

Optimization of Sentiment Analysis for MBKM Program using Naïve Bayes with Particle Swarm Optimization

Divi Aliyah, Zilrahmi, Yenni Kurniawati, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: zilrahmi@fmipa.unp.ac.id

Submitted : 02 Agustus 2024

Revised : 12 Agustus 2024

Accepted : 16 Agustus 2024

ABSTRACT

In early 2020, Kemendikbudristek launched the MBKM to improve the quality of higher education through a student-focused learning approach. The launch of this program triggered various reactions on social media, especially on Twitter, both positive and negative. This study aims to analyze the sentiment of Twitter users towards the MBKM program using the Naive Bayes algorithm optimized with Particle Swarm Optimization (PSO). The data used are Indonesian tweets containing the keywords "MBKM" and "Merdeka Campus" from July to December 2022. The research stages include data collection through crawling, manual labeling of data into positive and negative sentiments, data preprocessing, applying the Naive Bayes algorithm, and feature selection with PSO. The results showed that the group of tweets categorized based on positive and negative sentiments towards the implementation of the MBKM program in Indonesia in 2022, showed that the NB-PSO experiment achieved an accuracy of 90.87%, an increase of 7.12% compared to the Naive Bayes algorithm alone. Thus, using the Particle Swarm Optimization algorithm in the Naive Bayes classification algorithm is proven to improve classification performance, especially in the case of sentiment analysis.

Keywords: *Sentiment Analysis, Merdeka Belajar Kampus Merdeka (MBKM), Twitter, Naive Bayes, Particle Swarm Optimization (PSO).*



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Perkembangan dalam ilmu pengetahuan dan teknologi telah memberikan pengaruh yang signifikan di berbagai sektor, termasuk pendidikan. Pendidikan adalah usaha strategis yang memainkan peran penting dalam meningkatkan kualitas dan kecerdasan bangsa. Di tingkat perguruan tinggi, proses pembelajaran perlu berfokus pada mahasiswa (*student-centered learning*), sehingga mahasiswa dapat menyesuaikan diri dengan perubahan dan kemajuan ilmu pengetahuan serta teknologi, sehingga diharapkan menjadi profesional di bidangnya (Pramayasa dkk, 2023). Menanggapi kebutuhan ini, pada awal tahun 2020, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia (Kemendikbudristek RI) meluncurkan program Merdeka Belajar Kampus Merdeka (MBKM). Program ini adalah implementasi dari pendekatan pembelajaran yang berpusat pada mahasiswa, dengan memberikan kebebasan kepada mahasiswa untuk menjelajahi dan memperoleh pengetahuan yang sesuai dengan minat serta bakat masing-masing.

MBKM merupakan program yang relatif baru diimplementasikan dalam sistem pendidikan di Indonesia. Seperti halnya program baru, penerapan MBKM turut mengundang berbagai opini, baik yang mendukung maupun yang menentang. Pendapat dari masyarakat, terutama mahasiswa mengenai pelaksanaan program ini dapat ditemukan dari berbagai sumber, termasuk media sosial Twitter. Twitter adalah salah media sosial yang penggunaanya cukup banyak di Indonesia (Sarasvananda, 2022) dan berfungsi sebagai platform untuk menyampaikan opini masyarakat terhadap kebijakan publik yang dibuat oleh pemerintah. Opini masyarakat dalam menyampaikan pandangan mereka, khususnya terkait kebijakan program MBKM di Twitter, dapat dimanfaatkan sebagai bahan analisis sentimen untuk memahami persepsi masyarakat terkait program MBKM.

Analisis sentimen adalah metode untuk mengevaluasi apakah dalam sekumpulan data teks memiliki makna positif atau negatif (Indrayuni, 2016). Analisis sentimen diterapkan untuk mengetahui tanggapan masyarakat pengguna Twitter

terhadap pelaksanaan program MBKM. Dalam proses analisis sentimen, algoritma naïve bayes sering digunakan sebagai metode klasifikasi. Algoritma *Naïve Bayes* tetap mampu menghasilkan model klasifikasi yang baik meski menggunakan jumlah data yang sedikit, namun, algoritma *Naïve Bayes* juga memiliki kelemahan yaitu sangat sensitif terhadap pemilihan fitur yang dapat menyebabkan hasil optimasi probabilitas tidak berjalan secara optimal. Kelemahan ini dapat diatasi dengan berbagai teknik optimasi seperti penerapan seleksi fitur (Chen dkk, 2009). Algoritma *Particle Swarm Optimization* (PSO) sebagai salah satu metode dalam proses pemilihan fitur diharapkan dapat meningkatkan akurasi untuk mengatasi kelemahan *Naïve Bayes* pada klasifikasi sentimen. Pada penelitian ini dilakukan pengoptimalan metode *Naïve Bayes* dengan menerapkan metode seleksi fitur PSO untuk meningkatkan akurasi dari analisis opini twitter untuk mengetahui opini publik terhadap pelaksanaan program MBKM.

II. METODE PENELITIAN

A. Jenis Penelitian dan Sumber Data

Penelitian ini adalah penelitian terapan yang menggunakan data sekunder dari aplikasi Twitter. Pengumpulan data dilakukan melalui teknik *web crawling*, yaitu proses pengambilan data dari sebuah situs web dengan menggunakan kata kunci tertentu (Ruhyana, 2019)

B. Tahapan Penelitian

Langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut.

1. Pengumpulan data
Mengumpulkan data dari aplikasi twitter menggunakan teknik *crawling*.
2. Melakukan pelabelan data
Data diberi label secara manual oleh penulis, dengan menyesuaikan opini terhadap sentimen positif atau negatif. Kriteria pelabelan *tweet* adalah jika *tweet* mengandung kata-kata positif, disampaikan dengan baik, serta menunjukkan persetujuan, dukungan atau apresiasi, maka diberi label 'Positif'. Sebaliknya, jika *tweet* mengandung kata-kata negatif dan menunjukkan ketidaksetujuan atau protes, maka diberi label 'Negatif'.
3. Melakukan *preprocessing* data
Preprocessing adalah proses yang bertujuan untuk mengurangi *noise* yang ada dalam data, meningkatkan kualitas data dan menghasilkan representasi teks yang lebih terstruktur. Tahapan *preprocessing* yang diterapkan dalam penelitian ini adalah sebagai berikut.
 - a) *Case Folding*, yaitu tahapan mengubah semua huruf dalam data menjadi huruf kecil.
 - b) *Cleaning* data, yaitu tahap membersihkan dataset dengan menghapus kata dari atribut yang tidak diperlukan, tidak memiliki arti atau mempengaruhi sentimen seperti simbol, tanda baca, emotikon, angka dan sebagainya.
 - c) *Tokenizing*, yaitu tahap memotong teks menjadi kata-kata individual.
 - d) *Filtering*, yaitu tahap mengambil kata-kata penting dari hasil tokenisasi.
 - e) *Stemming*, yaitu proses memetakan dan mengurai kata dasar yang telah mengalami penambahan imbuhan dengan cara menghilangkan imbuhan tersebut. Dataset kata Bahasa Indonesia yang digunakan pada tahap ini adalah sastrawi, yang memiliki kosakata berdasarkan Kamus Besar Bahasa Indonesia (KBBI).
4. Melakukan klasifikasi dengan *Naïve Bayes Classifier*
Naïve Bayes merupakan metode pembelajaran mesin (*machine learning*) yang menggunakan perhitungan statistik dan perhitungan probabilitas, yang dikembangkan oleh ilmuwan Inggris, Thomas Bayes. Metode *Naïve Bayes* bekerja dengan memprediksi probabilitas kejadian di masa depan berdasarkan dari pengalaman sebelumnya (Tuhuteru & Iriani, 2018). Proses klasifikasi *Naïve Bayes* pada data dilakukan dengan merepresentasikan setiap data ke dalam " $X_1, X_2, X_3, \dots, X_n$ " (Liu dkk, 2017). Persamaan utama yang digunakan dalam metode ini adalah teorema bayes, yang dinyatakan sebagai berikut.

$$P(X|H) = \frac{P(H|X)P(H)}{P(X)} \quad (1)$$

Dimana X adalah data yang kelasnya tidak diketahui, H adalah data hipotesis, $P(H|X)$ adalah probabilitas hipotesis H berdasarkan kondisi X (*posterior probability*), $P(H)$ adalah probabilitas hipotesis H (*prior probability*), $P(X|H)$ adalah probabilitas X sesuai dengan kondisi pada hipotesis H (*likelihood*), $P(X)$ adalah probabilitas X secara keseluruhan (*marginal probability*).

5. Metode *K-Folds Cross Validation*

K-Folds Cross Validation adalah teknik yang digunakan untuk membagi data menjadi data latih dan data uji. Teknik ini sering digunakan oleh peneliti karena mampu mengurangi bias yang muncul saat mengambil sampel. Dalam *K-*

Folds Cross Validation, data terus dibagi menjadi data latih dan data uji, sehingga setiap data memiliki kesempatan untuk menjadi data uji. “K” merujuk pada jumlah pembagian data yang digunakan untuk membagi data latih dan data uji. Tabel 1 menggambarkan pembagian data menggunakan metode *K-fold cross validation*.

Tabel 1. Pembagian data metode *K-folds cross validation*

<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	...	<i>Fold K</i>
Test	Train	Train	...	Train
Train	Test	Train	...	Train
Train	Train	Test	...	Train
...
Train	Train	Train	...	Test

6. Melakukan tahapan seleksi fitur PSO

Particle Swarm Optimization (PSO) adalah metode optimasi berbasis populasi yang memanfaatkan sekumpulan individu (partikel) yang diperbarui pada setiap iterasi. Setiap partikel dalam metode ini bergerak menuju posisi terbaik yang pernah dicapainya sebelumnya (P_{best}) serta posisi terbaik secara global (G_{best}) untuk mencapai solusi optimal (Nurhadi, 2016). Selama proses iterasi, kecepatan dan posisi setiap partikel terus diperbarui untuk meningkatkan hasil optimasi.

Untuk mengoptimalkan *Naïve Bayes* menggunakan PSO, terdapat empat langkah utama yang dilakukan, pertama, melakukan inialisasi populasi (*swarm*). Kedua, menghitung akurasi berdasarkan fitur yang dipilih. Ketiga, menentukan akurasi klasifikasi yang terbaik dan keempat, memperbarui posisi serta kecepatan partikel. Beberapa parameter yang digunakan dalam proses optimasi PSO adalah sebagai berikut.

- a) *Population Size* (Q), adalah jumlah partikel dalam *swarm* atau populasi yang digunakan dalam proses optimasi. Setiap partikel mewakili solusi potensial di ruang pencarian.
- b) *Inertia Weight* (ω), adalah parameter yang mengontrol seberapa banyak pengaruh kecepatan partikel sebelumnya terhadap kecepatan saat ini dan dilambangkan dengan ω . Berikut persamaan dari update kecepatan.

$$v_{ij}^{(t)} = \omega_{ij}^{(t)} + c_1 r_{1j}^{(t)} (P_{best,i}^{(t)} - X_{ij}^{(t)}) + c_2 v_{2j}^{(t)} (G_{best,i}^{(t)} - X_{ij}^{(t)}) \tag{2}$$

- c) *Max Iteration*, adalah jumlah maksimum iterasi yang akan dilakukan selama proses optimasi.

7. Evaluasi kinerja algoritma dengan *Confusion Matrix*

Confusion matrix adalah tabel yang digunakan untuk menganalisis sejauh mana metode klasifikasi mampu memprediksi kelas-kelas data dengan akurat. *Confusion Matrix* dan rumus perhitungannya dapat dilihat pada Tabel 2.

Tabel 2. *Confusion Matrix*

Pred.Kelas	Kelas	
	Positif	Negatif
Negatif	TN	FN
Positif	FP	TP

$$Accuracy = \frac{(TP+TN)}{TP+TN+FP+FN} 100\%$$

$$Precision = \frac{TP}{(TP+F)} 100\%$$

$$Recall = \frac{TP}{(TP+FN)} 100\%$$

(3)

Keterangan:

TP : *True Positive*, jumlah data yang benar positif dan diprediksi sebagai positif oleh model.

- TN : *True Negative*, jumlah data yang benar negatif dan diprediksi sebagai negatif oleh model.
- FP : *False Positive*, jumlah data yang sebenarnya negatif tetapi diprediksi sebagai positif oleh model.
- FN : *False Negative*, jumlah data yang sebenarnya positif tetapi diprediksi sebagai negatif oleh model.

III. HASIL DAN PEMBAHASAN

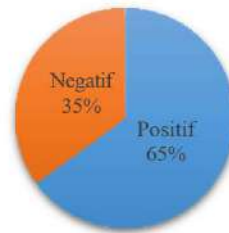
A. Pengumpulan Data

Data dikumpulkan dengan menggunakan teknik *crawling* melalui bahasa pemrograman Python. Langkah pertama adalah menghubungkan koneksi *Twitter API* ke dalam *Python*. Kemudian, memasukkan kata kunci “Kampus Merdeka” dan “MBKM”. Proses *crawling* menghasilkan 1200 data yang akan digunakan dalam penelitian ini. Data yang diperoleh akan digunakan sebagai dasar analisis sentimen terkait program MBKM.

B. Labeling data

Pelabelan data dilakukan secara manual dengan membaca dan memahami konteks dan makna setiap data secara teliti, kemudian diberikan label yang sesuai.

Persentase hasil labeling data dapat dilihat pada Gambar 1.



Gambar 1. Persentase Kategori Sentimen Program MBKM

Gambar 1 menunjukkan persentase sentimen masyarakat terhadap pelaksanaan program MBKM, dengan 65% sentimen positif dan 35% untuk sentimen negatif. Hal ini menunjukkan bahwa sentimen terhadap pelaksanaan MBKM cenderung bersifat positif.

C. Preprocessing Data

Preprocessing data bertujuan untuk membersihkan dan menyiapkan data untuk proses selanjutnya, yaitu proses klasifikasi. Terdapat 5 proses yang dilakukan pada tahapan ini yaitu *case folding*, *cleaning*, *tokenizing*, *filtering* dan *stemming*. Hasil dari tahapan *preprocessing* dapat dilihat pada tabel 3.

Tabel 3. Hasil *Preprocessing* Data

Tweet	Preprocessing
@collegemenfess Sebagai mahasiswa yang pernah mengikuti kampus merdeka, kebijakan bagus kok kalau kampusnya jg mendukung dengan segala fasilitasnya. Aku liat pengalaman selain kampus banyak yang dipersulit buat ikut program ini bahkan ga diizinin sama sekali 😊👍	Data Mentah
@collegemenfess sebagai mahasiswa yang pernah mengikuti kampus merdeka, kebijakan bagus kok kalau kampusnya jg mendukung dengan segala fasilitasnya. aku liat pengalaman selain kampus banyak yang dipersulit buat ikut program ini bahkan ga diizinin sama sekali 😊👍	Case Folding
sebagai mahasiswa yang pernah mengikuti kampus merdeka, kebijakan bagus kok kalau kampusnya jg mendukung dengan segala fasilitasnya. aku liat pengalaman selain kampus banyak yang dipersulit buat ikut program ini bahkan ga diizinin sama sekali	Cleansing

“sebagai” “mahasiswa” “yang” “pernah” “mengikuti” “kampus” “merdeka” “kebijakan” “bagus” “kok” “kalau” “kampus” “jg” “mendukung” “dengan” “segala” “fasilitas” “aku” “lihat” “pengalaman” “selain” “kampus” “banyak” “yang” “dipersulit” “buat” “ikut” “program” “ini” “bahkan” “ga” “diizinin” “sama” “sekali”	<i>Tokenizing</i>
[sebagai, mahasiswa, pernah, mengikuti, kampus, merdeka, kebijakan, bagus, kampus, mendukung, dengan, segala, fasilitas, aku, lihat, pengalaman, selain, kampus, banyak, dipersulit, ikut, program, ini, bahkan, diizin, sama, sekali]	<i>Filtering</i>
[sebagai, mahasiswa, pernah, ikut, kampus, merdeka, bijak, bagus, kampus, dukung, dengan, segala, fasilitas, lihat, alam, selain, kampus, banyak, sulit, program, ini, bahkan, izin, sama, sekali]	<i>Stemming</i>

D. Implementasi Algoritma

Pengujian awal dilakukan pada dataset yang telah disiapkan dengan menggunakan metode klasifikasi *Naïve Bayes* tanpa menerapkan seleksi fitur. Algoritma *Naïve Bayes* diterapkan melalui teknik *10-fold cross-validation*, dengan memvariasikan jumlah *folds* pada parameter untuk memperoleh akurasi terbaik. Teknik ini bertujuan untuk menghindari *overfitting* pada data (Mukherjee, 2017). Hasil pengujian dapat dilihat pada Tabel 4.

Tabel 4. Hasil Akurasi dengan *10-fold cross validation*

<i>Number of Folds</i>	<i>Accuracy(%)</i>
2	77,60%
3	78,09%
4	79,59%
5	80,10%
6	81,36%
7	83,75%
8	82,88%
9	82,37%
10	82,98%

Hasil akurasi *Naïve Bayes* dengan *10-fold cross validation* menunjukkan bahwa rata-rata akurasi untuk algoritma *Naïve Bayes* lebih dari 75% dengan nilai akurasi terendah adalah 77,60%. Sebagai perbandingan, nilai akurasi tertinggi *Naïve Bayes* adalah 83,75% untuk validasi 7.

Pengukuran performa *Naïve Bayes* dengan algoritma *Particle Swarm Optimization* (PSO) dilakukan dengan mengkombinasikan beberapa parameter, yaitu *Population Size* (Q), *Inertia Weight* (ω), dan jumlah iterasi maksimum untuk mencapai hasil optimal dan akurasi tertinggi. Pengujian awal adalah dengan memvariasikan nilai Q atau *population size*. Akurasi terbaik didapatkan dengan ukuran populasi (Q) sebesar 10, seperti yang ditunjukkan pada Tabel 5.

Tabel 5. Hasil *Naïve Bayes* PSO dengan *Population Size*

<i>Population Size</i>	<i>Accuracy(%)</i>
5	83,26%
10	86,67%
20	84,31%
40	85,71%

Pengujian kedua dilakukan dengan meragamkan nilai parameter *inertia weight* (ω) pada rentang 0,1 hingga 1,0 dengan ukuran populasi = 10 berdasarkan hasil pengujian pertama. Tabel 6 menunjukkan hasil akurasi pada pengujian kedua.

Tabel 6. Hasil *Naïve Bayes* PSO dengan *Inertia Weight* (ω)

<i>Inertia weight</i> (ω)	NB-PSO Accuracy
0,1	83,40%
0,2	84,34%
0,3	85,17%
0,4	85,46%
0,5	85,97%
0,6	86,88%
0,7	85,39%
0,8	86,89%
0,9	86,76%
1,0	86,37%

Pengujian ketiga dilakukan dengan meragamkan nilai jumlah maksimum iterasi. Pada percobaan ini digunakan *population size* (Q) = 10 dan *inertia weight* (ω) = 0,8, berdasarkan hasil pengujian pertama dan kedua. Sedangkan jumlah maksimum iterasi yang digunakan berkisar antara 3 hingga 50. Tabel 7 menunjukan hasil dari pengujian ketiga.

Tabel 7. Hasil pengujian NB-PSO dengan Maximum Iterasi

<i>Population Size</i> (Q)	<i>Inertia weight</i> (ω)	<i>Max Iteration</i>	NB-PSO Accuracy
10	0,8	3	89,17%
10	0,8	5	89,98%
10	0,8	10	90,87%
10	0,8	15	90,87%
10	0,8	20	90,87%
10	0,8	25	90,87%
10	0,8	30	90,87%
10	0,8	40	90,87%
10	0,8	45	90,87%
10	0,8	50	90,87%

Berdasarkan Tabel 7, akurasi terendah terjadi pada nilai dengan jumlah iterasi maksimum 3, yaitu sebesar 89,17%, sedangkan untuk jumlah iterasi maksimum pada rentang 10 hingga 50 menghasilkan nilai akurasi yang sama yaitu sebesar 90,87% untuk $Q = 10$ dan $\omega = 0,8$.

Berdasarkan pengujian yang telah dilakukan, parameter terbaik untuk PSO dalam metode klasifikasi *Naïve Bayes* adalah dengan *population Size* = 10, *inertia weight* = 0,8 dan *maximum iterasi* = 10.

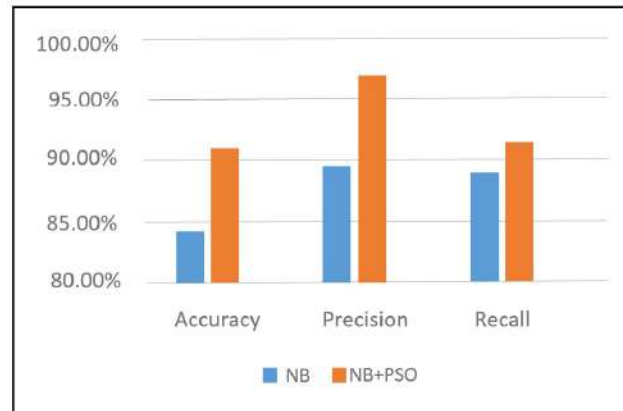
Hasil pengujian *Naïve Bayes* yang dioptimasi menggunakan parameter PSO menunjukkan peningkatan performa dengan mencapai nilai 90,87%. nilai *Confusion Matrix* NB-PSO dijelaskan pada Tabel 8.

Tabel 8. *Confusion Matrix* NB-PSO

Pred.Kelas	Kelas	
	Positif	Negatif
Negatif	141	48
Positif	25	586

E. Evaluasi

Penerapan PSO dapat meningkatkan akurasi dari akurasi algoritma Naïve Bayes. Peningkatan yang signifikan pada hasil kinerja pengukuran dapat dilihat dari grafik pada Gambar 2.



Gambar 2. Grafik Perbandingan Performa Pengukuran

Gambar 2 membandingkan kinerja model *Naïve Bayes* dengan model yang dioptimasi menggunakan PSO berdasarkan tiga metrik evaluasi yaitu *accuracy*, *precision* dan *recall*. Hasil pengujian menunjukkan bahwa optimasi dengan PSO meningkatkan kinerja *Naïve Bayes* secara signifikan. Model *Naïve Bayes* tanpa optimasi menghasilkan *accuracy* sebesar 83,75%, *precision* 89,86% dan *recall* 88,99%. Setelah diterapkan optimasi dengan PSO, *accuracy* meningkat menjadi 90,87%, menunjukkan kemampuan yang lebih baik dalam mengklasifikasikan data dengan benar. *Precision* juga mengalami peningkatan signifikan menjadi 97,02% yang berarti model NB-PSO lebih akurat dalam memberikan prediksi positif yang benar. Selain itu, *recall* meningkat menjadi 91,68%, menunjukkan bahwa model yang dioptimasi lebih baik dalam mengidentifikasi data yang sesuai.

IV. KESIMPULAN

Berdasarkan pengujian dan analisis, dapat disimpulkan bahwa penggunaan metode seleksi fitur *Particle Swarm Optimization* mampu meningkatkan performa dan akurasi serta mengatasi kelemahan *Naive Bayes* pada klasifikasi sentimen. Berdasarkan hasil evaluasi menggunakan *confusion matrix*, hasil *Naive Bayes* dengan PSO untuk analisis sentimen adalah 90,87%. Dapat disimpulkan bahwa metode yang diusulkan dapat meningkatkan akurasi sebesar 7,12% dari hasil analisis twitter untuk mengetahui opini publik terhadap pelaksanaan program MBKM di Indonesia tahun 2022. Saran untuk penelitian selanjutnya, mengombinasikan metode klasifikasi *Naive Bayes* dengan metode pemilihan fitur lainnya untuk mencapai hasil dan kinerja yang lebih maksimal. Selain itu, variasi dalam penentuan nilai indikator juga dapat dilakukan untuk mengoptimalkan hasil dari pemilihan fitur menggunakan *Particle Swarm Optimization* (PSO).

DAFTAR PUSTAKA

- Chen, J., Huang, H., Tian, S., & Qu, Y. (2009). Feature Selection for Text Classification with Naïve Bayes. *Expert Systems with Applications*, 36(3), 5432-5435.
- Indrayuni, E. (2016). Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization. *EVOLUSI: Jurnal Sains dan Manajemen*, 4(2).
- Liu, P., Yu, H., Xu, T., & Lan, C. (2017). Research on archives text classification based on Naive Bayes. In *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (pp. 187-190). IEEE.
- Mukherjee, S., & Bala, P. K. (2017). Detecting sarcasm in customer tweets: an NLP based approach. *Industrial Management & Data Systems*, 117(6), 1109-1126.

- Nurhadi, A. (2016). Implementasi Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization (PSO) untuk Klasifikasi Konten Berita Digital Bahasa Indonesia. *Sari*, 2(3), 48-56.
- Pramayasa, K., Maysanjaya, I. M. D., & Indradewi, I. G. A. A. D. (2023). Analisis Sentimen Program Mbkm Pada Media Sosial Twitter Menggunakan KNN Dan SMOTE. *SINTECH (Science and Information Technology) Journal*, 6(2), 89-98.
- Ruhyana, N. (2019). Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil/Genap Pada Twitter Dengan Metode Klasifikasi Naive Bayes. *IKRA-ITH Informatika: Jurnal Komputer dan Informatika*, 3(1), 94-99.
- Sarasvananda, I. B. G., Selivan, D., Radhitya, M. L., & Putra, I. N. T. A. (2022). Analisis Sentimen Pada Pembelajaran Daring Di Indonesia Melalui Twitter Menggunakan Naive Bayes Classifier. *SINTECH (Science and Information Technology) Journal*, 5(2), 227-233.
- Tuhuteru, H., & Iriani, A. (2018). Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier. *Jurnal Informatika: Jurnal Pengembangan IT*, 3(3), 394-401.