

Perbandingan Metode Naive Bayes, Random Forest dan SVM Untuk Analisis Sentimen Pada Twitter Tentang Kenaikan Gaji Guru

Eny Yuniar, Nirwana Hendrastuty*

Fakultas Teknik dan Ilmu Komputer, Program Studi Sistem Informasi, Universitas Teknokrat Indonesia, Lampung, Indonesia

Email: ¹yuniar_eni@teknokrat.ac.id, ^{2,*}nirwanahendrastuty@teknokrat.ac.id

Email Penulis Korespondensi: nirwanahendrastuty@teknokrat.ac.id

Submitted: 11/02/2025; Accepted: 03/03/2025; Published: 07/03/2025

Abstrak—Kenaikan gaji guru menjadi isu yang banyak diperdebatkan di kalangan masyarakat, dengan berbagai pendapat yang disampaikan melalui media sosial, terutama Twitter. Penelitian ini bertujuan untuk menganalisis sentimen komentar publik mengenai kebijakan kenaikan gaji guru dengan menggunakan tiga algoritma machine learning: Naive Bayes, Random Forest, dan Support Vector Machine (SVM). Data yang digunakan berjumlah 6011 tweet yang dikumpulkan mengenai topik tersebut, yang kemudian diproses menjadi 5531 data setelah tahap pembersihan dan pengolahan. Penelitian ini mengukur kinerja setiap algoritma menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil menunjukkan bahwa SVM memiliki akurasi tertinggi (86%) sebelum penerapan teknik SMOTE, diikuti oleh Random Forest (85%) dan Naive Bayes (84%). Setelah penerapan SMOTE untuk mengatasi ketidakseimbangan data, Random Forest menunjukkan peningkatan kinerja yang signifikan, dengan akurasi mencapai 99%, diikuti oleh SVM (98%) dan Naive Bayes (89%). Hasil ini mengindikasikan bahwa teknik SMOTE dapat secara efektif memperbaiki kinerja model, khususnya dalam mengatasi ketidakseimbangan antara data sentimen positif, negatif, dan netral. Penelitian ini memberikan wawasan baru tentang bagaimana publik merespons kebijakan kenaikan gaji guru, serta memperkenalkan penggunaan SMOTE untuk meningkatkan akurasi dalam analisis sentimen di media sosial.

Kata Kunci: Analisis Sentimen; Kenaikan Gaji Guru; Twitter; SMOTE; Naive Bayes; Random Forest; SVM.

Abstract—The increase in teacher salaries has become a highly debated issue within the community, with various opinions being expressed through social media, particularly Twitter. This study aims to analyze public sentiment regarding the teacher salary increase policy using three machine learning algorithms: Naive Bayes, Random Forest, and Support Vector Machine (SVM). The data used consists of 6010 tweets collected on the topic, which were processed into 5531 data points after cleaning and preprocessing. This study evaluates the performance of each algorithm using accuracy, precision, recall, and F1-score metrics. The results show that SVM achieved the highest accuracy (86%) before applying the SMOTE technique, followed by Random Forest (85%) and Naive Bayes (84%). After applying SMOTE to address data imbalance, Random Forest showed a significant performance improvement, with accuracy reaching 99%, followed by SVM (98%) and Naive Bayes (89%). These results indicate that the SMOTE technique can effectively improve model performance, particularly in handling the imbalance between positive, negative, and neutral sentiment data. This study provides new insights into how the public responds to the teacher salary increase policy, while also introducing the use of SMOTE to enhance accuracy in sentiment analysis on social media.

Keywords: Sentiment Analysis; Teacher Salary Increase; Twitter; SMOTE; Naive Bayes; Random Forest; SVM.

1. PENDAHULUAN

Kebijakan kenaikan gaji guru adalah isu yang sering dibahas dalam masyarakat, karena dampaknya yang luas terhadap kualitas pendidikan dan kesejahteraan tenaga pendidik. Kesejahteraan yang lebih baik diharapkan dapat meningkatkan motivasi dan kinerja guru, yang pada gilirannya berkontribusi pada peningkatan kualitas pengajaran[1]. Namun, meskipun gaji menjadi faktor penting dalam menentukan kesejahteraan guru, masih ada ketimpangan yang signifikan dalam sistem penggajian di berbagai daerah di Indonesia. Ketidakseimbangan ini menyebabkan disparitas dalam kualitas pendidikan antar wilayah, yang menjadi tantangan besar bagi kebijakan pendidikan nasional. Selain itu, ketidaksesuaian antara kebijakan kenaikan gaji dan persepsi publik terhadap kebijakan tersebut juga menjadi masalah yang perlu dianalisis lebih lanjut, terutama dalam konteks masyarakat yang lebih luas melalui media sosial.

Twitter sebagai salah satu platform media sosial yang banyak digunakan oleh publik, memungkinkan masyarakat untuk menyampaikan pendapat secara langsung mengenai berbagai isu, termasuk kebijakan kenaikan gaji guru. Oleh karena itu, analisis sentimen pada komentar-komentar Twitter terkait kebijakan ini dapat memberikan gambaran yang lebih jelas mengenai pandangan masyarakat[2]. Namun, meskipun banyaknya penelitian yang telah menggunakan Twitter untuk menganalisis sentimen terkait isu-isu sosial dan kebijakan publik, penelitian yang secara khusus mengkaji sentimen publik terhadap kebijakan kenaikan gaji guru masih sangat terbatas.

Penelitian terkait analisis sentimen menggunakan algoritma machine learning telah dilakukan oleh berbagai peneliti dalam beberapa tahun terakhir. Misalnya, penelitian [3] membandingkan kinerja algoritma Naive Bayes dan *Long Short-Term Memory* (LSTM) dalam analisis sentimen terhadap kebijakan sosial, dengan hasil yang menunjukkan bahwa LSTM lebih efektif dalam menangkap pola kompleks dalam data teks. Penelitian lainnya yaitu [4] mengaplikasikan algoritma Naive Bayes untuk analisis sentimen terhadap ujian seleksi CPNS di Indonesia, dengan fokus pada efektivitas metode dalam menganalisis komentar publik. Dalam konteks yang lebih luas, penelitian [5] membahas perbandingan antara algoritma Multinomial Naive Bayes dan Support Vector Machine (SVM) untuk analisis sentimen dalam opini publik, dengan hasil yang menunjukkan bahwa SVM unggul dalam hal akurasi. Sementara itu, penelitian [6] membandingkan efektivitas Random Forest dan SVM dalam analisis sentimen terkait pemilu, menunjukkan bahwa SVM memberikan hasil yang lebih konsisten. Terakhir, penelitian [7] mengkaji sentimen

publik terkait bencana gempa megathrust di Indonesia, menggunakan SVM dan Naïve Bayes, dengan fokus pada tantangan ketidakseimbangan data dalam analisis sentimen.

Namun, meskipun terdapat berbagai penelitian terkait analisis sentimen pada berbagai kebijakan publik, masih ada gap yang signifikan dalam penelitian ini, yakni analisis sentimen terhadap kebijakan kenaikan gaji guru. Selain itu, sebagian besar penelitian yang ada belum secara optimal menangani ketidakseimbangan data dalam analisis sentimen. Ketidakseimbangan ini sering kali menjadi masalah utama, di mana kelas sentimen positif mendominasi data, sementara sentimen negatif dan netral kurang terwakili. Oleh karena itu, penelitian ini berfokus pada analisis sentimen terhadap kebijakan kenaikan gaji guru di Indonesia, dengan menggunakan tiga algoritma machine learning yang umum digunakan yaitu Naïve Bayes, Random Forest, dan SVM.

Solusi yang diusulkan dalam penelitian ini adalah penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) untuk menangani masalah ketidakseimbangan data. SMOTE adalah metode yang telah terbukti efektif dalam meningkatkan jumlah data kelas minoritas dengan menghasilkan data sintesis, sehingga dapat membantu model dalam menghasilkan prediksi yang lebih akurat, terutama pada kelas sentimen negatif dan netral[8]. Penelitian ini bertujuan untuk membandingkan efektivitas ketiga algoritma tersebut dalam menganalisis sentimen Twitter terkait kebijakan kenaikan gaji guru, serta mengevaluasi dampak penerapan SMOTE terhadap kinerja model dalam mengklasifikasikan sentimen.

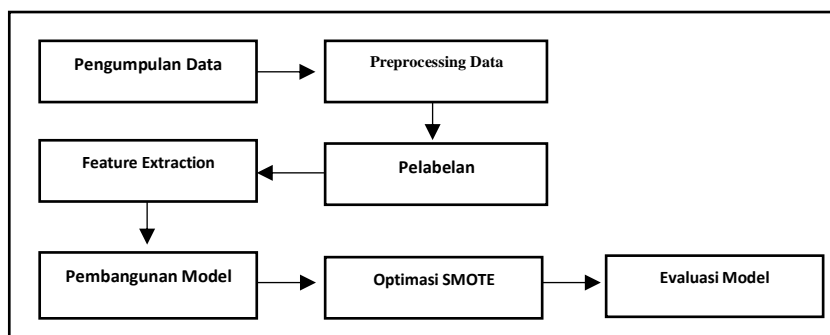
Tujuan utama dari penelitian ini adalah untuk memberikan wawasan yang lebih mendalam mengenai pandangan publik terhadap kebijakan kenaikan gaji guru melalui analisis sentimen pada media sosial, khususnya Twitter. Penelitian ini juga bertujuan untuk mengidentifikasi algoritma yang paling efektif dalam menganalisis sentimen terkait kebijakan pendidikan dan memperkenalkan penerapan SMOTE untuk meningkatkan kinerja model machine learning dalam analisis sentimen, terutama dalam menghadapi ketidakseimbangan data. Diharapkan, hasil penelitian ini dapat memberikan kontribusi dalam pemilihan metode terbaik untuk analisis sentimen di media sosial, serta memberikan rekomendasi yang berguna bagi kebijakan pendidikan di Indonesia.

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan dengan tujuan untuk menganalisis sentimen terhadap komentar Twitter mengenai kebijakan kenaikan gaji guru, serta untuk membandingkan kinerja tiga algoritma machine learning, yaitu Naïve Bayes, Random Forest, dan Support Vector Machine (SVM). Dalam penelitian ini, kami menggunakan pendekatan kuantitatif dengan tahapan yang meliputi pengumpulan data, preprocessing, pelabelan data, penerapan teknik SMOTE untuk menangani ketidakseimbangan data, dan evaluasi model. Setiap tahapan dijelaskan secara rinci di bawah ini.

2.1 Tahapan Penelitian

Proses penelitian ini terdiri dari beberapa tahapan yang saling berkaitan, yang dimulai dari pengumpulan data, diikuti dengan pembersihan dan pengolahan data (*preprocessing*), pelabelan sentimen, penerapan teknik SMOTE, hingga evaluasi model. Diagram alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

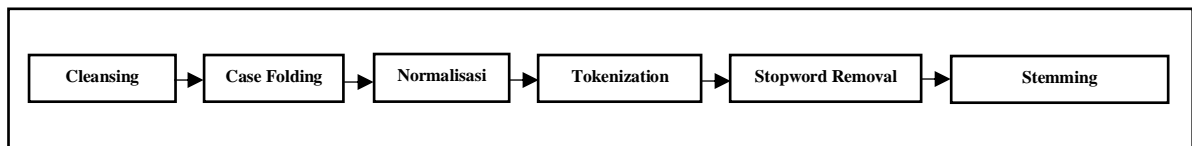
2.2 Pengumpulan Data

Langkah pertama dalam penelitian ini adalah mengumpulkan data komentar dari Twitter terkait kebijakan kenaikan gaji guru. Pengumpulan data dilakukan menggunakan pustaka *Harvest* pada platform Google Colab[1]. Data yang dikumpulkan berjumlah 6010 tweet yang relevan dengan topik tersebut. Data dikumpulkan selama periode antara 3 April 2024 hingga 30 Januari 2025, dengan menggunakan kata kunci yang relevan, seperti "kenaikan gaji guru" dan "gaji guru". Setelah data dikumpulkan, tahap berikutnya adalah proses pembersihan data (*preprocessing*).

2.3 Preprocessing Data

Tahap *preprocessing* merupakan langkah krusial dalam memastikan kualitas data yang akan digunakan dalam analisis. Pada tahap ini, data yang telah dikumpulkan dibersihkan dari elemen-elemen yang tidak relevan, seperti emoji, tanda

baca yang berlebihan, atau *username* Twitter yang tidak penting. Proses ini dilakukan dengan langkah-langkah berikut yang dapat dilihat pada Gambar 2[9].



Gambar 2. Alur Preprocessing Data

- Cleansing* merupakan tahapan pertama pada *Preprocessing* Data yaitu proses penghapusan duplikasi data seperti menghapus komentar yang berulang agar tidak ada bias dalam analisis[10]. Pada tahapan ini juga dilakukan pembersihan teks yang meliputi:
 - Menghapus emoji, angka, dan simbol khusus.
 - Menghapus username Twitter (@username) untuk menghilangkan identifikasi pengguna.
- Case Folding* yaitu mengubah semua teks menjadi huruf kecil untuk menghindari ketidakkonsistenan dalam penulisan kata[11].
- Normalisasi yaitu mengubah kata-kata tidak baku menjadi bentuk standar menggunakan kamus kata baku untuk bahasa Indonesia[12].
- Tokenization* untuk memisahkan teks menjadi kata-kata individual untuk analisis lebih lanjut[13].
- Stopword removal* yaitu tahap untuk menghapus kata-kata umum yang tidak memberi kontribusi signifikan terhadap analisis sentimen, seperti "ini", "adalah", dan sebagainya[14].
- Stemming* untuk mengubah kata-kata ke bentuk dasarnya, misalnya "kenaikan" menjadi "naik", untuk menyederhanakan variasi kata yang ada[15].

2.4 Pelabelan Data

Setelah tahap preprocessing, data yang siap untuk dianalisis berkurang menjadi 5531 entri. Pelabelan data dilakukan untuk mengkategorikan sentimen pada setiap tweet yang telah diproses. Proses pelabelan ini menggunakan pendekatan berbasis kamus, di mana setiap tweet diklasifikasikan sebagai sentimen positif, negatif, atau netral. Kamus sentimen yang digunakan dalam penelitian ini adalah kamus *Lexicon-Based*, yang telah banyak diterapkan dalam analisis sentimen berbahasa Indonesia. Pendekatan *Lexicon-Based* melibatkan penggunaan frasa, ekspresi, atau konten teks yang biasanya ditemukan dalam obrolan, dialog, posting, ulasan, dan lainnya. Pendekatan ini menggunakan kamus sentimen yang berisi kata-kata positif dan negatif, yang kemudian dicocokkan dengan kata-kata dalam kalimat untuk menentukan polaritasnya. Klasifikasi sentimen dengan *Lexicon-Based* dilakukan dengan menganalisis kata positif dan negatif dalam tweet yang telah dibersihkan[16]. Jika jumlah kata positif lebih banyak dari kata negatif, tweet tersebut dikategorikan sebagai positif. Jika kata negatif lebih dominan, tweet dikategorikan sebagai negatif, dan jika kedua jenis kata seimbang, tweet tersebut dikategorikan sebagai netral.

2.5 Feature Extraction

Feature Extraction dalam penelitian ini dilakukan menggunakan TF-IDF (*Term Frequency - Inverse Document Frequency*) untuk mengubah teks menjadi representasi numerik yang dapat diolah oleh model *machine learning*. TF-IDF menghitung bobot setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen relatif terhadap keseluruhan *dataset*, sehingga kata yang sering muncul tetapi kurang informatif memiliki bobot lebih rendah. Teknik ini memungkinkan model untuk lebih fokus pada kata-kata yang memiliki signifikansi tinggi dalam menentukan sentimen teks[17].

2.6 Pembangunan Model

Model klasifikasi dalam penelitian ini dibangun menggunakan tiga algoritma utama, yaitu Naïve Bayes, Random Forest, dan Support Vector Machine (SVM). Model dilatih menggunakan 80% data latih dan 20% data uji untuk memastikan performa model dapat dievaluasi dengan baik. Berikut merupakan penjelasan dari algoritma-algoritma yang digunakan pada penelitian ini:

- Naïve Bayes: Algoritma ini bekerja berdasarkan hukum probabilitas Bayes, yang menghitung kemungkinan suatu kategori berdasarkan distribusi kata dalam *dataset*. Algoritma ini dikenal karena efisiensinya dalam klasifikasi teks, terutama dalam kasus di mana asumsi independensi antar fitur dapat diterima[18].
- Random Forest: Algoritma ini merupakan teknik *ensemble learning* yang menggabungkan banyak pohon keputusan untuk meningkatkan akurasi prediksi. Dengan membangun beberapa pohon keputusan secara acak, Random Forest mampu mengurangi *overfitting* dan menghasilkan klasifikasi yang lebih stabil dibandingkan dengan metode pohon keputusan tunggal[19].
- Support Vector Machine (SVM): SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan kelas-kelas dalam data. Model ini dikenal unggul dalam menangani data berdimensi tinggi dan efektif dalam mengklasifikasikan data dengan margin yang jelas[20].



2.7 Optimasi SMOTE

Ketidakeimbangan data merupakan tantangan utama dalam analisis sentimen, terutama dalam kasus di mana sentimen positif mendominasi dibandingkan dengan sentimen negatif dan netral. Untuk mengatasi masalah ini, penelitian ini menerapkan teknik SMOTE, yang dapat meningkatkan jumlah data pada kelas minoritas (sentimen negatif dan netral) dengan menghasilkan data sintesis berdasarkan data yang ada. Teknik ini membantu model untuk belajar dengan lebih baik dalam mengidentifikasi sentimen negatif dan netral yang lebih jarang ditemukan, sehingga meningkatkan akurasi model secara keseluruhan.

2.8 Evaluasi Model

Setelah model dilatih, dilakukan evaluasi menggunakan beberapa metrik yang umum digunakan dalam analisis sentimen, yaitu:

- Akurasi untuk mengukur persentase prediksi yang benar dibandingkan dengan jumlah total sampel uji.
- Presisi untuk mengukur keakuratan model dalam mengklasifikasikan sentimen positif, yakni berapa banyak prediksi positif yang benar.
- Recall untuk mengukur seberapa banyak sentimen yang benar-benar ditemukan oleh model, terutama untuk sentimen negatif dan netral.
- F1-score merupakan rata-rata harmonis antara presisi dan recall, memberikan gambaran keseluruhan kinerja model dalam kondisi data yang tidak seimbang.

Hasil evaluasi ini membantu dalam menilai kekuatan dan kelemahan dari setiap algoritma yang digunakan dalam penelitian ini.

3. HASIL DAN PEMBAHASAN

Pada bagian ini, kami akan membahas hasil yang diperoleh dari penerapan tiga algoritma machine learning—Naïve Bayes, Random Forest, dan Support Vector Machine (SVM)—dalam analisis sentimen terhadap kebijakan kenaikan gaji guru di Indonesia. Selain itu, akan dijelaskan bagaimana penerapan teknik SMOTE berhasil meningkatkan kinerja model, terutama dalam mengatasi ketidakseimbangan data. Setiap tahapan, mulai dari pengumpulan data, preprocessing, pelabelan, penerapan SMOTE, hingga evaluasi model, akan dibahas secara rinci.

3.1 Pengumpulan Data

Data dalam penelitian ini diperoleh dari platform media sosial *Twitter* menggunakan teknik *Crawling*. Total data mentah yang berhasil dikumpulkan sebanyak 6010 komentar yang membahas kenaikan gaji guru yang dikumpulkan mulai dari 3 april 2024 hingga 30 januari 2025. Namun, data ini masih mengandung elemen-elemen yang tidak relevan seperti duplikasi teks, tanda baca berlebihan, dan karakter khusus yang dapat memengaruhi kualitas analisis sentimen. Oleh karena itu, diperlukan tahap *preprocessing* untuk menyaring dan membersihkan data sebelum dilakukan klasifikasi. Tabel 1 berikut merupakan contoh dari hasil pengumpulan data.

Tabel 1. Hasil Pengumpulan Data

No	Hasil Pengumpulan Data
1	Kenaikan gaji menyemangati guru untuk terus berkarya! #PemerintahNaikkanGajiGuru
2	@Yunisyaaahhh33 Guru SD 14 Bekasi sangat senang dengan adanya kenaikan gaji ini.
...	...
6010	pak ini nasib guru honorer di Indonesia jangan hanya IKN aja
6011	Kenaikan gaji guru adalah investasi masa depan untuk dunia pendidikan yang lebih baik.

3.2 Preprocessing Data

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum masuk ke proses klasifikasi sentimen. Setelah dilakukan beberapa tahap pembersihan data, jumlah data yang siap digunakan dalam analisis berkurang menjadi 5531 entri. Berikut adalah hasil dari setiap tahap *preprocessing*:

a. *Cleansing*

Pada tahap ini, karakter khusus seperti tanda baca yang tidak diperlukan, angka, serta simbol dihapus dari teks. Hal ini bertujuan untuk memastikan bahwa data hanya mengandung kata-kata yang relevan dalam analisis sentimen. Tabel 2 berikut merupakan hasil dari data komentar yang telah melalui tahap *Cleansing*.

Tabel 2. Hasil *Cleansing*

<i>Comment</i>	<i>Cleansing</i>
@nasilodho Tumpah ruah! Segini besaran kenaikan gaji guru!!!!!!111	Tumpah ruah Segini besaran kenaikan gaji guru



b. Case Folding

Semua teks diubah menjadi huruf kecil untuk memastikan tidak ada perbedaan antara kata yang memiliki makna sama namun ditulis dengan format yang berbeda. Misalnya, kata "Hingga" akan dianggap identik setelah dilakukan *Case Folding*. Tabel 3 dibawah ini adalah contoh perbedaan data setelah dilakukan tahap *Case Folding*.

Tabel 3. Hasil *Case Folding*

<i>Cleansing</i>	<i>Case Folding</i>
Hingga saat ini nihil Pak sama sekali tdk ada kenaikan gaji guru ASN	hingga saat ini nihil pak sama sekali tdk ada kenaikan gaji guru asn

c. Normalisasi

Pada tahap ini, kata-kata yang memiliki bentuk tidak baku atau singkatan diubah menjadi bentuk standar. Misalnya, kata "ilang" diubah menjadi "hilang". Tabel 4 dibawah ini adalah contoh hasil perbedaan data komentar setelah dilakukan tahap Normalisasi.

Tabel 4. Hasil Normalisasi

<i>Case Folding</i>	Normalisasi
kenaikan gaji guru honorer bangun fakultas kedokteran bagian ini kok ilang	kenaikan gaji guru honorer bangun fakultas kedokteran bagian ini kok hilang

d. Tokenization

Proses ini memisahkan teks menjadi kata-kata individual agar lebih mudah dianalisis oleh model *machine learning*. Sebagai contoh, kalimat "Kenaikan gaji guru" akan diubah menjadi daftar kata: ['kenaikan', 'gaji', 'guru']. Tabel 5 dibawah ini adalah contoh hasil dan perbedaan data setelah dilakukan tahap *Tokenization*.

Tabel 5. Tokenization

Normalisasi	<i>Tokenization</i>
kayak kenaikan gaji guru kemarin sampai menangis nyatanya diprank	['kayak', 'kenaikan', 'gaji', 'guru', 'kemarin', 'sampai', 'menangis', 'nyatanya', 'diprank']

e. Stopword Removal

Pada tahap ini, kata-kata umum yang tidak memiliki nilai informatif tinggi dalam klasifikasi sentimen, seperti "saat", "ini", "ada", dihapus. Hal ini bertujuan untuk mengurangi beban komputasi dan meningkatkan relevansi kata dalam analisis sentimen. Tabel 6 dibawah ini adalah contoh perbedaan data setelah dilakukan tahap ini.

Tabel 6. *Stopword Removal*

<i>Tokenization</i>	<i>Stopword Removal</i>
['hingga', 'saat', 'ini', 'nihil', 'pak', 'sama', 'sekali', 'tidak', 'ada', 'kenaikan', 'gaji', 'guru', 'asn']	['nihil', 'kenaikan', 'gaji', 'guru', 'asn']

f. Stemming

Stemming dilakukan untuk mengubah kata menjadi bentuk dasarnya agar model dapat mengenali variasi kata sebagai satu entitas yang sama. Sebagai contoh, kata "kenaikan" akan dikonversi menjadi "naik". Tabel 7 dibawah ini adalah contoh hasil dan perbedaan data setelah dilakukan tahap *Stemming*.

Tabel 7. *Stemming*

<i>Stopword Removal</i>	<i>Stemming</i>
['kenaikan', 'gaji', 'guru', 'fokus', 'pengembangan', 'kualitas', 'pemerintahnaikkangajiguru']	naik gaji guru fokus kembang kualitas ajar 'pengajaran', pemerintahnaikkangajiguru

3.3 Pelabelan Data

Pelabelan data sentimen menggunakan pendekatan *lexicon-based*. Sistem ini bekerja dengan membandingkan kata-kata dalam komentar dengan kamus sentimen *InSet*, yang telah dikategorikan sebagai positif, negatif, atau netral. Hasil pelabelan dapat dilihat pada Tabel 8.

Tabel 8. Pelabelan Data

<i>Stemming</i>	Score	Label
terima kasih perintah naik gaji tingkat sejahtera guru pemerintahnaikkangajiguru	4	Positif

tersebut serta seluruh dataset. Kata yang sering muncul namun kurang informatif akan mendapatkan bobot yang lebih rendah, sementara kata-kata yang jarang namun signifikan akan mendapatkan bobot lebih tinggi. Dengan demikian, fitur yang dihasilkan akan lebih fokus pada kata-kata yang memiliki kontribusi besar dalam menentukan sentimen dari teks. Hasil dari transformasi ini berupa *matriks sparse* yang berisi vektor-vektor fitur untuk setiap teks. *Matriks* ini memiliki banyak kolom, yang masing-masing mewakili kata unik dalam seluruh dataset, dan setiap baris mewakili sebuah teks atau dokumen. Dalam implementasi pada dataset ini, setelah diterapkan teknik TF-IDF, diperoleh *matriks* fitur dengan ukuran 5531 x 5000 (di mana 5000 adalah jumlah kata unik yang ditemukan dalam *dataset* setelah proses normalisasi dan *tokenization*). Proses ini menghasilkan fitur yang sangat kaya dan memungkinkan model untuk melakukan klasifikasi sentimen secara efektif. Hasil dari *Feature Extraction* TF-IDF dapat dilihat pada Tabel 9.

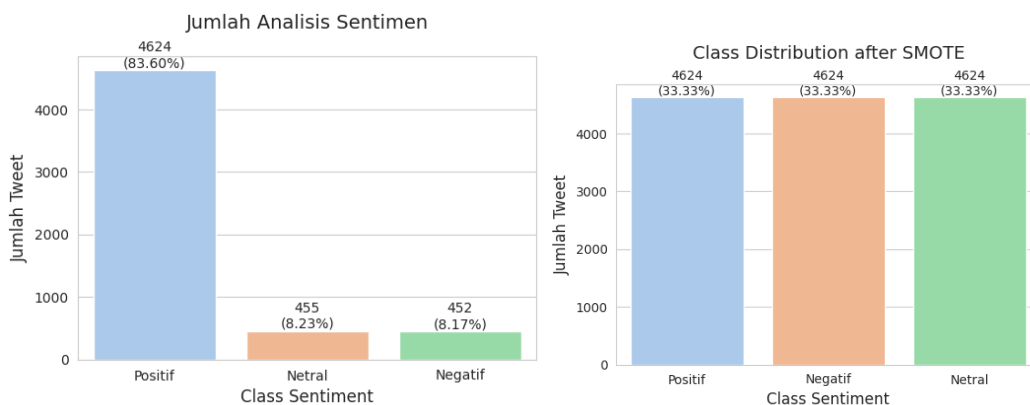
Tabel 9. Hasil TF-IDF

	Term	TFIDF_SUM
2086	guru	462.168239893028
1818	gaji	425.6245796712907
6595	pemerintahnaikkangajiguru	358.9368663270745
5972	naik	335.70554308329025
1318	didik	244.52277754237517

Tabel 9 ini menampilkan 5 kata dengan nilai TF-IDF tertinggi yang ditemukan dalam komentar Twitter terkait kenaikan gaji guru. Nilai TF-IDF mengukur pentingnya kata dalam konteks dokumen dibandingkan dengan seluruh *dataset*. Kata-kata dengan nilai tinggi, seperti "kenaikan", "gaji", dan "guru", menunjukkan topik utama yang dominan dalam opini masyarakat. Kata-kata ini memainkan peran penting dalam membedakan sentimen dalam analisis.

3.6 Optimasi SMOTE

SMOTE (*Synthetic Minority Over-sampling Technique*) diterapkan untuk menangani masalah ketidakseimbangan kelas dalam *dataset*. Pada Gambar 5 adalah hasil dari sebelum Optimasi SMOTE dan setelah Optimasi SMOTE.



Gambar 5. Visualisasi Sebelum SMOTE dan Setelah SMOTE

Gambar 5 merupakan hasil dari sebelum dan sesudah penerapan SMOTE, sebelum penerapan SMOTE *dataset* cenderung didominasi oleh sentimen positif, sementara sentimen negatif dan netral lebih sedikit jumlahnya. Hal ini menyebabkan model lebih cenderung mengklasifikasikan komentar sebagai sentimen positif, sehingga menurunkan presisi dalam klasifikasi sentimen negatif dan netral. Setelah SMOTE diterapkan, jumlah data untuk kelas minoritas (negatif dan netral) diperbanyak dengan menghasilkan data sintesis baru. Hasilnya, distribusi kelas menjadi lebih seimbang, yang memungkinkan model untuk lebih sensitif terhadap sentimen negatif dan netral.

3.7 Hasil Pengujian Model

Pada bagian ini, Hasil dari pengujian model ini menggambarkan kinerja masing-masing algoritma dalam mengklasifikasikan sentimen positif, negatif, dan netral berdasarkan data Twitter yang telah diproses. Proses pengujian dilakukan dengan menggunakan 80% data latih dan 20% data uji. *Metriks* yang digunakan untuk evaluasi adalah akurasi, presisi, recall, dan F1-score. Berikut adalah hasil pengujian sebelum dan setelah penerapan SMOTE.

3.7.1 Hasil Sebelum Penerapan SMOTE:

Sebelum penerapan teknik SMOTE, ketiga model menunjukkan perbedaan kinerja yang cukup signifikan. Tabel 10 di bawah ini memperlihatkan hasil evaluasi untuk Naïve Bayes, Random Forest, dan SVM.

Tabel 10. Hasil Sebelum Penerapan SMOTE

Model	Accuracy	Precision (Negatif)	Recall (Negatif)	Precision (Positif)	Recall (Positif)	F1-Score (Negatif)	F1-Score (Positif)
Naive Bayes	84%	0%	0%	87%	99%	0%	93%
Random Forest	85%	0%	0%	88%	100%	0%	94%
SVM	86%	0%	7%	89%	100%	0%	94%

Berdasarkan Tabel 10, dapat dilihat bahwa meskipun akurasi keseluruhan dari ketiga model relatif tinggi (di atas 80%), presisi dan recall untuk sentimen negatif sangat rendah. Ini menunjukkan bahwa ketidakseimbangan data menjadi masalah utama dalam mengklasifikasikan sentimen negatif. Model cenderung lebih baik dalam mengidentifikasi sentimen positif, karena jumlah data sentimen positif jauh lebih banyak dalam dataset.

3.7.2 Hasil Setelah Penerapan SMOTE:

Setelah teknik SMOTE diterapkan untuk mengatasi ketidakseimbangan data, hasil pengujian menunjukkan peningkatan yang signifikan pada semua model. SMOTE menghasilkan data sintesis untuk kelas minoritas, yang meningkatkan representasi sentimen negatif dan netral dalam dataset. Hasil setelah penerapan SMOTE dapat dilihat pada Tabel 11.

Tabel 11. Hasil Setelah Penerapan SMOTE

Model	Accuracy	Precision (Negatif)	Recall (Negatif)	Precision (Positif)	Recall (Positif)	F1-Score (Negatif)	F1-Score (Positif)
Naive Bayes	89%	85%	87%	91%	92%	86%	91%
Random Forest	99%	98%	97%	100%	99%	98%	100%
SVM	98%	97%	98%	99%	97%	97%	98%

Setelah SMOTE diterapkan, peningkatan yang signifikan terlihat pada semua model. Akurasi meningkat di semua algoritma, dengan Random Forest mencapai akurasi tertinggi (99%). Selain itu, nilai presisi, recall, dan F1-score untuk sentimen negatif dan netral juga meningkat, yang menunjukkan bahwa SMOTE berhasil memperbaiki ketidakseimbangan data dan memungkinkan model untuk lebih akurat dalam mengklasifikasikan semua kelas sentimen.

3.8 Confusion Matrix - Evaluasi Model

Confusion Matrix adalah alat evaluasi yang sangat berguna dalam analisis model klasifikasi. Ini membantu kita memahami bagaimana model mengklasifikasikan data ke dalam tiga kelas sentimen: positif, negatif, dan netral. Dengan melihat perbandingan antara prediksi model dan nilai aktual dari data uji, kita dapat mengetahui seberapa akurat model dalam mengklasifikasikan sentimen. Secara sederhana, *Confusion Matrix* membagi hasil prediksi model menjadi empat kategori utama:

- True Positives* (TP): Jumlah data yang benar-benar positif dan diprediksi sebagai positif oleh model.
- True Negatives* (TN): Jumlah data yang benar-benar negatif dan diprediksi sebagai negatif oleh model.
- False Positives* (FP): Jumlah data yang sebenarnya negatif, tetapi diprediksi sebagai positif oleh model.
- False Negatives* (FN): Jumlah data yang sebenarnya positif, tetapi diprediksi sebagai negatif oleh model.

Dengan menggunakan *Confusion Matrix*, kita bisa mendapatkan gambaran yang jelas tentang performa model, terutama dalam menghadapi ketidakseimbangan data. Tabel di bawah ini menunjukkan *Confusion Matrix* untuk masing-masing algoritma sebelum dan setelah penerapan teknik SMOTE.

3.8.1 Confusion Matrix Sebelum SMOTE:

Berikut adalah perbandingan hasil *Confusion Matrix* untuk Naïve Bayes, Random Forest, dan SVM sebelum penerapan SMOTE. Tabel ini menggambarkan bagaimana masing-masing model mengklasifikasikan sentimen positif, negatif, dan netral. Hasil sebelum penerapan SMOTE dapat dilihat pada Tabel 12.

Tabel 12. Hasil Sebelum Penerapan SMOTE

Model	Prediksi Positif	Prediksi Negatif	Prediksi Netral	Aktual Positif	Aktual Negatif	Aktual Netral
Naive Bayes	1386 (TP)	0	0	1386 (TP)	145 (FP)	129 (FP)
Random Forest	1384 (TP)	0	2	1386 (TP)	10 (FP)	129 (FP)
SVM	1377 (TP)	8 (FN)	1	1386 (TP)	94 (FP)	129 (FN)

True Positives (TP) menunjukkan jumlah tweet yang benar-benar positif dan diprediksi sebagai positif oleh model. Pada model Naïve Bayes, 1386 tweet positif diprediksi dengan benar, sementara Random Forest berhasil memprediksi 1384 tweet positif dengan tepat, dan SVM berhasil memprediksi 1377 tweet positif dengan benar. Di sisi lain, *False Negatives* (FN) menunjukkan jumlah tweet yang seharusnya positif, tetapi diprediksi sebagai negatif.



Pada Naïve Bayes, tidak ada tweet yang salah diprediksi sebagai negatif, namun Random Forest mengalami kesalahan pada 2 tweet positif yang diprediksi sebagai netral. SVM, di sisi lain, salah memprediksi 8 tweet positif sebagai negatif. False Positives (FP) adalah jumlah tweet yang seharusnya negatif atau netral, namun diprediksi sebagai positif. Pada Naïve Bayes, 145 tweet negatif dan 129 tweet netral salah diprediksi sebagai positif. Random Forest juga membuat kesalahan dengan memprediksi 10 tweet negatif dan 129 tweet netral sebagai positif, sementara SVM salah memprediksi 94 tweet negatif dan 129 tweet netral sebagai positif. *True Negatives* (TN) menunjukkan jumlah tweet yang benar-benar negatif atau netral dan diprediksi dengan benar. Naïve Bayes tidak berhasil mengklasifikasikan tweet negatif atau netral dengan benar, sedangkan Random Forest berhasil mengklasifikasikan 10 tweet negatif dengan tepat. SVM memprediksi dengan benar 44 tweet negatif dan 13 tweet netral.

Berdasarkan tabel yang ada, terlihat bahwa semua model lebih efektif dalam mengklasifikasikan sentimen positif, tetapi kesulitan dalam mengklasifikasikan sentimen negatif dan netral. Hal ini sangat jelas terlihat pada Naïve Bayes, yang tidak mampu mengidentifikasi sentimen negatif dan netral dengan baik. Ini menunjukkan adanya bias terhadap sentimen positif akibat ketidakseimbangan data dalam dataset.

Model Naïve Bayes menunjukkan kesulitan yang besar dalam mengklasifikasikan sentimen negatif dan netral, dengan tidak ada *True Negatives* dan *True Positives* yang cukup untuk kelas tersebut. Di sisi lain, Random Forest sedikit memperbaiki kinerjanya dengan berhasil mengidentifikasi beberapa *True Negatives* untuk sentimen negatif, serta lebih banyak *True Positives* untuk sentimen positif. SVM memberikan hasil yang lebih seimbang dibandingkan dua model lainnya, dengan beberapa kesalahan prediksi pada sentimen negatif dan netral, meskipun masih terdapat beberapa kesalahan pada kelas netral.

Secara keseluruhan, sebelum penerapan SMOTE, semua model menunjukkan kecenderungan untuk lebih memprediksi sentimen positif, dengan sentimen negatif dan netral sering salah diklasifikasikan sebagai positif. SMOTE diharapkan dapat mengatasi masalah ketidakseimbangan kelas, sehingga meningkatkan kinerja model dalam mengklasifikasikan semua kelas dengan lebih akurat.

3.8.2 Confusion Matrix Setelah SMOTE:

Berikut adalah *Confusion Matrix* untuk Naïve Bayes, Random Forest, dan SVM setelah penerapan SMOTE, yang menunjukkan bagaimana masing-masing model mengklasifikasikan sentimen positif, negatif, dan netral. Hasil setelah penerapan SMOTE dapat dilihat pada Tabel 13.

Tabel 13. Hasil Setelah Penerapan SMOTE

Model	Prediksi Positif	Prediksi Negatif	Prediksi Netral	Aktual Positif	Aktual Negatif	Aktual Netral
Naive Bayes	716 (TP)	124 (FP)	16 (FP)	1386 (TP)	900 (TN)	922 (TN)
Random Forest	935 (TP)	11 (FN)	7 (FP)	1386 (TP)	893 (TN)	908 (TN)
SVM	898 (TP)	17 (FP)	38 (FN)	1386 (TP)	900 (TN)	922 (TN)

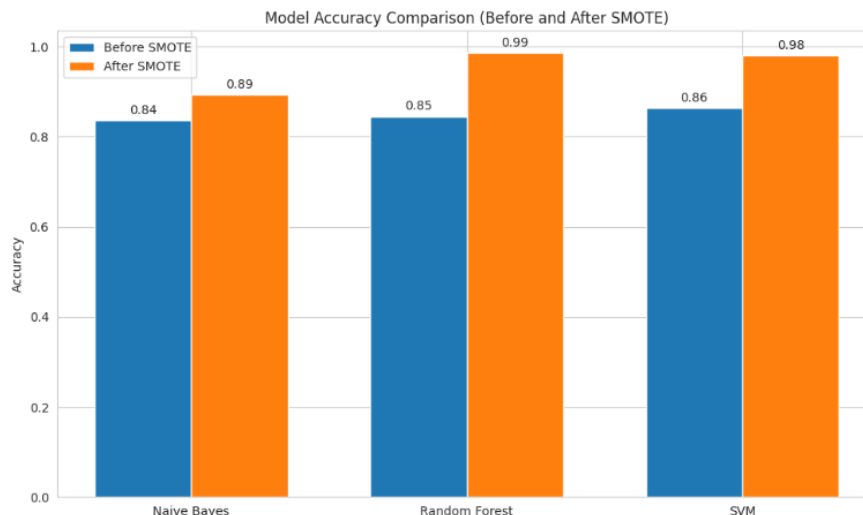
Setelah penerapan teknik SMOTE, kinerja model Naïve Bayes mengalami peningkatan yang signifikan dalam mengklasifikasikan sentimen. Sebelumnya, model ini kesulitan dalam mengidentifikasi sentimen negatif dan netral karena masalah ketidakseimbangan data. Namun, setelah SMOTE diterapkan, model berhasil mengklasifikasikan 716 tweet positif dengan tepat, meskipun masih ada beberapa kesalahan prediksi, terutama antara sentimen negatif dan netral. Meskipun demikian, model ini berhasil mengurangi kesalahan klasifikasi dengan memperoleh akurasi sekitar 89%. Secara keseluruhan, penerapan SMOTE memungkinkan Naïve Bayes untuk memperbaiki performanya dalam mengatasi data yang tidak seimbang, terutama pada kelas minoritas.

Pada model Random Forest, penerapan SMOTE memberikan hasil yang luar biasa. Model ini berhasil mengklasifikasikan sentimen negatif dan netral dengan sangat baik, memperoleh akurasi 99% setelah SMOTE diterapkan. Random Forest dapat memprediksi dengan tepat 935 tweet positif, 893 tweet negatif, dan 908 tweet netral. Walaupun masih ada beberapa kesalahan prediksi, seperti *False Negatives* dan *False Positives* pada sentimen negatif dan netral, model ini menunjukkan performa terbaik dibandingkan dengan model lainnya. Hal ini menunjukkan bahwa Random Forest memiliki keunggulan dalam menangani ketidakseimbangan data yang ada.

Sementara itu, SVM juga menunjukkan perbaikan yang signifikan setelah penerapan SMOTE, dengan akurasi 98%. Model ini berhasil mengklasifikasikan 898 tweet positif dengan tepat dan mampu mengidentifikasi sentimen negatif dan netral dengan baik. Meskipun ada beberapa *False Negatives* pada sentimen netral, model ini tetap menunjukkan performa yang baik dalam memprediksi kelas-kelas tersebut. Secara keseluruhan, meskipun SVM menunjukkan kinerja yang baik, Random Forest tetap lebih unggul dalam hal akurasi.

Secara keseluruhan, penerapan SMOTE berhasil memperbaiki akurasi ketiga model dalam mengklasifikasikan sentimen negatif dan netral, yang sebelumnya terabaikan karena ketidakseimbangan data. Random Forest menunjukkan hasil terbaik, diikuti oleh SVM, dan Naïve Bayes yang juga memperlihatkan perbaikan signifikan. Penerapan SMOTE memungkinkan ketiga model ini untuk mengklasifikasikan sentimen lebih seimbang dengan tingkat kesalahan yang lebih rendah pada kelas minoritas.

3.9 Visualisasi Perbandingan Algoritma



Gambar 6. Visualisasi Perbandingan Akurasi Algoritma

Gambar 6 merupakan grafik batang yang menyajikan perbandingan akurasi model Naive Bayes, Random Forest, dan SVM sebelum dan sesudah penerapan SMOTE (*Synthetic Minority Over-sampling Technique*). Terlihat bahwa ketiga model mengalami peningkatan akurasi setelah penerapan SMOTE. Peningkatan paling signifikan terlihat pada model Random Forest, di mana akurasi meningkat dari 85% menjadi 99%. Model Naive Bayes juga mengalami peningkatan yang cukup besar, dari 84% menjadi 89%. Sementara itu, model SVM menunjukkan peningkatan moderat dari 86% menjadi 98%. Grafik ini secara visual mengonfirmasi bahwa SMOTE efektif dalam mengatasi masalah ketidakseimbangan kelas dan meningkatkan performa ketiga algoritma dalam mengklasifikasikan sentimen.

4. KESIMPULAN

Penelitian ini menganalisis sentimen publik terhadap kebijakan kenaikan gaji guru di Indonesia menggunakan algoritma Naive Bayes, Random Forest, dan SVM, serta mengevaluasi penerapan teknik SMOTE untuk mengatasi ketidakseimbangan data dalam *dataset*. Hasil penelitian menunjukkan bahwa penerapan SMOTE berhasil meningkatkan kinerja semua model, terutama dalam mengklasifikasikan sentimen negatif dan netral yang sebelumnya kurang teridentifikasi. Setelah penerapan SMOTE, Random Forest mencapai akurasi tertinggi sebesar 99%, diikuti oleh SVM dengan akurasi 98% dan Naive Bayes dengan akurasi 89%. Penerapan SMOTE secara efektif menyeimbangkan distribusi kelas sentimen, meningkatkan akurasi model secara keseluruhan. Temuan ini memberikan kontribusi penting terhadap penggunaan teknik SMOTE dalam analisis sentimen pada dataset yang tidak seimbang, khususnya dalam konteks media sosial, dan menyarankan bahwa teknik ini dapat diterapkan untuk meningkatkan prediksi pada model-model analisis sentimen di masa depan. Selain itu, hasil dari analisis sentimen ini memberikan wawasan mengenai persepsi publik terhadap kebijakan kenaikan gaji guru, yang dapat digunakan sebagai alat untuk mengevaluasi dan memperbaiki kebijakan pendidikan. Hasil ini menunjukkan bahwa masyarakat lebih cenderung memiliki pandangan positif terhadap kebijakan ini, meskipun masih ada segmen masyarakat yang merasa tidak puas, terutama terkait dengan keadilan distribusi dan pelaksanaan kebijakan. Meskipun demikian, penelitian ini juga menunjukkan keterbatasan pada pengumpulan data yang hanya berasal dari platform Twitter, yang mungkin tidak sepenuhnya mewakili opini publik secara luas. Untuk itu, penelitian selanjutnya dapat mengeksplorasi penggunaan data yang lebih beragam dari berbagai platform media sosial, serta menerapkan teknik analisis yang lebih kompleks seperti *deep learning* untuk menangani konteks emosional yang lebih mendalam dan beragam. Dengan demikian, kebijakan pendidikan, termasuk kebijakan kenaikan gaji guru, dapat disesuaikan lebih baik dengan keinginan masyarakat, berlandaskan pada analisis data yang lebih komprehensif dan representatif.

REFERENCES

- [1] J. I. Manajemen, H. Jurnal, M. Deka Faathir, and F. Fauziya, "Analisis Tingkat Gaji pada Kinerja Guru di MIS Muslimin Cihurang," *JURNAL EKONOMIKA45*, vol. 12, no. 1, 2024.
- [2] Y. Wang, J. Guo, C. Yuan, and B. Li, "Sentiment Analysis of Twitter Data," *MDPI*, Nov. 01, 2022, doi: 10.3390/app122211775.



- [3] N. R. Ramadhan and N. Hendrastuty, “Perbandingan Algoritma Naïve Bayes dan LSTM untuk Analisis Sentimen Terhadap Opini Masyarakat Tentang Sandwich Generation,” *Technology and Science (BITS)*, vol. 6, no. 3, pp. 1677–1687, 2024, doi: 10.47065/bits.v6i3.6385.
- [4] R. Nurhidayat and N. Hendrastuty, “Analisis Sentimen Komentar Media Sosial Twitter Terhadap Tes CPNS dengan Algoritma Naive Bayes,” *Technology and Science (BITS)*, vol. 6, no. 3, pp. 1477–1489, 2024, doi: 10.47065/bits.v6i3.6148.
- [5] I. Bagus, G. Sarasvananda, D. Selivan, M. L. Radhitya, N. Tri, and A. Putra, “Analisis Sentimen Pada Pembelajaran Daring Di Indonesia Melalui Twitter Menggunakan Naïve Bayes Classifier”, *SINTECH*, vol 5
- [6] I. Septiana and D. Alita, “Perbandingan Random Forest dan SVM dalam Analisis Sentimen Quick Count Pemilu 2024,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 9, no. 3, pp. 224–233, Dec. 2024, doi: 10.30591/jpit.v9i3.6640.
- [7] D. Kurniawan, M. Najib, and D. Satria, “Analisis Sentimen Opini Publik Tentang Gempa Megathrust di Indonesia Menggunakan Metode Support Vector Machine dan Naïve Bayes,” *Technology and Science (BITS)*, vol. 6, no. 3, 2024, doi: 10.47065/bits.v6i3.6213.
- [8] Y. A. Singgalen, “A Hybrid CNN-LSTM Model with SMOTE for Enhanced Sentiment Analysis of Hotel Reviews,” *Technology and Science (BITS)*, vol. 6, no. 3, pp. 1363–1373, 2024, doi: 10.47065/bits.v6i3.6301.
- [9] E. Septiana Pane, C. Caroline, “Optimalisasi Evaluasi Pelaksanaan Pelatihan Melalui Analisis Sentimen Otomatis Dengan Model Text Classification.” *Prosiding PITNAS Widyaiswara*, 2024
- [10] M. Syiarul Amrullah and A. Yuniarti, “Effective Coronary Artery Disease Prediction Using Bayesian Optimization Algorithm and Random Forest,” *Technology and Science (BITS)*, vol. 6, no. 2, 2024, doi: 10.47065/bits.v6i2.5554.
- [11] A. Erfina and M. R. N. R. Alamsyah, “Implementation of Naive Bayes classification algorithm for Twitter user sentiment analysis on ChatGPT using Python programming language,” *Data and Metadata*, vol. 2, Jan. 2023, doi: 10.56294/dm202345.
- [12] L. Rofiqi and M. Akbar, “Analisis Sentimen Terkait RUU Perampasan Aset dengan Support Vector Machine,” *JEKIN - Jurnal Teknik Informatika*, vol. 4, no. 3, pp. 529–538, Aug. 2024, doi: 10.58794/jekin.v4i3.824.
- [13] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, “Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, Jan. 2022, doi: 10.23960/jitet.v10i1.2262.
- [14] S. Jessica Angelina, A. Bijaksana Putra Negara, H. Muhardi, J. H. Nawawi, and K. Barat, “Analisis Pengaruh Penerapan Stopword Removal Pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia,” *Jurnal Aplikasi dan Riset Informatika*, vol. 02, no. 1, 2023, doi: 10.26418/juara.v2i1.69680.
- [15] R. Ulgasesa, A. B. P. Negara, and T. Tursina, “Pengaruh Stemming Terhadap Performa Klasifikasi Sentimen Masyarakat Tentang Kebijakan New Normal,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 3, p. 286, Sep. 2022, doi: 10.26418/justin.v10i3.53880.
- [16] M. Al Khadafi, K. Paranita Kartika, and F. Febrinita, “Penerapan Metode Naïve Bayes Classifier Dan Lexicon Based Untuk Analisis Sentimen Cyberbullying Pada BPJS,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol 6 no 2, 2022, doi 10.36040/jati.v6i2.5633
- [17] A. Apriani, H. Zakiyudin, and K. Marzuki, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta,” *Jurnal Bumigora Information Technology (BITE)*, vol. 3, no. 1, pp. 19–27, Jul. 2021, doi: 10.30812/bite.v3i1.1110.
- [18] A. Salsabila and M. Nasution, “Implementation of the Naïve Bayes Algorithm to Predict New Student Admissions,” *Technology and Science (BITS)*, vol. 6, no. 1, 2024, doi: 10.47065/bits.v6i1.5363.
- [19] M. Putra and Erwin Harahap, “Machine Learning pada Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Random Forest,” *Jurnal Riset Matematika*, pp. 127–136, Dec. 2024, doi: 10.29313/jrm.v4i2.5102.
- [20] T. T. Thet, J. C. Na, and C. S. G. Khoo, “Aspect-based sentiment analysis of movie reviews on discussion boards,” *J Inf Sci*, vol. 36, no. 6, pp. 823–848, Dec. 2010, doi: 10.1177/0165551510388123.
- [21] B. F. S. Supriyanto and S. Rosalin, “Analisis Sentimen Program Merdeka Belajar dengan Text Analysis Wordcloud & Word Frequency,” *Jurnal Minfo Polgan*, vol. 12, no. 1, pp. 25–32, Mar. 2023, doi: 10.33395/jmp.v12i1.12312.