Sentiment Analysis of the Convict Assimilation Program on Handling Covid-19

Aniq Noviciatie Ulfah[1], M. Khairul Anam[2], Novi Yona Sidratul Munti[3], Saleh Yaakub[4], Muhammad Bambang Firdaus[5]

[1]*Informatics Engineering, STMIK Amik Riau, Indonesia*
[2]*Information Technology, STMIK Amik Riau, Indonesia*
[3]*Informatics Engineering, Pahlawan Tuanku Tambusai University, Indonesia*
[4]*Information System, Muhammdiyah Jambi University, Indonesia*
[5]*Informatics, Mulawarman University, Indonesia*

[1]aniqnoviciatieulfah@sar.ac.id, [2]khairulanam@sar.ac.id,
[3]noviyona@universitaspahlawan.ac.id, [4]saleh@umjambi.ac.id, [5]bambangf@fkti.unmul.ac.id

Abstract - Coronavirus Disease-19 (Covid-19) is an infectious disease caused by the SARS-CoV-2 virus. The rapid spread of this disease has affected 216 other countries and regions, including Indonesia. In minimizing the spread and increasing losses, it is necessary to have several policies made by the Indonesian government in dealing with this. One of the policies taken by the government is the Convict Assimilation Program to prevent the spread of the virus in prisons. The Prisoner Assimilation Program fosters inmates by integrating prisoners into social life. Many media reported on the assimilation program in various media, including news portals, so that it became a forum for the public to express their opinions. News portals can be a source for getting public opinion. Therefore, sentiment analysis can be done to determine the sentiment of any existing public opinion. In this study, the analysis was carried out by applying one of the data mining methods, namely the Support Vector Machine, with positive, negative, and neutral sentiment labeling. The data used is audience comments in Indonesian with a dataset of 404 comments and then resampled so that the number of data becomes 669. The analysis uses the kernel Radial Basis Function (RBF), RBF with Grid Search, Polynomials, and Polynomials with grid search. Kernel RBF and Kernel Polynomial with Grid Search comparing test and training data 80%:20% with the highest accuracy of 95%.

Keywords: Covid-19, Sentiment Analysis, Support Vector Machine, RBF, Polynomial

I. INTRODUCTION

The Prisoners' Assimilation Program has been regulated in the Law of the Republic of Indonesia number 12 of 1995 concerning Corrections. Assimilation as a form of extramural coaching in prisons fosters Correctional Inmates who have met specific requirements by integrating them into community life [1]. The Open Penitentiary is a Penitentiary UPT that specifically carries out further development of prisoners at the assimilation stage, namely with a criminal period of between 2/3 of the criminal period that prisoners must serve [2]. The implementation of this assimilation caused many pros and cons in the community because it was applied during the Corona Virus Dieses 19 (Covid-19) pandemic [3]. but this is the right step to reduce the risk of the coronavirus [4].

Coronavirus Dieses 19 (Covid-19) is currently the most dangerous virus in the whole world [5]. Covid-19 was first discovered in Wuhan, and the case began with pneumonia or mysterious pneumonia in December 2019 [6]. From Wuhan, then covid-19 spread throughout the world, including Indonesia. The president immediately responded to the entry of covid in Indonesia through Presidential Decree No. 11 of 2020 concerning the Determination of a Public Health Emergency. Corona Virus Disease 2019 (COVID-19) has declared COVID-19 a public health emergency that must be taken care of [7]. Countermeasures have been carried out in many sectors: correctional institutions, namely the prisoner assimilation program during the COVID-19 pandemic [8].

The public's response to the assimilation of prisoners during the Covid-19 pandemic received many negative, positive, and neutral responses. Several news portals have reported on this issue and received few responses. This study will use comments on news portals as tested data. Research that analyzes positive, negative, and neutral is usually called sentiment analysis. Many researchers have done research related to sentiment analysis.

Several studies have been conducted, including analyzing hate speech sentiment on news portals using the Support Vector Machine (SVM), but this research

still produces a reasonably low accuracy, 53.88% [9]. Another SVM study resulted in an accuracy of 80.7% [10], 98.66% [11], 89%, 98% [12][13], respectively. Furthermore, the sentiment analysis used to analyze online transportation using SVM produces an accuracy of 96.04% [14]. This study will also use SVM because it has pretty high accuracy [15][16].

This study slightly differs from previous research, optimizing using a grid search. Experiments were carried out using two different kernels by combining them with grid search, and the two kernels were polynomial and Radial Basis Function (RBF). Some of these experiments were carried out to see the highest level of accuracy using the comparison of training data and 80:20 and 70:30 test data, respectively. RBF and Polynomial are used because in previous researchers these two kernels can help improve accuracy results on SVM [17].

II. METHOD

The process of conducting this research can be described in a plot, as shown in Fig. 1.

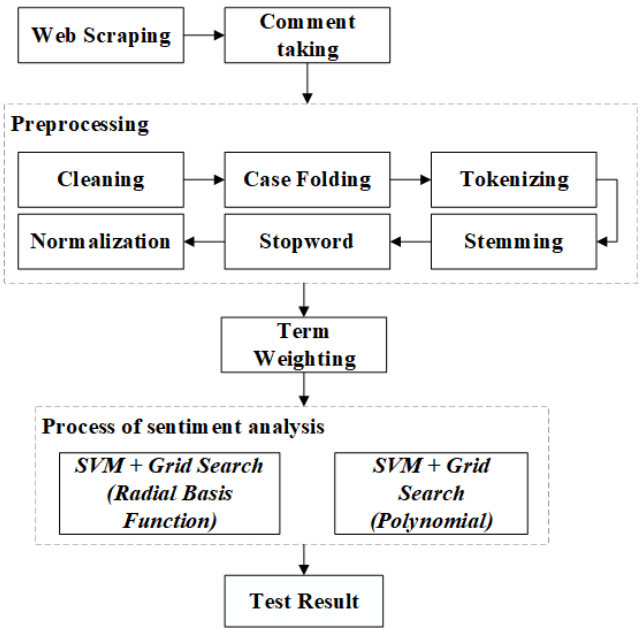

Fig. 1 Research methodology

A. Web Scrapping

The data is taken from the news portal Detik.com by applying the Web scrapping technique. This technique is applied using the Python programming language to discuss the Assimilation of Prisoners and COVID-19. This technique is used to retrieve audience comment data. We are collecting data using comment-API, which is a method that makes it easier to find comments. After the comments are obtained, then the comments are extracted.. In Table I are some examples of comments.

It can be seen in table 1 that the extracted documents are audience comments and are in the form of text. The extracted document is still in its original form because it has not gone through the pre-processing process, so many numbers, characters, and symbols are mixed, complicating the research if it is not cleaned. More complete data can be seen in the attachment.

B. Sampling Technique

Submissions must be in 1 column format with file type *.doc or *.docx. The data taken is the public opinion on the prisoner assimilation program in dealing with COVID-19, so the appropriate sampling technique in this study is purposive sampling. This technique does not take data based on strata, random, or regional, but data collection follows the criteria for the sample data needed for research [18].

C. Training Data and Test Data

This study divides the data into training data (training) and testing data (test). Training data contains a collection of comments/opinions used as training data that already has positive, negative, and neutral categories (classes). The training data used is taken from a collection of opinions that have been given labels and their sentiment classes which are done manually. This study will apply the ratio of training data and data testing of 70%:30%, 80%:20%, 90%:10%, respectively. The SVM method has high accuracy in research conducted by [19] applying the ratios of 60%:40%, 70%:30%, 80%:20%, and 90%:10%, respectively.

TABLE I
SOME AUDIENCE COMMENTS

No	Comments
1	Bukannya malah lebih besar resiko penyebaran Covid-9 nya kalau 30r napi dibebaskan?
2	skenario apa lg ini...???? apa sengaja utk buat Chaos di indonesia sehingga dberlakukan darurat sipil..?? Ya Allah lindungi Negara ini dr kezaliman
3	Terus mereka mau makan apa pak? Diluar lagi sulit.. jangan2 nanti jadi pemantik keributan diluar..🤦
4	Rezim ini sudah semakin kehilangan akal sehatnya. segala yg diperbuat sudah semakin aneh dan tdk wajar.

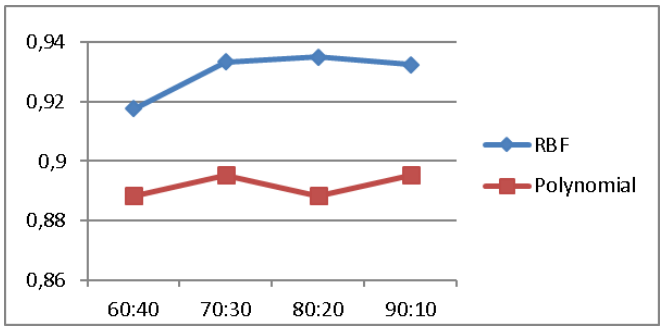

Fig. 2 Ratio accuracy of training data and test data

In Fig. 2, we can see that the RBF kernel gets a significant increase in accuracy at a ratio of 70%:30% and continues to increase at a ratio of 80%:20% and a slight decrease at a ratio of 90%:10%, respectively. While in the Polynomial kernel, the increase occurred in the ratio of 70%:30% and the ratio of 90%:10%, respectively. However, it slightly decreased at a ratio of 80%:20%. Therefore, in this study, the author will use the ratios of 70%:30%, 80%:20%, and 90%:10%, respectively, as a reference.

D. Resample

This stage will improve accuracy, which balances the data for model building. The 404 data have been manually labeled with the sentiment, 82 data labeled positive, 223 data labeled negative, and 82 data labeled neutral, respectively. The data is then resampled to balance the positive, negative, and neutral data. The number of data that has been resampled is 669 data. Table II is the result of the resample process in Python programming.

E. Pre-Processing

In the pre-processing stage, several steps will be passed, namely:

1) Cleansing, At this stage, comments are cleaned from words that are not needed to reduce noise. Emissions in comments are HTML characters, emoticons, hashtags (#), username (@username), and URLs [20].

2) Case Folding, all characters in the text are changed to lowercase at this stage. The processed characters are only letters 'a' to 'z's, and other than these characters will be omitted, such as punctuation marks (.), commas (,), and numbers [21].

3) Tokenizing, the sentence is broken at this stage, which was initially a sentence into pieces like words based on each word that composes it. So it can be said to return the conjunction [22].

4) Normalization or slang word conversion, at this stage, the incomplete or extended words are converted into ordinary words according to the Big Indonesian Dictionary (KBBI) [23].

5) Stopwords, at this stage, the vocabulary is deleted, which is not a unique word or does not convey any message significantly in the sentence. The intended vocabulary is connecting words and adverbs that are not unique words, such as "a," "by," "at," and so on [24].

6) Stemming, at this stage, the process is carried out to get essential words by removing prefixes, suffixes, insertions, and combinations of prefixes and suffixes [25].

F. Term Weighting

The next step in this research is the weighting of each word using TF-IDF. This stage is Numeric Statistics, which converts opinions into numerical data. TF-IDF uses the formula that we can see in (1), which is used to find the value of TF (term frequency), DF (document frequency), and IDF (inverse document frequency)—counting this document or term based on the frequency of occurrence of the term or document. The word/term has calculated the probability of its occurrence in one document (D1 to D5). The way to get the IDF, the following equation is used:

$$IDF = \log(D/DF) \quad (1)$$

Next, a document information table contains the term frequency (TF), document frequency (DF), and IDF for each term. Then look for the TF*IDF value of each term. The IDF value used is the IDF value obtained after the system training process (Table III).

TABLE II
RESAMPLE PROCESS RESULTS

Before		**After**	
Classification	**Total**	**Classification**	**Total**
negative	223	negative	223
neutral	99	neutral	223
positive	82	positive	223

TABEL III
TF-IDF PROCESS AND WORD WEIGHTING

Term/word	Documents				df	IDF(log D/df)	Documents			
	D1	D2	D3	D4			D1	D2	D3	D4
resiko	1	0	0	0	1	1	1	0	0	0
sebar	1	0	0	0	1	1	1	0	0	0
covid	1	0	0	0	1	1	1	0	0	0
nya	1	0	0	0	1	1	1	0	0	0
narapidana	1	0	0	0	1	1	1	0	0	0
bebas	1	0	0	0	1	1	1	0	0	0
skenario	0	1	0	0	1	1	0	1	0	0
sengaja	0	1	0	0	1	1	0	1	0	0
chaos	0	1	0	0	1	1	0	1	0	0
indonesia	0	1	0	0	1	1	0	1	0	0
laku	0	1	0	0	1	1	0	1	0	0
darurat	0	1	0	0	1	1	0	1	0	0
sipil	0	1	0	0	1	1	0	1	0	0
ya	0	1	0	0	1	1	0	1	0	0
allah	0	1	0	0	1	1	0	1	0	0
lindung	0	1	0	0	1	1	0	1	0	0
negara	0	1	0	0	1	1	0	1	0	0
zalim	0	1	0	0	1	1	0	1	0	0
makan	0	0	1	0	1	1	0	0	1	0
luar	0	0	2	0	1	1	0	0	2	0
sulit	0	0	1	0	1	1	0	0	1	0
mantik	0	0	1	0	1	1	0	0	1	0
ribut	0	0	1	0	1	1	0	0	1	0
rezim	0	0	0	1	1	1	0	0	0	1
hilang	0	0	0	1	1	1	0	0	0	1
akal	0	0	0	1	1	1	0	0	0	1
sehat	0	0	0	1	1	1	0	0	0	1
aneh	0	0	0	1	1	1	0	0	0	1
wajar	0	0	0	1	1	1	0	0	0	1

In Table III, word weighting has been carried out. This word weighting will be continued for the classification process using the Support Vector Machine.

G. Data Classification with Support Vector Machine

Text data will be classified using the SVM classification process at this stage. The goal is to classify public sentiment so that the public's views on the Prisoner's Assimilation Program can be known in dealing with Covid-19. The test flow using SVM can be seen in Fig. 3.

From Fig. 3, it can be seen that the TF IDF weighting results will be classified using SVM kernel Radial Basis Function (RBF) and Polynomial because the RBF kernel and Polynomial kernel are kernel functions that are commonly used in analysis when the data are not linearly separated [27]. If the classification score is greater than 0, the opinion is positive. If the score is less than 0, then the opinion is negative. If the score is equal to 0, then the opinion is neutral. After that, it will be continued with calculations to determine the accuracy and percentage of positive, negative, and neutral numbers.

III. RESULT AND DISCUSSION

After going through the pre-processing and weighting stages, the next stage is classification using the Support Vector Machine. The data will divide into three sentiment classes: positive, negative, and neutral. Tests will use the Radial Basis Function (RBF) kernel and Polynomial using the default parameters to find the best accuracy. The test will two times, comparing training data and test data of 80%: 20% and 70%: 30%, respectively.

Experiment using the RBF kernel. Kernel RBF is a kernel function that uses when data is not linearly separated. The parameters used in the RBF kernel are the cost (C) and gamma (γ) parameters. The parameter settings are set by default where the C parameter is one, and the gamma parameter is 'scale.' Experiment with RBF using a grid search. This experiment uses parameter

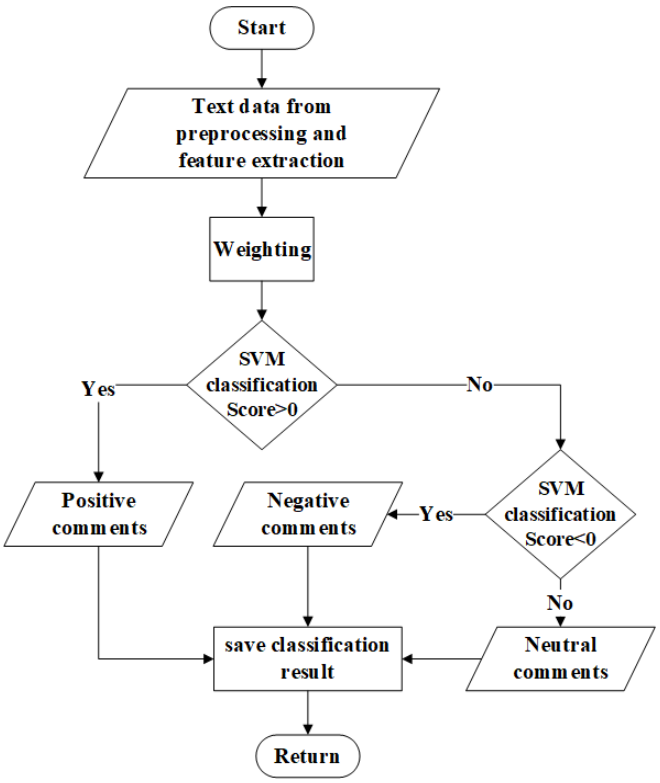

Fig. 3 Test flow using SVM [26]

values 0.1, 1, 10, 100, 1000 with gamma 1, 0.1, 0.01, 0.001, 0.0001, respectively.

Experiment using the Polynomial kernel. This kernel uses the cost (C) and degree (d) parameters. Parameters are set in the default settings where the value of parameter C is one, and the parameter degree is 3. This experiment uses 0.1, 1, 10, 100, 1000 with degrees 2 and 3, respectively. Experiment with Polynomials using a grid search.

A. Experiment Using the RBF kernel.

The typeface used in the script is Times New Roman, except for the writing of e-mail accounts. Other fonts can be used if needed for certain purposes.

The first experiment used training data and test data with a ratio of 70%:30%. The amount of data is 669, divided into 468 training data and 201 test data, respectively.

1) Experiment using the RBF kernel. In the experiment using RBF, the accuracy obtained was 92%. Figure 4 is the result obtained through the python program (Fig. 4). The accuracy results obtained are higher when compared to previous studies [28], which have an accuracy of 79.19%. The effect of RBF managed to increase more than 1% of the previous researchers to 92%, respectively.

2) Experiment using the RBF kernel with grid search. The next experiment added a selection grid search feature, but the results obtained did not affect the previous experiment, which was 92% (Fig. 5). This proves that feature selection does not continuously improve accuracy results.

3) Experiment using the polynomial kernel. The use of SVM and polynomial kernels in previous studies obtained a not high accuracy, namely 50% [29]. In this study, the results of SVM using a polynomial kernel increased (Fig. 6). The results in this study increased from previous studies to 91%, quite far compared to previous studies compared to the results obtained.

4) Experiment using the polynomial kernel with grid search. In the last experiment with 70:30 data, SVM uses a polynomial kernel with a selection grid search feature added. The results obtained increased from before (Fig. 7). These results prove that feature selection increases the accuracy results obtained even though the results are not significant, namely 1%. Table III is the conclusion of the with 70% training data and 30% test data.

	precision	recall	f1-score	support
negatif	0.98	0.82	0.89	76
netral	0.86	0.97	0.91	62
positif	0.91	0.98	0.95	63
accuracy			0.92	201
macro avg	0.92	0.92	0.92	201
weighted avg	0.92	0.92	0.91	201

Fig. 4 SVM using RBF kernel

	precision	recall	f1-score	support
negatif	0.98	0.82	0.89	76
netral	0.86	0.97	0.91	62
positif	0.91	0.98	0.95	63
accuracy			0.92	201
macro avg	0.92	0.92	0.92	201
weighted avg	0.92	0.92	0.91	201

Fig. 5 SVM using RBF kernel and grid search

	precision	recall	f1-score	support
negatif	1.00	0.81	0.89	78
netral	0.84	0.97	0.90	61
positif	0.90	0.98	0.94	62
accuracy			0.91	201
macro avg	0.91	0.92	0.91	201
weighted avg	0.92	0.91	0.91	201

Fig. 6 SVM using RBF polynomial

	precision	recall	f1-score	support
negatif	1.00	0.82	0.90	77
netral	0.84	0.97	0.90	61
positif	0.91	0.98	0.95	63
accuracy			0.92	201
macro avg	0.92	0.92	0.92	201
weighted avg	0.92	0.92	0.91	201

Fig. 7 SVM using polynomial kernel and grid search

TABEL III
EXPERIMENT WITH 70% TRAINING DATA AND 30% TEST DATA

Experiment	Accuracy
Experiment using the RBF kernel	92%
Experiment using the RBF kernel with Grid Search	92%
Experiment using the Polynomial kernel	91%
Experiment using the Polynomial kernel with Grid Search	92%

B. Experiment With 80% Training Data And 20% Test Data

Table IV is the result of the experiments conducted. The second experiment used training data and test data with a ratio of 80%:20%. The amount of data is 669 data will divide into 535 training data and 134 test data, respectively.

In the SVM experiment using the RBF kernel, there was a decrease in accuracy when using feature selection from 95% to 94%. Then the opposite happens when using polynomials an increase. This has increased considerably compared to previous studies that used the SVM method with the RBF kernel to get an accuracy of 71.55% [30]. Then this research is also significantly increased when compared to research using SVM with a Polynomial kernel, which is 51% [31].

TABEL IV
EXPERIMENT WITH 80% TRAINING DATA AND 20% TEST DATA

Experiment	Accuracy
Experiment using the RBF kernel	95%
Experiment using the RBF kernel with Grid Search	94%
Experiment using the Polynomial kernel	94%
Experiment using the Polynomial kernel with Grid Search	95%

IV. CONCLUSION

There is an increase and decrease in the accuracy results obtained. The highest accuracy in this study reached 95% with a data comparison of 80:20 using the RBF kernel and Polynomial with feature selection grid search. This study finds that SVM using the RBF kernel with feature selection grid search unrecommended to be applied. That is because there is no increase in accuracy, even a decrease in inaccuracy.

REFERENCES

[1] Presiden Republik Indonesia, *Undang-undang Repiblik Indonesia Nomor: 12 Tahun 1995 Tentang Pemasyarakatan*. Indonesia: The Ministry of Law and Human Rights is an Indonesian ministry that administers laws and human rights., 1995.

[2] A. Sanusi, "Evaluasi Pelaksanaan Pembinaan Narapidana di Lembaga Pemasyarakatan Terbuka," *J. Ilm. Kebijak. Huk.*, vol. 13, no. 2, p. 123, 2019, doi: 10.30641/kebijakan.2019.v13.123-138.

[3] D. R. Putri, Johari, and Husni, "Kebijakan Asimilasi Terhadap Narapidana Dimasa Pandemi COVID-19," *J. Ilm. Mhs. Fak. Huk. Univ. Malikussaleh*, vol. 2, no. 3, pp. 138–145, 2021, doi: 10.29103/jimfh.v4i2.4450.

[4] N. Y. I. Alfons, D. J. A. Hehanussa, and L. Lokollo, "Aspek Yuridis Pemberian Remisi Terhadap Narapidana Di Masa Pandemi COVID-19," *TATOHI J. Ilmu Huk.*, vol. 1, no. 4, pp. 337–345, 2021, [Online]. Available: https://fhukum.unpatti.ac.id/jurnal/tatohi/article/view/610.

[5] T. Alam and S. Qamar, "Coronavirus Disease (COVID-19): Reviews, Applications, and Current Status," *J. Inform. Univ. Pamulang*, vol. 5, no. 3, p. 213, 2020, doi: 10.32493/informatika.v5i3.6563.

[6] N. H. Nasution *et al.*, "Gambaran Pengetahuan Masyarakat Tentang Pencegahan Covid-19 Kecamatan Padangsidimpuan Batunadua, Kota Padangsidimpuan," *J. Biomedika dan Kesehat.*, vol. 4, no. 2, pp. 47–49, 2021, doi: 10.5193/health.v6i1.419.

[7] A. Syauqi, "Jalan Panjang COVID19 (sebuah refleksi dikala wabah merajalela berdampak pada perekonomian)," *JKUBS*, vol. 1, no. 1, pp. 1–19, 2020.

[8] S. Marthaningtiyas, "IMPLEMENTASI KEBIJAKAN ASIMILASI NARAPIDANA DI TENGAH PANDEMI COVID-19," *SUPREMASI J. Huk.*, vol. 3, no. 2, pp. 51–65, 2020, doi: 10.36441/supremasi.v3i2.130.

[9] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.35957/jatisi.v7i1.196.

[10] E. S. Romaito, M. K. Anam, Rahmaddeni, and A. N. Ulfah, "Perbandingan Algoritma SVM Dan NBC Dalam

Analisa Sentimen Pilkada Pada Twitter," *CSRID J.*, vol. 13, no. 3, pp. 169–179, 2021, doi: 10.22303/csrid.13.3.2021.169-179.

[11] P. H. Prastyo, I. Ardiyanto, and R. Hidayat, "A Combination of Query Expansion Ranking and GA-SVM for Improving Indonesian Sentiment Classification Performance," in *Procedia CIRP*, 2021, vol. 189, pp. 108–115, doi: 10.1016/j.procs.2021.05.074.

[12] E. Tyagi and A. K. Sharma, "Sentiment Analysis of Product Reviews using Support Vector Machine Learning Algorithm," *Indian J. Sci. Technol.*, vol. 10, no. 35, pp. 1–9, 2017, doi: 10.17485/ijst/2017/v10i35/118965.

[13] R. Maulana, P. A. Rahayuningsih, W. Irmayani, D. Saputra, and W. E. Jayanti, "Improved Accuracy of Sentiment Analysis Movie Review Using Support Vector Machine Based Information Gain," in *Journal of Physics: Conference Series*, 2020, vol. 1641, no. 1, doi: 10.1088/1742-6596/1641/1/012060.

[14] V. K. S. Que, A. Iriani, and H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 9, no. 2, pp. 162–170, 2020, doi: 10.22146/jnteti.v9i2.102.

[15] F. F. Irfani, M. Triyanto, A. D. Hartanto, and Kusnawi, "Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine," *JBMI (Jurnal Bisnis, Manajemen, dan Inform.*, vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.

[16] A. Veluchamy, H. Nguyen, M. L. Diop, and R. Iqbal, "Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches," *SMU Data Sci. Rev.*, vol. 1, no. 4, pp. 1–22, 2018,

[17] A. N. Sutranggono, "KLASIFIKASI EMOSI PADA CUITAN DI TWITTER DENGAN PRINCIPAL COMPONENT ANALYSIS DAN SUPPORT VECTOR MACHINE," *J. Ilm. Mat.*, vol. 10, no. 1, pp. 13–20, 2022, doi: 10.26740/mathunesa.v10n1.p13-20.

[18] I. Etikan, "Comparison of Convenience Sampling and Purposive Sampling," *Am. J. Theor. Appl. Stat.*, vol. 5, no. 1, p. 1, 2016, doi: 10.11648/j.ajtas.20160501.11.

[19] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 2, pp. 247–253, 2020, doi: 10.32736/sisfokom.v9i2.903.

[20] M. K. Anam, M. I. Mahendra, W. Agustin, Rahmaddeni, and Nurjayadi, "Framework for Analyzing Netizen Opinions on BPJS Using Sentiment Analysis and Social Network Analysis (SNA)," *Intensif*, vol. 6, no. 1, pp. 2549–6824, 2022, doi: 10.29407/intensif.v6i1.15870.

[21] Z. U. Siregar, R. R. A. Siregar, and R. Arianto, "Klasifikasi Sentiment Analysis Pada Komentar Peserta Diklat Menggunakan Metode K-Nearest Neighbor," *Kilat*, vol. 8, no. 1, pp. 81–92, 2019, doi: 10.33322/kilat.v8i1.421.

[22] D. Rustiana and N. Rahayu, "Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 1, pp. 113–120, 2017, doi: 10.24176/simet.v8i1.841.

[23] H. Tuhuteru, "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine," *Inf. Syst. Dev.*, vol. 5, no. 2, pp. 7–13, 2020.

[24] R. Julianto, E. D. Bintari, and Indrianti, "Analisis Sentimen Layanan Provider Telepon Seluler pada Twitter menggunakan Metode Naïve Bayesian Classification," *J. Big Data Anal. Artif. Intell.*, vol. 3, no. 1, 2017.

[25] F. O. Putri, Indriati, and R. C. Wihandika, "Analisis Sentimen pada Ulasan Pengguna MRT Jakarta Menggunakan Metode Neighbor-Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 7, pp. 2195–2203, 2020.

[26] R. Dehghani, H. Torabi Poudeh, H. Younesi, and B. Shahinejad, "Forecasting Daily River Flow Using an Artificial Flora–Support Vector Machine Hybrid Modeling Approach (Case Study: Karkheh Catchment, Iran)," *Air, Soil Water Res.*, vol. 13, 2020, doi: 10.1177/1178622120969659.

[27] H. C. S. Ningrum, "Perbandingan Metode Support Vector Machine (SVM) Linear, Radial Basis Function (RBF), dan Polinomial Kernel dalam Klasifikasi Bidang Studi Lanjut Pilihan Alumni UII," *Stat. UII*, 2018.

[28] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm," *J. Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi: 10.14710/j.gauss.v9i3.28932.

[29] I. Taufik and S. A. Pamungkas, "ANALISIS SENTIMEN TERHADAP TOKOH PUBLIK MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," *J. "LOG!K@,"* vol. 8, no. 1, pp. 69–79, 2018.

[30] R. Indraswari, A. Z. Arifin, and D. Herumurti, "Rbf Kernel Optimization Method With Particle Swarm Optimization On Svm Using The Analysis Of Input Data's Movement," *J. Ilmu Komput. dan Inf.*, vol. 10, no. 1, pp. 36–42, 2017, doi: 10.21609/jiki.v10i1.410.

[31] R. Mukarramah, D. Atmajaya, and L. B. Ilmawan, "Performance comparison of support vector machine (SVM) with linear kernel and polynomial kernel for multiclass sentiment analysis on twitter," *Ilk. J. Ilm.*, vol. 13, no. 2, pp. 168–174, 2021, doi: 10.33096/ilkom.v13i2.851.168-174.