

Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka

Erina Undamayanti¹, Teguh Iman Hermanto², Ismi Kaniawulan³

^{1,2,3}Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana Purwakarta, Indonesia

e-mail: erinaundamayanti73@wastukencana.ac.id¹, teguhiman@wastukencana.ac.id², ismi@stt-wastukencana.ac.id³

Abstract

During the MBKM program running at several universities in Indonesia, several problems occurred, namely the implementation of the curriculum that did not have a reference, the disbursement of pocket money given was not on schedule, the policies of each partner were different, and the existence of the covid-19 pandemic. The way to find out public opinion or opinion about the MBKM program is to summarize public opinion on Twitter social media. This study aims to analyze the results of the classification of twitter users opinions on the MBKM program in Indonesia through sentiment analysis using the Naive Bayes method based on Particle Swarm Optimization. The research methodology carried out in this study was through the stages of data crawling, text preprocessing, feature extraction, classification, and evaluation. The data used in this study are 428 data. The results of the research in the form of sentiment analysis obtained are positive sentiments of 61.92%, it can be concluded that the MBKM program can be well received by the Twitter user community, especially students. Although there are some negative sentiments that appear around 38.08%. The results of this study can be used as a reference for the MBKM policy development team, especially the Kemendikbud POKJA team, because this program can provide benefits and experiences for students while the results of this research can be used as evaluation material for the team in the future to be even better.

Keywords: Sentiment Analysis, MBKM, Naive Bayes, Particle Swarm Optimization

Abstrak

Penelitian ini bertujuan untuk menganalisis hasil klasifikasi opini pengguna twitter terhadap program MBKM di Indonesia melalui analisis sentimen menggunakan metode Naive Bayes berbasis Particle Swarm Optimization. Metodologi penelitian yang dilakukan dalam penelitian ini adalah melalui tahapan Crawling data, Text Preprocessing, Ekstrasi Fitur, Klasifikasi, dan Evaluasi. Data yang digunakan dalam penelitian adalah 428 data. Hasil penelitian berupa analisis sentimen yang didapatkan adalah sentimen positif sebesar 61.92% dapat disimpulkan bahwa program MBKM ini dapat diterima dengan baik oleh masyarakat pengguna Twitter terutama mahasiswa. Walaupun ada beberapa sentimen negatif yang muncul sekitar 38.08%. Adapun hasil penelitian ini bisa dijadikan rujukan untuk tim pembuat pengembangan kebijakan MBKM terutama tim POKJA Kemendikbud, karena program ini bisa memberikan manfaat dan pengalaman bagi mahasiswa adapun hasil penelitian ini bisa dijadikan bahan evaluasi untuk tim kedepannya supaya menjadi lebih baik lagi.

Kata Kunci: Analisis Sentimen, MBKM, Naive Bayes, Particle Swarm Optimization

1. PENDAHULUAN

Di era *digital* perkembangan *internet* telah mengambil langkah besar dalam perkembangan *teknologi* dunia dan banyak aspek komunikasi juga proses komputer telah mengalami banyak perubahan serta perkembangan dan berharap akan lebih banyak jejaring sosial di *internet* [1]. Jejaring sosial adalah media yang digunakan pengguna sebagai tempat untuk menampilkan diri melalui dialog dengan pengguna lain, melakukan kegiatan kerja sama, dan menyelesaikan berbagai masalah yang muncul di jejaring sosial [1]. *Twitter* adalah jejaring sosial yang *independen* dimana semua pengguna dapat mengekspresikan pendapat mereka dengan lebih bebas data *twitter* adalah data yang sangat kuat dan dapat mempengaruhi pengguna di Indonesia [1]. Salah satu keunggulan situs jejaring sosial *twitter* adalah menyediakan *API* yang hebat sehingga siapa pun dapat dengan mudah mendapatkan data dari *twitter* [1].

Salah satu topik yang ramai diperbincangkan masyarakat umum khususnya mahasiswa dan perguruan tinggi Indonesia di *twitter* adalah program Merdeka Belajar Kampus Merdeka, kampus merdeka sangat relevan dan dinilai baik diterapkan di era demokrasi saat ini [2]. Konsep kemandirian dapat diterapkan dalam proses perkuliahan di perguruan tinggi, mahasiswa dapat memilih dari delapan program MBKM yang telah disetujui oleh Pemerintahan Pendidikan dan Kebudayaan, mahasiswa dibebaskan dan diberi kesempatan untuk memilih dan mengikuti mata kuliah non-kurikulum yang diambil selama satu semester (20 sks) [2].

Ada beberapa permasalahan selama program MBKM berjalan dari beberapa perguruan tinggi, serta kendala yang dialami oleh mahasiswa selama mengikuti program ini yaitu masalah implementasi kurikulum masih belum jelas, belum memiliki acuan berupa aturan, buku panduan, petunjuk pelaksanaan, dan prosedur operasional selama kegiatan berlangsung. Mitra yang bekerja sama dengan program ini rata-rata berada di Ibukota, hal ini menimbulkan dampak bagi mahasiswa di perguruan tinggi daerah mengalami kesulitan dalam biaya transportasi, biaya batuan hidup serta tempat tinggal. Keterlambatan pemberian dana bantuan biaya hidup yang tidak sesuai dengan jadwal seharusnya menjadi hambatan untuk mahasiswa yang merantau. Kebijakan setiap mitra yang berbeda dari segi waktu kegiatan, tugas, dan uang saku atau biaya bantuan hidup. Dan adanya pandemi Covid-19, karena ada beberapa aktivitas seperti pembelajaran ke lapangan harus di desain ke arah *virtual* dan kurikulum harus tetap disusun sejalan dengan kebutuhan pemerintah maupun industri [2].

Salah satu cara untuk mengetahui opini publik terhadap program MBKM adalah dengan menganalisis dan merangkum opini yang ada di jejaring sosial *twitter*. Jejaring ini berisi informasi dan opini tentang banyak topik [3]. Pada penelitian ini, *naive bayes* dikombinasikan dengan metode *particle swarm optimization* sebagai *optimize weight* yang diterapkan pada proses *klasifikasi* pengguna *twitter* dalam pelaksanaan program MBKM. Dalam melakukan penelitian tahapan yang harus dilakukan adalah *Studi*

lapangan, *Studi Pustaka, Crawling data, Ekstrasi Fitur, Klasifikasi, dan Evaluasi* [4], [5], [6].

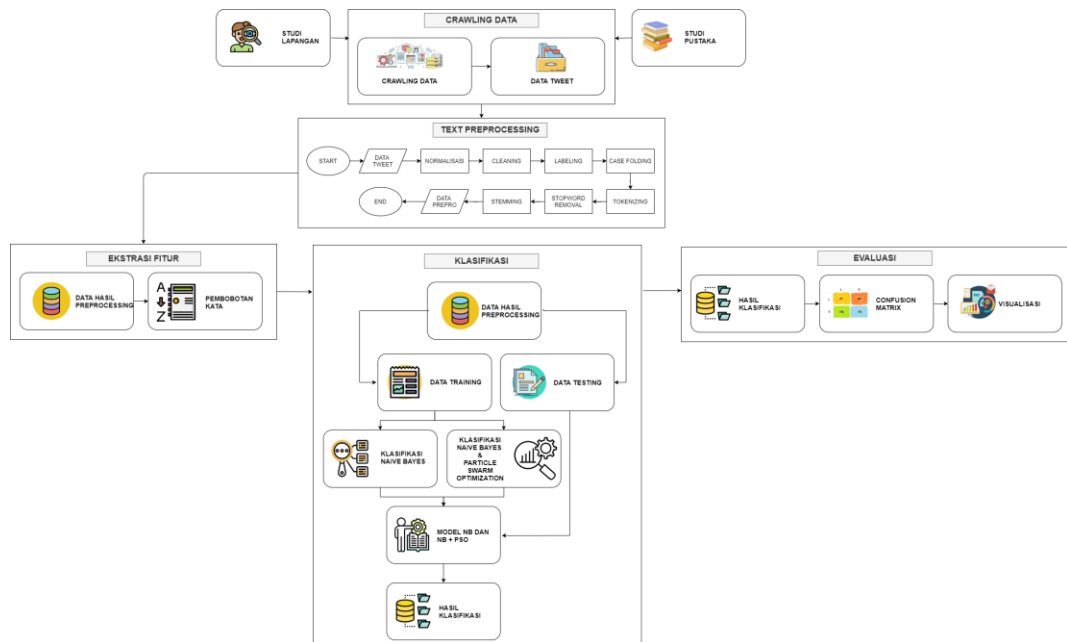
Studi lapangan adalah proses mengamati secara langsung kegiatan yang diteliti, mengamatinya di lingkungan kampus atau di jejaring sosial *Twitter*, dan mengidentifikasi masalah apa yang terjadi [7]. *Studi pustaka* adalah proses mencari dan membaca beberapa *referensi* dari berbagai sumber dari website dan jurnal penelitian sebelumnya yang terkait dengan topik *Text Mining, Analisis Sentimen, Naive Bayes, Particle Swarm Optimization*, dan MBKM [7]. *Crawling data* adalah proses untuk pengunduhan atau pengambilan data dalam format pengguna dan *tweet* menggunakan *API Twitter* [8]. *Text Preprocessing* adalah proses transformasi dokumen data tidak terstruktur yang mengandung banyak *noise*, menjadi data terstruktur yang memudahkan peneliti untuk menganalisis [9]. *Ekstrasi Fitur* merupakan faktor penting, salah satunya dapat mempengaruhi tingkat akurasi pada tahap klasifikasi. *Ekstrasi Fitur* yang digunakan dalam penelitian ini adalah *TF-IDF*, metode ini bekerja dengan cara menghitung bobot setiap kata yang digunakan untuk menghitung bobot setiap kata dalam dokumen, metode ini menghitung kemunculan sebuah kata dalam dokumen [10].

Klasifikasi adalah proses persiapan pencarian model yang menggambarkan data dan mengklasifikasikannya ke kelas-kelas yang ada. *Naive Bayes* adalah metode *klasifikasi* yang biasa digunakan karena keunggulan kecepatan dan akurasinya, *naive bayes* dianggap berpotensi sebagai metode *klasifikasi* data terbaik dalam hal akurasi dan perhitungan [11]. *Particle Swarm Optimization* memiliki beberapa teknik optimasi, antara lain meningkatkan bobot atribut, pemilihan atribut, pemilihan fitur dari variabel yang digunakan. Keuntungan dari metode optimasi *PSO* ini adalah mudah diimplementasikan, sederhana, dan cepat [12]. *Evaluasi* adalah tahapan akhir dalam penelitian ini, proses evaluasi dilakukan menggunakan metode perhitungan akurasi yaitu *confusion matrix* untuk mengukur nilai *accuracy, precision, recall, specificity, positive predictive value, negative predictive value, false negative rate, false positive rate*, dan *f-score* [13]. *Visualisasi* membantu penggunaannya dalam melakukan analisa dan penalaran tentang informasi dan bukti sehingga informasi yang kompleks menjadi lebih mudah untuk dipahami dan dimengerti [5].

Berdasarkan penjelasan di atas, maka penelitian ini bertujuan untuk mengetahui bagaimana sentimen atau opini pengguna *twitter* terhadap pelaksanaan program MBKM di Indonesia melalui postingan dan komentar di jejaring sosial *twitter*, kemudian data tersebut diklasifikasikan menjadi dua jenis *klasifikasi* yaitu *Positive* dan *Negative* menggunakan metode *Naive Bayes* berbasis *Particle Swarm Optimization* untuk meningkatkan akurasi dan melakukan optimalisasi pada bobot atribut [10].

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode *klasifikasi* menggunakan *Naive Bayes berbasis PSO*. Penelitian ini terbagi menjadi lima tahapan yaitu *Crawling Data*, *Text Preprocessing*, *Ekstrasi Fitur*, *Klasifikasi*, dan *Evaluasi*. Metodologi penelitian bisa dilihat pada Gambar 1 yang di adopsi dari penelitian [4],[5],[6].



Gambar 1. Metodologi Penelitian

Berikut adalah tahapan yang dilakukan dalam penelitian ini dijelaskan sebagai berikut.

- Studi Lapangan* proses untuk melakukan pengamatan secara langsung melalui teman-teman yang mengikuti program MBKM di lingkungan kampus dan melakukan pengamatan di media sosial *Twitter* untuk mendapatkan sebuah informasi dan permasalahan apa saja yang terjadi selama mengikuti pelaksanaan program MBKM.
- Studi Pustaka* proses untuk melakukan pengumpulan data sebagai *referensi* untuk mendukung penelitian melalui jurnal, artikel, yang berhubungan dengan Analisis Sentimen, *Klasifikasi*, MBKM, *Naive Bayes*, dan *Particle Swarm Optimization*.
- Crawling Data* proses untuk melakukan pengunduhan dan pengambilan data menggunakan *Twitter API*, baik berupa data nama pengguna ataupun data postingan pengguna. *Crawling* data dilakukan dengan menggunakan *tools Google Colaboratory* dengan menggunakan bahasa pemrograman *Python*.
- Text Preprocessing* proses untuk mengubah dokumen teks yang sudah di dapatkan di proses *crawling*, data yang didapatkan tidak terstruktur setelah melalui tahapan ini data dirubah menjadi lebih terstruktur dan

dapat digunakan untuk proses selanjutnya. *Text preprocessing* memiliki beberapa proses yaitu dokumen data *tweet* dilakukan proses *normalisasi, cleaning, labeling, case folding, tokenizing, stopword removal, stemming*. *Text Processing* dilakukan dengan menggunakan tools *Microsoft Excel* untuk proses *normalisasi* dan *labeling*, *RapidMiner* untuk proses *cleaning*, dan *Google Colaboratory* untuk proses *case folding, tokenizing, stopword removal*, dan *stemming* menggunakan bahasa pemrograman *Python*.

- e) *Ekstrasi Fitur* proses untuk mempermudah klasifikasi, data yang sudah di dapatkan dalam tahapan *text preprocessing* dalam tahapan ini dilakukan pembobotan kata menggunakan *TF-IDF*. *Ekstrasi Fitur* dilakukan dengan menggunakan tools *Microsoft Excel*.
- f) *Klasifikasi* proses untuk pembuatan model, data yang sudah dilakukan tahapan *text preprocessing* maka dilakukan pembagian data menjadi data *training* dan data *testing*. Data *training* akan dilakukan proses *klasifikasi* menggunakan metode *Naive Bayes* dan *Naive Bayes* berbasis *Particle Swarm Optimization*, model data *training* sudah didapatkan maka akan dilakukan pengujian model menggunakan data *testing* dan mendapatkan hasil *klasifikasi* tersebut. *Klasifikasi* dilakukan dengan menggunakan tools *RapidMiner*.
- g) *Evaluasi* proses akhir dalam penelitian, hasil *klasifikasi* yang sudah di dapatkan akan dilakukan pengujian akurasi menggunakan *confusion matrix*, dan akan dilakukan proses *visualisasi*. *Evaluasi* untuk pengujian akurasi dilakukan dengan menggunakan tools *Microsoft Excel*, dan *visualisasi* dilakukan dengan menggunakan tools *PowerBI*.

3. HASIL DAN PEMBAHASAN

Pada point ini membahas hasil analisis berdasarkan pengolahan data yang telah dilakukan pada jejaring sosial *Twitter* yang memuat postingan atau komentar mengenai program MBKM dimulai dari tahapan *Crawling Data, Text Preprocessing, Ekstrasi Fitur, Klasifikasi, dan Evaluasi*.

3.1. *Crawling Data*

Crawling dilakukan pada bulan februari, maret, dan april 2022, dalam proses *crawling* pertama peneliti harus memiliki akun di *Twitter API* dan melakukan *registrasi token API*. Setelah mendapatkan *token api key, api secret key, access token*, dan *access token secret*, selanjutnya adalah melakukan *import library* yang dibutuhkan untuk proses *crawling*, lalu masukan *token api tadi*, dan *keyword "MBKM"*. Semua sudah berhasil dimasukan maka data *twitter* dengan *keyword* tadi akan terunduh, data yang sudah diunduh disimpan dalam bentuk *csv* disajikan pada link berikut: <https://github.com/erinaundamayanti/Dataset/blob/main/Dataset/DataCrawling.xlsx>

3.2. Text Preprocessing

Sebelum melakukan proses *klasifikasi* dokumen data *tweet* harus melewati tahapan *text preprocessing* untuk melakukan *transformasi* data, data *tweet* akan dirubah menjadi data *tweet* yang terstruktur agar lebih mudah untuk dilakukan analisis.

3.2.1. Normalisasi

Proses normalisasi dilakukan untuk menghapus atribut yang tidak digunakan dan hanya mengambil atribut berisi *tweet* yang diunggah oleh pengguna *Twitter* [14]. Contoh data *Tweet* hasil *crawling* yang digunakan dalam penelitian dan belum dilakukan proses *normalisasi*, disajikan pada tabel 1.

Tabel 1. Data Hasil *Crawling*

No	Username	Komentar
0	zakka_f	@siti mbkm energi emang asik
1	lynntahh	bisa-bisanya mbkm lotre wkkwkw https://t.co/dhdiDX9gUU
2	Jvxcxwh	Mbkm magang ini rollercoaster banget asu!

Data hasil *crawling* memiliki tiga atribut yaitu nomor, username, dan komentar. Penelitian ini hanya membutuhkan satu atribut yaitu komentar yang berisi postingan atau *tweet* yang diunggah oleh pengguna *twitter*. Setelah dilakukan proses *normalisasi* data hasil *crawling* disajikan pada tabel 2.

Tabel 2. Hasil *Normalisasi*

komentar
@siti mbkm energi emang asik
bisa-bisanya mbkm lotre wkkwkw https://t.co/dhdiDX9gUU
Mbkm magang ini rollercoaster banget asu!

3.2.2. Cleaning

Proses *cleaning* dilakukan untuk menghapus kata-kata yang tidak diinginkan seperti karakter HTML, simbol, tagar, nama pengguna, URL, email [15]. Bisa dilihat pada tabel 3 komentar yang didapatkan masih terdapat simbol dan URL. Setelah dilakukan proses *cleaning* data disajikan pada tabel 3.

Tabel 3. Hasil *Cleaning*

komentar
mbkm energi emang asik
bisa-bisanya mbkm lotre wkkwkw
Mbkm magang ini rollercoaster banget asu

3.2.3. Labeling

Proses *labeling* dilakukan untuk mengidentifikasi emosi dari setiap *tweet*, *positive* atau *negative*. *Tweet Positive* adalah komentar yang mendukung, merekomendasikan, memberikan saran dan berkontribusi atau

mengikuti program MBKM ini [16]. *Tweet Negative* adalah komentar yang berisi kata-kata kasar, hinaan, cacian, makian, dan tuduhan yang menyinggung mengenai program MBKM ini. Bisa dilihat pada tabel 3 komentar yang didapatkan masih belum diberikan label, setelah dilakukan proses labeling data disajikan pada tabel 4.

Tabel 4. Hasil *Labeling*

komentar	sentimen
mbkm energi emang asik	Positif
bisa-bisanya mbkm lotre wkkwkw	Positif
Mbkm magang ini rollercoaster banget asu	Negatif

3.2.4. Case Folding

Proses *case folding* dilakukan untuk mengubah semua teks atau kalimat dalam data *tweet* menjadi huruf kecil (*lower case*) sehingga memudahkan untuk mencocokkan dokumen dalam proses penyamaan sebuah dokumen [16]. Bisa dilihat pada tabel 5 komentar yang didapatkan masih terdapat teks yang menggunakan kapital didalamnya, setelah dilakukan proses *case folding* data disajikan pada Gambar 2.

1	mbkm energi emang asik	Positif	mbkm energi emang asik
2	bisa-bisanya mbkm lotre wkkwkw	Positif	bisa-bisanya mbkm lotre wkkwkw
3	Mbkm magang ini rollercoaster banget asu	Negatif	mbkm magang ini rollercoaster banget asu

Gambar 2. Hasil *Case Folding*

3.2.5. Tokenizing

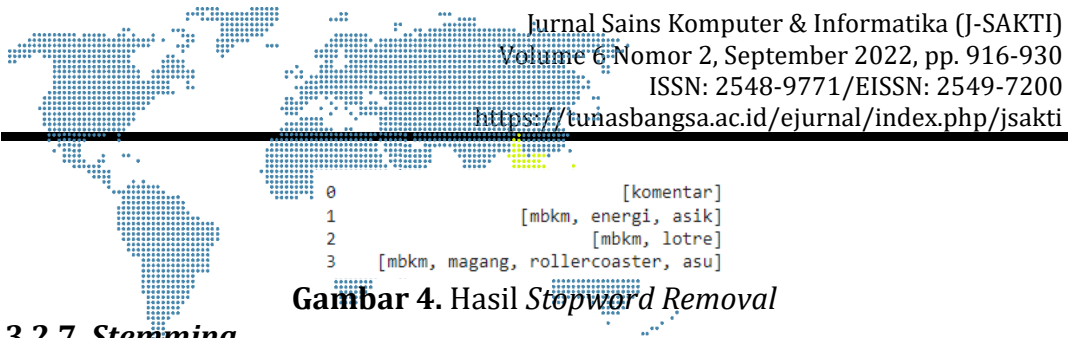
Proses *tokenizing* dilakukan untuk memecah kumpulan kata dalam kalimat, paragraf, halaman menjadi sebuah token, atau sebuah kata yang disingkat [16]. Pada proses *tokenizing*, semua kata dalam setiap dokumen dipisahkan dan menghapus apapun yang terdeteksi bukan huruf. Bisa dilihat pada Gambar 2 komentar yang didapatkan masih belum terpecah kumpulan kata nya, setelah dilakukan proses *tokenizing* data disajikan pada Gambar 3.

1	mbkm energi emang asik	Positif	mbkm energi emang asik	[mbkm, energi, emang, asik]
2	bisa-bisanya mbkm lotre wkkwkw	Positif	bisabisanya mbkm lotre wkkwkw	[bisabisanya, mbkm, lotre, wkkwkw]
3	Mbkm magang ini rollercoaster banget asu	Negatif	mbkm magang ini rollercoaster banget asu	[mbkm, magang, ini, rollercoaster, banget, asu]

Gambar 3. Hasil *Tokenizing*

3.2.6. Stopword Removal

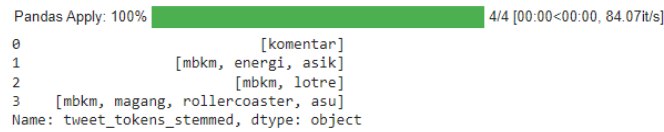
Proses *stopword removal* dilakukan untuk mengidentifikasi istilah umum atau kata-kata yang tidak relevan dengan penelitian akan dihapus atau dihilangkan [16]. Bisa dilihat pada Gambar 3 komentar yang didapatkan masih mengandung kata-kata yang tidak relevan seperti “wkkwkw”, “emang”, “ini”, “banget” , setelah dilakukan proses *tokenizing* data disajikan pada Gambar 4.



Gambar 4. Hasil *Stopword Removal*

3.2.7. Stemming

Proses *stemming* dilakukan untuk menghilangkan *sufiks* dan mengembalikannya ke kata dasar [16]. Setelah dilakukan proses *stemming* data disajikan pada Gambar 5.



Gambar 5. Hasil *Stemming*

Data keseluruhan yang sudah dilakukan tahapan *text preprocessing* dari 2702 data menjadi 428 data disajikan pada link berikut:
<https://github.com/erinaundamayanti/Dataset/blob/main/Dataset/DataTextPreprocessing.xlsx>

3.3. Ekstrasi Fitur

Dalam tahapan *ekstrasi fitur* melakukan pembobotan kata menggunakan metode *TF-IDF* dapat dirumuskan seperti dibawah ini [10].

$$IDF (Word) = \log \frac{td}{df} \quad (1)$$

$$TF - IDF = TF \times IDF \quad (2)$$

Dokumen yang digunakan dalam proses *TF-IDF* disajikan pada tabel 5.

Tabel 5. Dokumen *TF-IDF*

D1	mbkm energi asik
D2	mbkm lotre
D3	mbkm magang rollercoaster asu

Setelah melakukan tahapan *text preprocessing* selanjutnya adalah mencari frekuensi kemunculan (*tf*) suatu kata dalam suatu dokumen, dengan cara menghitung jumlah kata yang muncul pada dokumen. Hasil *tf* disajikan pada tabel 6.

Tabel 6. Hasil *Term Frequency*

Kata	TF		
	D1	D2	D3
mbkm	1	1	1
energi	1	0	0
asik	1	0	0
lotre	0	1	0
magang	0	0	1
rollercoaster	0	0	1
asu	0	0	1

Selanjutnya adalah menghitung jumlah frekuensi kata terpilih yang sudah diketahui kemudian menghitung jumlah dokumen (*idf*), dimana *idf* merupakan hasil *inverse* dari *df*. Hasil *idf* disajikan pada tabel 7.

Tabel 7. Hasil Inverse Document Frequency

Kata	TF			DF	TD/DF	IDF
	D1	D2	D3			
mbkm	1	1	1	3	1	0
energi	1	0	0	1	3	0.48
asik	1	0	0	1	3	0.48
lotre	0	1	0	1	3	0.48
magang	0	0	1	1	3	0.48
rollercoaster	0	0	1	1	3	0.48
asu	0	0	1	1	3	0.48

Setelah nilai *idf* didapatkan selanjutnya adalah mencari nilai *TF-IDF* yang merupakan hasil perkalian antara *tf* dan *idf*. Hasil *TF-IDF* disajikan pada tabel 8.

Tabel 8. Hasil TF-IDF

TF*IDF		
D1	D2	D3
0	0	0
0.48	0	0
0.48	0	0
0	0.48	0
0	0	0.48
0	0	0.48
0	0	0.48

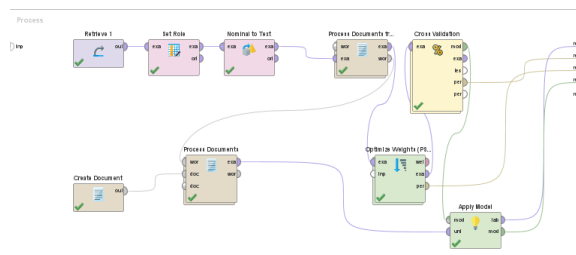
Metode ini berfungsi untuk mencari representasi nilai dari masing-masing dokumen dari sekumpulan data, dari tabel diatas menghasilkan jumlah keseluruhan *TF-IDF* dari tiga dokumen *tweet* disajikan pada tabel 9.

Tabel 9. Jumlah TF-IDF

D1	D2	D3
0.95	0.48	1.44

3.4. Klasifikasi

Model klasifikasi *Naive Bayes* dan *Naive Bayes + PSO* disajikan pada Gambar 6.



Gambar 6. Model Klasifikasi NB dan NB + PSO

Pada Gambar 6 proses diawali dengan *Retrieve* yang berfungsi untuk menginputkan data *tweet* yang sudah dilakukan proses *Text Preprocessing* ke dalam *RapidMiner*. Setelah itu, dilakukan *set Role* yaitu untuk membedakan baris penamaan atribut dan prediksi posisi yang akan di masukan kedalam kategori label untuk kolom sentimen, agar pada saat pengkategorian data label tidak ikut serta terhitung dan merubah hasil. *Nominal to Text* berfungsi untuk mengubah jenis atribut nominal yang dipilih menjadi teks dan juga berfungsi untuk memetakan semua nilai atribut ke nilai *string* yang sesuai. *Process Document from Data* berfungsi untuk menghasilkan vektor kata dari atribut *string*.

Create Document membuat dokumen yang berisi teks dijadikan sebagai parameter atau dijadikan sebagai data *testing* untuk mengetahui prediksi data tersebut termasuk kedalam kelas positif atau negatif. *Process Document* menggunakan objek teks tunggal sebagai *input* untuk menghasilkan suatu vektor, operator ini sangat berguna untuk menerapkan model pada suatu teks. Selanjutnya dilakukan *Cross Validation*, didalam *validation* terdapat dua proses yaitu *training* dan *testing*, *training* digunakan untuk melatih model dan *testing* untuk pengujian sekaligus mengukur kinerja model. *Optimize Weight* digunakan untuk penerapan *PSO* untuk mengoptimalisasi hasil *klasifikasi*. *Apply Model* digunakan untuk menerapkan model yang telah dilatih sebelumnya menggunakan data *training* pada data *testing*. Berikut adalah hasil *klasifikasi* berupa tabel *confusion matrix* dari model yang sudah dibuat disajikan pada Gambar 7 dan 8.

accuracy: 69.64% +/- 4.95% (micro average: 69.63%)

	true Positif	true Negatif	class precision
pred. Positif	184	49	78.97%
pred. Negatif	81	114	58.46%
class recall	69.43%	69.94%	

Gambar 7. Hasil Klasifikasi Naive Bayes

accuracy: 71.93% +/- 6.66% (micro average: 71.96%)

	true Positif	true Negatif	class precision
pred. Positif	177	32	84.69%
pred. Negatif	88	131	59.82%
class recall	66.79%	80.37%	

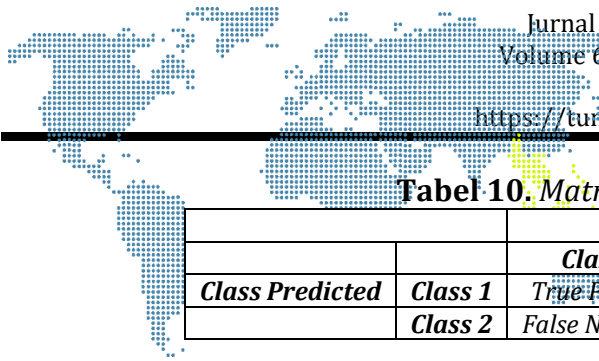
Gambar 8. Hasil Klasifikasi Naive Bayes + PSO

3.5. Evaluasi

Tahapan ini adalah untuk melakukan proses pengujian akurasi menggunakan *confusion matrix* dan *visualisasi*.

3.5.1. Confusion Matrix

Evaluasi dilakukan menggunakan metode perhitungan akurasi yaitu *confusion matrix* untuk mengukur nilai *accuracy*, *precision*, *recall*, *specificity*, *positive predictive value*, *negative predictive value*, *false negative rate*, *false positive rate*, dan *f-score*. Berikut adalah tabel *Confusion matrix* yang terdiri dari empat sel disajikan pada tabel 10 [18].



Tabel 10. *Matrix Confusion*

		Class True	
		Class 1	Class 2
Class Predicted	Class 1	True Positive	False Positive
	Class 2	False Negative	True Negative

Berikut adalah rumus dalam pengukuran *evaluasi performance* berdasarkan tabel *confusion matrix*.

$$\text{Accuracy} = \frac{TN+TP}{TP+TN+FP+FN} \times 100\% \quad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (5)$$

$$\text{SPEC} = \frac{TN}{TN+FP} \times 100\% \quad (6)$$

$$\text{PPV} = \frac{TP}{TP+FP} \times 100\% \quad (7)$$

$$\text{NPV} = \frac{TN}{TN+FN} \times 100\% \quad (8)$$

$$\text{FNR} = \frac{FN}{FN+TP} \times 100\% \quad (9)$$

$$\text{FPR} = \frac{FP}{FP+TN} \times 100\% \quad (10)$$

$$\text{F - Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

Keterangan:

- Accuracy = Kinerja klasifikasi secara keseluruhan
- Precision = Rasio prediksi benar positif dari hasil prediksi positif
- Recall = Persentase prediksi benar positif dari populasi
- Specificity = Kebenaran prediksi negatif dengan keseluruhan data negatif
- PPV = Rasio prediksi benar positif
- NPV = Rasio prediksi benar negatif
- FNR = Rasio positif yang diketahui hasilnya negatif
- FPR = Rasio negatif yang diketahui hasilnya positif

Setelah diuji dengan perhitungan akurasi maka peneliti mendapatkan hasil pengujian *klasifikasi naive bayes* dan *naive bayes+ pso* dari pengujian menggunakan *confusion matrix* disajikan pada tabel 11.

Tabel 11. Hasil *Confusion Matrix NB dan NB+PSO*

	NB	NB + PSO
Accuracy	69.63%	71.96%
Precision	68.71%	72.25%
Recall	69.68%	73.58%
Spec	69.94%	80.37%
PPV	78.97%	84.69%
NPV	58.46%	59.82%
FNR	30.57%	33.21%
FPR	30.06%	19.63%
F-Score	68.91%	71.63%

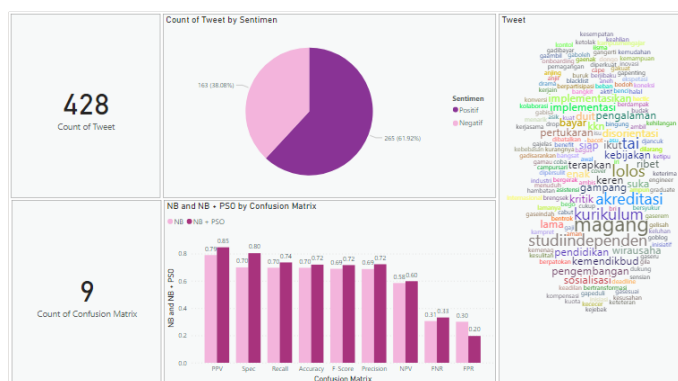
Kinerja pengklasifikasian secara keseluruhan menggunakan metode naive bayes sebesar 69.63% setelah menerapkan metode untuk melakukan optimalisasi menggunakan particle swarm optimization akurasi pengklasifikasian meningkat sebesar 2.33% menjadi 71.96% hasil tersebut menunjukan bahwa dengan menerapkan pso kinerja pengklasifikasian menjadi semakin baik. Rasio untuk memprediksi benar positif dari populasi hasil prediksi positif secara keseluruhan menggunakan naive bayes didapatkan hasil sebesar 68.71% setelah menerapkan pso mendapatkan hasil sebesar 72.25%, hasil tersebut menunjukan bahwa hasil prediksi positif dari hasil prediksi secara keseluruhan lebih baik setelah menerapkan metode pso. Persentase hasil prediksi yang benar positif dari populasi menggunakan naive bayes didapatkan hasil sebesar 69.68% setelah menerapkan pso mendapatkan hasil sebesar 73.58%.

Kebenaran untuk memprediksi negatif dibandingkan dengan keseluruhan data negatif menggunakan naive bayes didapatkan hasil sebesar 69.94% setelah menerapkan pso mengalami kenaikan signifikan menjadi 80.37%. Rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif menggunakan naive bayes didapatkan hasil sebesar 78.97% setelah menerapkan pso mendapatkan hasil sebesar 84.69%. Rasio prediksi benar negatif dibandingkan dengan keseluruhan hasil yang diprediksi negatif menggunakan naive bayes didapatkan hasil sebesar 58.46% setelah menerapkan pso mendapatkan hasil sebesar 59.82%.

Miss rate atau rasio positif yang diketahui hasilnya negatif menggunakan naive bayes didapatkan hasil sebesar 30.57% setelah menerapkan pso mendapatkan hasil sebesar 33.21%. Rasio negatif yang diketahui hasilnya positif menggunakan naive bayes didapatkan hasil sebesar 30.06% sedangkan setelah menerapkan pso mendapatkan hasil sebesar 19.63%. Dan terakhir adalah perbandingan rata-rata antara precision dan recall menggunakan naive bayes didapatkan hasil sebesar 68.91% setelah menerapkan pso mendapatkan hasil sebesar 71.63%.

3.5.2. Visualisasi

Berikut ini merupakan visualisasi berupa dashboard disajikan pada Gambar 9.



Gambar 9. Visualisasi Dashboard

Pada Gambar 9 menunjukkan jumlah tweet yang dilakukan dalam penelitian ini adalah 428 *tweet*, 61.92% berkategori positif dan 38.08% berkategori negatif. Dalam penelitian ini pengguna twitter terutama mahasiswa memberikan *opini* atau informasi mengenai program MBKM ini cenderung lebih kearah yang positif dibandingkan kearah yang negatif. Dalam penelitian ini kategori sentimen yang dihasilkan lebih banyak kategori positif yang mengandung kata positif “lolos”, “akreditasi”, “kurikulum”, “magang”, “studiindependen”. Dan kata negatif “duit”, “lama”, “disorientasi”, “ribet”.

Pengujian *klasifikasi confusion matrix* dengan melakukan 9 perhitungan menggunakan metode *naive bayes* dan *naive bayes* berbasis *particle swarm optimization* lebih baik menggunakan metode *naive bayes* berbasis *particle swarm optimization* karena setiap pengujian yang menerapkan metode *pso* mendapatkan hasil yang lebih baik, hanya saja dalam perhitungan *False Positive Rate* atau rasio negatif yang diketahui hasilnya positif setelah diterapkan metode *pso* mendapatkan hasil yang lebih kecil hanya 20% dibandingkan menggunakan metode *naive bayes*.

4. SIMPULAN

Hasil penelitian dengan menggunakan 428 data *tweet* di klasifikasikan kedalam sentimen positif dan negatif, dan hasil penelitian menghasilkan sentimen yang positif sekitar 61.92% jadi dapat disimpulkan bahwa program MBKM ini dapat diterima dengan baik oleh masyarakat pengguna *Twitter* terutama mahasiswa. Walaupun ada beberapa sentimen negatif yang muncul sekitar 38.08% memberikan opini bahwa program MBKM yang di duga bersifat disorientasi, ribet dalam proses konversi dibeberapa fakultas, dan masalah mengenai keterlambatan uang saku yang diberikan.

Adapun hasil penelitian ini bisa dijadikan rujukan untuk tim pembuat pengembangan kebijakan MBKM terutama tim POKJA Kemendikbud, karena program ini bisa memberikan manfaat dan pengalaman yang tidak didapatkan mahasiswa di lingkungan kampus, serta memberikan pengaruh terhadap akreditasi kampus yang mahasiswanya ikut berpartisipasi dalam program MBKM. Adapun dalam hasil penelitian ini bisa dijadikan bahan evaluasi untuk tim kedepannya dari sisi prosedur sistem konversi sks setiap fakultas yang masih tidak sesuai bahkan ada yang tidak bisa dikonversi karena tidak sesuai dengan mata kuliah pilihan yang diambil oleh mahasiswa, serta waktu pembayaran uang saku atau biaya bantuan hidup yang dalam proses pembayarannya ini masih mengalami kendala dan keterlambatan dalam proses pendistribusian dana tersebut.

DAFTAR PUSTAKA

- [1]. Nugroho, A. “Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstraksi Fitur N-Gram”, Jurnal Sains Komputer Dan Informatika, Vol. 2, No. 2, <https://doi.org/10.30645/j-sakti.v2i2.83>, pp 200, 2021.



- [2]. Fuadi, T. M., & Aswita, D. "Merdeka Belajar Kampus Merdeka (MbkM): Bagaimana Penerapan Dan Kedala Yang Dihadapi Oleh Perguruan Tinggi Swasta Di Aceh", *Jurnal Dedikasi Pendidikan*, Vol. 5, No. 2, <http://jurnal.abulyatama.ac.id/index.php/dedikasi>, pp 603-614, 2021.
- [3]. Wandani, A., Fauziah, F., & Andrianingsih, A. "Sentimen Analisis Pengguna Twitter pada Event Flash Sale Menggunakan Algoritma K-NN, Random Forest, dan Naive Bayes", *Jurnal Sains Komputer Dan Informatika*, Vol. 5, No. 2, pp 651-665, 2021.
- [4]. Prabowo, W. A., & Wiguna, C. "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM", *Jurnal Media Informatika Budidarma*, Vol. 5, No. 1, <https://doi.org/10.30865/mib.v5i1.2604>, pp 149, 2021 .
- [5]. Taofik Krisdiyanto, E. M. O. N. "Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naive Bayes Clasifiers", *Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, Vol. 7, No. 1, <http://ejournal.uin-suska.ac.id/index.php/coreit/article/view/12945>, pp 32-37, 2021.
- [6]. Eni Tri Handayani, Ari Sulistiyawati, "Analisis Sentimen Respon Masyarakat Terhadap Kabar Harian Covid-19 Pada Twitter Kementerian Kesehatan Dengan Metode Klasifikasi Naive Bayes", *Jurnal Teknologi dan Sistem Informasi*, Vol. 2, No. 3, <http://jim.teknokrat.ac.id/index.php/JTSI>, pp 32-37, September 2021.
- [7]. Aziz, A., Fauziah, F., & Fitri, I. "Analisis Sentimen Terhadap Kebijakan Pemerintah Tentang Larangan Mudik Hari Raya Idulfitri di Indonesia Tahun 2021 Menggunakan Metode Naive Bayes", *Jurnal Sains Komputer Dan Informatika*, Vol. 5, No. 2, <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/381>, pp 842-851, 2021.
- [8]. Normawati, D., & Prayogi, S. A. "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter", *Jurnal Sains Komputer & Informatika*, Vol. 5, No. 2, pp 697-711, 2021.
- [9]. Ahmadi, M. I., Gustian, D., & Sembiring, F. "Analisis Sentiment Masyarakat terhadap Kasus Covid-19 pada Media Sosial Youtube dengan Metode Naive bayes" *Jurnal Sains Komputer & Informatika*, Vol. 5, No. 2, pp 807-814, 2021.
- [10]. Rizki Aulianita, Achmad Rifai, "Optimasi Particle Swarm Optimization Pada Naive Bayes Untuk Sentiment Analysis Furniture", *Information Management For Educators And Professionals*, Vol. 3, No. 1, pp 31-40, Desember 2018.
- [11]. Utami, D. S., & Erfina, A. "Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes Pada dasarnya Indonesia", *Jurnal Sains Komputer & Informatika*, Vol. 6, No. 1, pp 418-427, 2022.
- [12]. Suwanda Aditya Saputra, Didi Rosiyadi, Windu Gata, Syepry Maulana Husain, "Analisis Sentimen E-Wallet Pada Google Play Menggunakan



- Algoritma Naive Bayes Berbasis Particle Swarm Optimization”, Jurnal Resti, Vol. 3, No.3, <http://jurnal.iaii.or.id/>, pp 377-382, 2021.
- [13]. Aziz, A. “Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength”. Jurnal Sains Komputer & Informatika, Vol. 6, No. 1, pp 115, 2022.
- [14]. Li, Z., Li, R., & Jin, G., “Sentiment Analysis Of Danmaku Videos Based On Naive bayesAnd Sentiment Dictionary”, IEEE Access, 8, <https://doi.org/10.1109/ACCESS.2020.2986582>, pp 75073–75084, 2020.
- [15]. Ernawati, S., Yulia, E. R., Friyadie, & Samudi, “Implementation of the Naive bayesAlgorithm with Feature Selection using Genetic Algorithm for Sentiment Review Analysis of Fashion Online Companies”, 6th International Conference on Cyber and IT Service Management, <https://doi.org/10.1109/CITSM.2018.8674286>, Citsm, 1–5, 2019.
- [16]. Retno Sari, Ratih Yulia Hayuningtyas, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website”, Indonesian Journal on Software Engineering, Vol. 5, No. 2, pp 51-60, Desember 2019.
- [17]. Ibrahim, M. N., & Yusoff, M. Z., “The Impact of Different Training Data Set on the Accuarcy of Sentiment Classification of Naive Bayes Technique”, IEEE Conference on Open Systems (ICOS), Miri Marriott Resort & SPA, Miri, Sarawak, Malaysia, pp 17-20, November 2021.
- [18]. Lisda Widiastuti, “Pemilihan Fitur Pada Analisis Sentimen Review Travel Online Menggunakan Algoritma Naive Bayes Dalam Penerapan Mutual Information Dan Particle Swarm Optimization (PSO)”, Indonesian Journal on Computer and Information Technology, Vol. 3, No.1, pp 91-100, Mei 2018.