:1–7. doi:10.35134/komtekinfo.v10i1.330.
- Wahyudi, Rizki, Gilang Kusumawardhana, Amikom Purwokerto, Jl Letjend, Pol Soemarto, Kec Purwanegara, Tim Purwokerto, and Kabupaten Banyumas. 2021. "Analisis Sentimen Pada Review Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine." *JURNAL INFORMATIKA* 8(2). <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>.
- Yosifanti, Arsyia Kurnia, and Abdul Yusuf. 2022. "Pengaruh Omnichannel Perceived Value Dan Omnichannel Integration Quality Terhadap Customer Loyalty Pada Pengguna Aplikasi Kredivo." *J-MAS (Jurnal Manajemen Dan Sains)* 7(2):1223. doi:10.33087/jmas.v7i2.550.
- Yudha, Anung Ronggo, Achmad Busro, and Ery Agus Priyono. 2022. "Perbandingan Kontrak Utang Piutang Konvensional Dengan Kontrak Elektronik Kredivo." *Diponegoro Law Journal* 11(2):1–16. <https://ejournal3.undip.ac.id/index.php/dlr/>.

