Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

Analysis of JKN Mobile User Satisfaction using SVM and KNN Methods Through PSO Optimization

Esty Purwaningsih[1]*, Ela Nurelasari[2]

[1]*Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika*
[2] *Program Studi Teknologi Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika*
esty.epw@bsi.ac.id [1*]

Abstract

This study was conducted to evaluate the service quality of the JKN Mobile application developed by the Health Social Security Administering Agency (BPJS Kesehatan) as a means of facilitating participants in accessing health services. Although the application provides convenience for users, there are still various complaints indicating that the service is not running optimally. Therefore, this study aims to analyze the positive and negative sentiments of JKN Mobile application users by comparing the Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) algorithms based on Particle Swarm Optimization (PSO). The research method was carried out by processing user review data using sentiment classification techniques. The test results showed that the SVM algorithm obtained an accuracy of 85.02% with an AUC value of 0.815, while the PSO-based SVM increased to 86.71% with an AUC of 0.831. The KNN algorithm obtained an accuracy of 39.54% with an AUC of 0.500, while the PSO-based KNN increased to 87.05% with an AUC of 0.736. The results of the study prove that the implementation of PSO is able to improve the accuracy performance of both algorithms.

Keywords: *user satisfaction; application; KNN; SVM; PSO*

1. Introduction

The development of information technology has increased very rapidly over time and has penetrated almost every aspect of human life. This is supported by the development of the internet, which allows for the rapid dissemination of information [1]. Indonesia itself is the country with the third largest number of internet users in Asia after China and India [2]. The number of internet users in Indonesia in 2020 reached 196.7 million people, or around 73.7% of the total population [3]. Utilization of information systems can now be implemented digitally through computerized systems and smartphone-based applications. The Central Statistics Agency (BPS) recorded that the number of smartphone users in Indonesia reached 355.62 million in 2022, a 4.2% increase compared to the previous year. Optimizing digital-based services is a solution to provide easy access while improving the quality of public services [4].

The development of mobile-based digital applications is also occurring in the health insurance sector. The Social Security Agency for Health (BPJS Kesehatan) is a legal entity established by the government to administer health insurance for the public based on Law Number 24 of 2011. As a form of commitment to providing easily accessible health services to the public, on November 15, 2017, BPJS Kesehatan launched a mobile application called the National Health Insurance (JKN Mobile). This application aims to make it easier for National Health Insurance-Indonesian Healthy Card (JKN-KIS) participants to obtain health services quickly and practically via smartphone [5]. JKN Mobile is an electronic-based health service innovation that allows people to access health information and services simply through the palm of their hand [6].

Over the course of more than five years of use, the JKN Mobile app has received various criticisms and suggestions from users regarding the quality of its services and available features. User reviews in the Google Play Store comments section can be used as evaluation material for BPJS Kesehatan, the developer.Sentiment analysis is used to automatically identify positive and negative user opinions [7]. Using this method, BPJS Kesehatan can evaluate the quality of JKN-KIS program services through user review analysis and understand public perception of the JKN Mobile application. This study employed supervised learning with the K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) algorithms, optimized using Particle Swarm Optimization (PSO) to improve accuracy. The sentiment analysis results are expected to be used as a consideration in improving the quality of JKN Mobile application services in the future.

2. Research Method

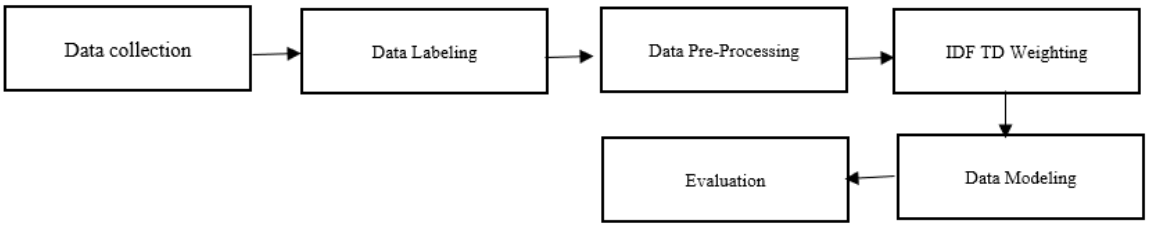

Fig.1: Research Stages

2.1. Data Collection

Collecting user rating data for the JKN mobile application service via Google Playstore around 2024. In this study, Google Colaboratory or better known as Google Colab, a web scraper, was used to carry out the data information retrieval process.

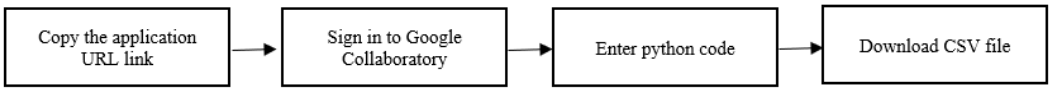

Fig.2: Data collection process

Figure 2 shows the stages that are gone through when data collection is carried out, the first is to copy the application URL, then enter Google Collaboratory, after that enter the following Python coding.

```
!pip install google-play-scraper

Collecting google-play-scraper
  Downloading google_play_scraper-1.2.6-py3-none-any.whl (28 kB)
Installing collected packages: google-play-scraper
Successfully installed google-play-scraper-1.2.6

[ ] from google_play_scraper import app

    import pandas as pd

    import numpy as np
```

Fig.3: Data import code snippet

```
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'app.bpjs.mobile',
    lang='id', # defaults to 'en'
    country='id', # defaults to 'us'
    sort=Sort.MOST_RELEVANT, # defaults to Sort.MOST_RELEVANT you can use Sort.NEWEST to get newst reviews
    count=600,
    filter_score_with=None # defaults to None(means all score) Use 1 or 2 or 3 or 4 or 5 to select certain score
)
```

Fig. 4: Data import code snippet

In this study, the data taken from the JKN Mobile application review on Google Playstore was 600 data, but we managed to take 597 data.

2.2. Data Labeling

Username	score	at	content	Sentimen
Fitria Nur Fajar Imami	3	17/05/2024 13:25	Kenapa aplikasi nya gak bisa login. Di bagian captcha mute	Negatif
ali imran chanel	1	17/05/2024 12:26	apliksh tidak pernah berubah kurang mendukung jelek bar	Negatif
zu zu	1	17/05/2024 12:26	ga bisa di pake, ada notif error tdk bisa dipakai krn perang	Negatif
Nono Bae	2	17/05/2024 12:22	Aplikasi daftar gagal terus.. Pada hal baru daftar..no HP uc	Negatif
naa 83	1	17/05/2024 11:23	Mau daftar kok susah banget ya.. Berkali-kali masukin no	Negatif
Adrian Nandi	1	17/05/2024 10:53	Daftar kok enggak dapat sms kode OTP ini gimana sih pad	Negatif
Asef Tegar	1	17/05/2024 08:07	Katanya nomor telpon tidak boleh di kosongkan, tapi kode	Negatif
Khoirul Anwar	5	17/05/2024 07:52	Aplikasi sangat bermanfaat dan sangat memudahkan dala	Positif
TARUT Channel	1	17/05/2024 07:48	Mau daftar pake no telpon susah banget,gagal terus "coba	Negatif
Irna Noviyanti	1	17/05/2024 07:18	Aplikasinya sulit buat reset password. Sangat tidak user fr	Negatif
zidni fauzan	3	17/05/2024 06:11	Aplikasi kurang terlalu memadai gaada tombol cancel unt	Negatif

Fig. 5: Results of manual labeling

The review data in the JKN Mobile application is unlabeled or unsupervised data. Therefore, a labeling process is required to determine the labels for the JKN Mobile reviews. After the data has gone through the pre-processing stage, the labeling process is carried out. This data labeling stage is carried out manually with a score criteria of 1 to 3 labeled with a negative value and 4 to 5 labeled with a positive value, resulting in 513 negative data and 84 positive data.

2.3. Data Preprocessing

Preprocessing: This is the initial stage of data collection before processing or analysis. In this step, the data is cleaned of unnecessary words. Next, modeling is performed using the Support Vector Machine and K-Nearest Neighbor methods. The research consists of several phases, including:

2.3.1. *Transform Case*

At this stage, the retrieved JKN Mobile evaluation data is converted entirely to lowercase, as both capital and lowercase letters are included in the evaluation. The purpose of this case transformation is to improve consistency and simplify word matching.

Table 1: Transform Case Process

Before	After
Saran, pada proses verifikasi based SMS ketika user tidak sengaja close pop up untuk mengisikan kode verifikasi, lalu user kembali ke menu input tersebut. Selama masih tersisa hitung mundur yang 300 sekon tersebut, lebih baik system langsung menampilkan input text kode verifikasi yang telah dikirimkan. Daripada kembali mengisikan ulang no. handphone, yang tau-taunya harus nunggu ratusan sekon lagi supaya bisa.	saran, pada proses verifikasi based sms ketika user tidak sengaja close pop up untuk mengisikan kode verifikasi, lalu user kembali ke menu input tersebut. selama masih tersisa hitung mundur yang 300 sekon tersebut, lebih baik system langsung menampilkan input text kode verifikasi yang telah dikirimkan. daripada kembali mengisikan ulang no. handphone, yang tau-taunya harus nunggu ratusan sekon lagi supaya bisa.

2.3.2. *Remove Stopword*

At this stopword stage, the National Health Insurance Mobile Application data review involves deleting irrelevant sentences or words that have no meaning in a sentence.

Table 2: Stopword Removal Process

Before	After
saran pada proses verifikasi based sms ketika user tidak sengaja close pop up untuk mengisikan kode verifikasi lalu user kembali ke menu input tersebut selama masih tersisa hitung mundur yang sekon tersebut lebih baik system langsung menampilkan input text kode verifikasi yang telah dikirimkan daripada kembali mengisikan ulang no handphone yang tau taunya harus nunggu ratusan sekon lagi supaya bisa	saran proses verifikasi based sms user sengaja close pop up mengisikan kode verifikasi user menu input tersisa hitung mundur sekon system langsung menampilkan input text kode verifikasi dikirimkan mengisikan ulang no handphone tau taunya nunggu ratusan sekon

2.3.3. *Tokenizing*

In the tokenization phase, the completed JKN Mobile review data is removed from the stopwords and then separated word by word from the sentence.

Table 3: Tokenizing Process

Before	After
saran, pada proses verifikasi based sms ketika user tidak sengaja close pop up untuk mengisikan kode verifikasi, lalu user kembali ke menu input tersebut. selama masih tersisa hitung mundur yang 300 sekon tersebut, lebih baik system langsung menampilkan input text kode verifikasi yang telah dikirimkan. daripada kembali mengisikan ulang no. handphone, yang tau-taunya harus nunggu ratusan sekon lagi supaya bisa.	saran pada proses verifikasi based sms ketika user tidak sengaja close pop up untuk mengisikan kode verifikasi lalu user kembali ke menu input tersebut selama masih tersisa hitung mundur yang sekon tersebut lebih baik system langsung menampilkan input text kode verifikasi yang telah dikirimkan daripada kembali mengisikan ulang no handphone yang tau taunya harus nunggu ratusan sekon lagi supaya bisa

2.3.4. *Filter Tokens* (*by Length*)

At the Filter Tokens (by Length) stage, the JKN Mobile review data is carried out by eliminating words in the text that have a word length of less than 4 characters and eliminating words that have a length of more than 25 characters.

Table 4: Proses Filter Tokens (by length)

Before	After
saran proses verifikasi based sms user sengaja close pop up mengisikan kode verifikasi user menu input tersisa hitung mundur sekon system langsung menampilkan input text kode verifikasi dikirimkan mengisikan ulang no handphone tau taunya nunggu ratusan sekon	saran proses verifikasi based user sengaja close mengisikan kode verifikasi user menu input tersisa hitung mundur sekon system langsung menampilkan input text kode verifikasi dikirimkan mengisikan ulang handphone taunya nunggu ratusan sekon

2.4. Feature Weighting with TF-IDF

The TF-IDF weighting method uses an estimation mechanism to calculate the weight of each feature. First, a term frequency (TF) estimate is performed, which determines how frequently a word appears in a given document. Then, an inverse document frequency (IDF) calculation is performed, which determines the importance of each word within the overall document collection. These two calculations are combined to produce a weight that indicates the relative importance of each word within the document and the collection as a whole [8].

2.5. Data Modeling

In this modeling, this pattern will be tested against opinion data that has undergone TF-IDF preprocessing and labeling. The K-Nearest Neighbor and Support Vector Machine classification algorithms are used in this process, assisted by Particle Swarm Optimization to improve data accuracy. By comparing the results of both algorithms, we can determine which method is more effective for this review classification task, thus providing deeper insights into the data and aiding the decision-making process.

2.6. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a global optimization method discovered by Kennedy and Eberhart in 1995 based on experiments on the behavior of groups of birds and fish. Each element in Particle Swarm Optimization has a speed of particles moving in the search space with a speed that is dynamically adjusted to their historical behavior [9].

2.7. Support Vector Machine (SVM)

This method is a common way to categorize text data through optimal accuracy. It works by finding the optimal hyperplane that can separate two different classes. The advantage of Support Vector Machine in text classification is its ability to handle unstructured data very well. Support Vector Machine is one of the supervised learning methods used in the classification stage [10].

2.8. K-Nearest Neighbor

K-Nearest Neighbor (K-NN) is an algorithm for grouping data based on training data obtained from nearest neighbors [11]. The KNN algorithm is used as a simple yet effective data classification method during the data modeling stage. The K-Nearest Neighbor algorithm is an instance-based machine learning algorithm that performs classification based on data similarity.

3. Result and Discussion

3.1. Data Processing and Calculation Methods

3.1.1. Support Vector Machine and PSO-based Support Vector Machine

The training cycles calculation in this experiment was determined by conducting a trial run on the data by entering the C and Epsilon values. The results of the research conducted to determine the training cycles value are as follows.

Table 5: Experiment of Determining SVM+PSO Training Values

C	Epsilon	SVM		Position Size	C	Epsilon	SVM + PSO	
		Accuracy	AUC				Accuracy	AUC
0.0	**0.0**	**85.02 %**	**0.815**	**5**	**0.0**	**0.0**	**86.71%**	**0.831**
0.1	0.1	76.94 %	0.845	5	0.1	0.1	81.83%	0.747
0.2	0.2	75.60 %	0.844	5	0.2	0.2	82.82%	0.787
0.3	0.3	75.26 %	0.844	5	0.3	0.3	82.84%	0.787
0.4	0.4	72.73 %	0.847	5	0.4	0.4	82.68%	0.807
0.5	0.5	71.38 %	0.847	5	0.5	0.5	80.79%	0.796
0.6	0.6	72.06 %	0.847	5	0.6	0.6	80.81%	0.788
0.7	0.7	71.72 %	0.847	5	0.7	0.7	81.64%	0.796

0.8	0.8	72.56 %	0.848	5	0.8	0.8	83.66%	0.799
0.9	0.9	71.56 %	0.848	5	0.9	0.9	80.50%	0.765
0.0	0.1	86.03 %	0.500	5	0.0	0.1	86.03%	0.500
1.1	1.1	86.03 %	0.500	5	1.1	1.1	86.03%	0.500
1.0	0.1	77.62 %	0.845	5	1.0	0.1	84.03%	0.769

The impact of the test results shows that the use of the SVM process in the table above with C = 0.0 and Epsilon E = 0.0 produces an Accuracy of 85.02% and an AUC of = 0.815, this value is the highest value of all training that has been done and the highest gain in the PSO-based SVM trial test above is C = 0.0 and Epsilon E = 0.0 and Population Size = 5 obtained from this accuracy is 86.71% and AUC = 0.831, This indicates that applying Particle Swarm Optimization (PSO) shows increasingly better accuracy results.

3.1.2. K-Nearest Neighbor and PSO-based K-Nearest Neighbor

The K value is the number of nearest neighbors involved in estimating the class label on the test data. Cross-validation techniques, also known as cross-validation, can be used to find the best K value by testing various K values and selecting the one that produces the best performance. The table below shows the results of the experiment to find the best K value.

Table 6: Experiment of Determining KNN+PSO Training Values

K	K-NN		Position Size	K	K-NN + PSO	
	Accuracy	AUC			Accuracy	AUC
1	39.54%	0.500	5	1	83.50%	0.500
2	**39.54%**	**0.587**	5	2	83.50%	0.990
3	28.44%	0.611	5	3	85.37%	0.527
4	28.26%	0.683	5	4	86.20%	0.513
5	22.38%	0.699	5	5	86.88%	0.586
6	22.72%	0.790	5	6	86.88%	0.622
7	18.51%	0.811	5	7	86.71%	0.687
8	19.18%	0.792	5	8	86.87%	0.678
9	16.83%	0.806	5	9	86.88%	0.628
10	17.50%	0.812	**5**	**10**	**87.05%**	**0.736**

The experimental results show that the use of the K-Nearest Neighbor method in Table III.2 uses a determinant between K = 1 to K = 10, so the value of K = 2 produces the highest accuracy value of = 39.54% and AUC = 0.587 and Based on the table above, the best results obtained in the KNN + PSO method are with Position size = 5 and K = 10 with the results obtained accuracy of 87.05% and AUC of 0.736, this indicates that conditions using Optimization (PSO) can show increasingly effective accuracy results.

3.1.3. Confusion Matrix

The confusion matrix provides an assessment of classification performance based on whether objects are correct or incorrect, given the conditions obtained from training and testing. In a classification system, the confusion matrix contains actual (real) and predicted (predicted) information. The following results are obtained.

accuracy: 85.02% +/- 2.12% (micro average: 85.02%)

	true Negatif	true Positif	class precision
pred. Negatif	500	78	86.51%
pred. Positif	11	5	31.25%
class recall	97.85%	6.02%	

Fig. 6: Confusion Matrix Support Vector Machine

$$Accuracy\ SVM = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} = \frac{5+500}{5+500+78+11} = \frac{505}{594} = 0{,}8502 \quad (1)$$

accuracy: 86.71% +/- 2.50% (micro average: 86.70%)

	true Negatif	true Positif	class precision
pred. Negatif	507	75	87.11%
pred. Positif	4	8	66.67%
class recall	99.22%	9.64%	

Fig. 7: Confusion Matrix Support Vector Machine + PSO

$$Accuracy\ SVM + PSO = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} = \frac{8+507}{8+507+75+4} = \frac{515}{594} = 0{,}8671 \quad (2)$$

accuracy: 39.54% +/- 6.61% (micro average: 39.56%)

	true Negatif	true Positif	class precision
pred. Negatif	152	0	100.00%
pred. Positif	359	83	18.78%
class recall	29.75%	100.00%	

Fig. 8: Confusion Matrix K-Nearest Neighbor

$$Accuracy\ KNN = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} = \frac{83+152}{83+125+0+359} = \frac{235}{594} = 0{,}395 \quad (3)$$

	true Negatif	true Positif	class precision
pred. Negatif	509	75	87.16%
pred. Positif	2	8	80.00%
class recall	99.61%	9.64%	

Fig. 9: Confusion Matrix K-Nearest Neighbor+PSO

$$Accuracy\ KNN + PSO = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} = \frac{8+509}{8+509+75+2} = \frac{517}{594} = 0{,}8704 \quad (4)$$

3.2. Evaluasi

After data processing using Support Vector Machine and K-Nearest Neighbor based on Particle Swarm Optimization, a comparison will be made with the grouping results as shown in the graph below.

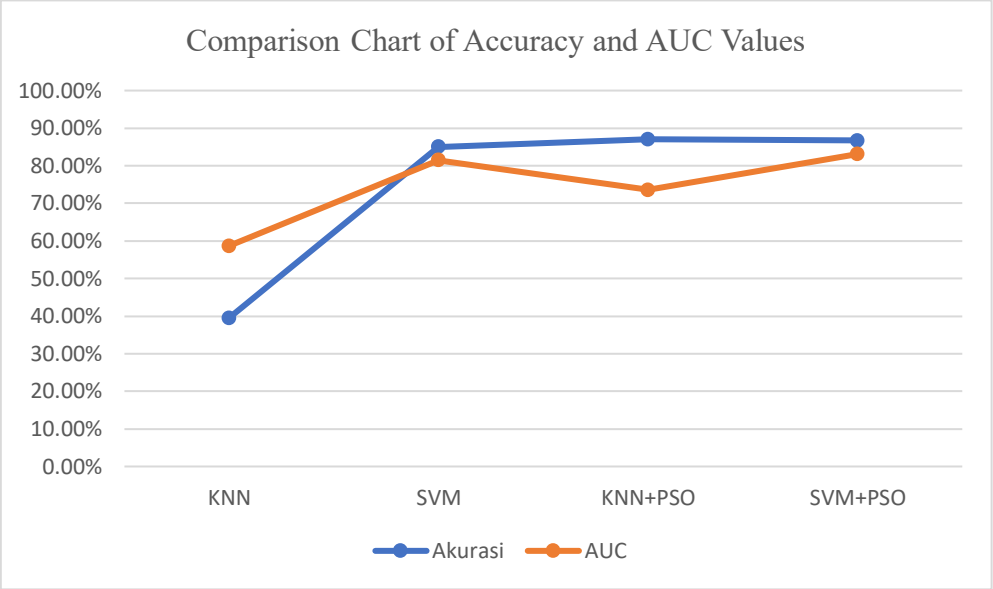

Fig. 10: Comparison Chart of Accuracy and AUC Values

The graph above shows a comparison of accuracy results between SVM, SVM with PSO, KNN, and KNN with PSO methods in two metrics: Accuracy Value and AUC. SVM and SVM with PSO methods performed almost equally well on both metrics, while KNN showed lower results but significantly improved when combined with PSO.

4. Conclusion

Based on the tests that have been carried out on the KNN and SVM algorithms based on PSO collected through Google Scrapper as many as 597 data, the accuracy results obtained for the SVM method are 85.02% and AUC is 0.815, and the PSO-based SVM method gets an accuracy result of 86.71% and AUC of 0.831, so it can be said that the application of PSO to the SVM method can increase the Accuracy and AUC Values.

Experiments were also conducted on the KNN method and the accuracy results obtained were 39.54% and the AUC results were 0.587, and for PSO-based KNN the accuracy results were 87.05% and the AUC was 0.736, so it can be said again that the application of PSO can increase the accuracy and AUC values.

Table 7: Results of the Tested Algorithm

No	Algoritma	Hasil Akurasi	Hasil Akurasi AUC
1	SVM	85.02%	0.815
2	**SVM + PSO**	**86.71%**	**0.831**
3	KNN	39.54%	0.587
4	**KNN + PSO**	**87.05%**	**0.736**

So overall the application of PSO on both methods succeeded in improving the performance of both, but it was the PSO-based KNN that showed a very high improvement compared to the PSO-based SVM. The suggestion from the next research is The suggestion for the next research considering the application of PSO in improving the performance of SVM and KNN is suggested to try the application of PSO on other algorithms such as Decision Tree, Random Forest or other models to see the potential improvement in these models.

References

[1] F. P. Sihotang, "Perbandingan Kualitas Layanan Dua Aplikasi Transportasi Online Menggunakan Metode Servqual," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 2, hal. 147–162, 2019, doi: 10.35957/jatisi.v6i2.190.

[2] Budy kusnandar Viva, "Pengguna Internet Indonesia Peringkat ke-3 Terbanyak di Asia," *14 september*, hal. 1, 21 Oktober 2021. [Daring]. Tersedia pada: https://databoks.katadata.co.id/datapublish/2021/10/14/pengguna-internet-indonesia-peringkat-ke-3-terbanyak-di-asia

[3] L. D. Jatmiko, "APJII: 196,7 Juta Warga Indonesia Sudah Melek Internet - Teknologi Bisnis.com," Bisnis.com.

[4] A. Aji Nugroho, E. Sutrisno, dan R. Ranintya Yusuf, "Optimalisasi Layanan Kesehatan Melalui Inovasi Pelayanan Publik Proyek Perubahan Pelatihan Kepemimpinan (Optimizing Health Services Through Public Service Innovation Project Change In Leadership Training)," *Setia Mengabdi J. Pengabdi. Kpd. Masy.*, vol. 1, no. 1, hal. 26–38, 2020, doi: 10.31113/setiamengabdi.v1i1.4.

[5] S. Ryan Aldi Prasetiyo, "EFEKTIVITAS PENGGUNAAN APLIKASI MOBILE JKN DALAM MENGURANGI ANTRIAN," vol. 7, no. 8.5.2017, hal. 2003–2005, 2022, [Daring]. Tersedia pada: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders

[6] A. Wulanadary, S. Sudarman, dan I. Ikhsan, "Inovasi Bpjs Kesehatan Dalam Pemberian Layanan Kepada Masyarakat : Aplikasi Mobile Jkn," *J. Public Policy*, vol. 5, no. 2, hal. 98, 2019, doi: 10.35308/jpp.v5i2.1119.

[7] R. Aryanti dan A. Saepudin, "Komparasi Algoritma Naive Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Pengguna Busway," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, hal. 174–180, 2022, doi: 10.31294/jtk.v4i2.

[8] M. I. Putri dan I. Kharisudin, "Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Analisis Sentimen Data Review Pengguna Aplikasi Marketplace Tokopedia," *Prism. Pros. Semin. Nas. Mat.*, vol. 5, hal. 759–766, 2022, [Daring]. Tersedia pada: https://journal.unnes.ac.id/sju/index.php/prisma/

[9] H. Muhamad *et al.*, "Optimasi Naïve Bayes Classifier Dengan Menggunakan PARTICLE SWARM OPTIMIZATION Pada data IRIS," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, hal. 180–184, 2017.

[10] D. Duei Putri, G. F. Nama, dan W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, Jan 2022, doi: 10.23960/jitet.v10i1.2262.

[11] H. Said, N. H. Matondang, dan H. N. Irmanda, "Penerapan Algoritma K-Nearest Neighbor Untuk Memprediksi Kualitas Air Yang Dapat Dikonsumsi," *Techno.Com*, vol. 21, no. 2, hal. 256–267, 2022, doi: 10.33633/tc.v21i2.5901.