



Komparasi Naive Bayes dan SVM Opini Kesetaraan Gender di Tiktok dengan SMOTE

Meida Amanda¹, Abdullah²

¹²Sistem Informasi, Teknik dan Ilmu Komputer, Universitas Islam Indragiri

¹meidaamanda240@gmail.com, ²abdullah@unisi.ac.id

Abstrak

Penelitian ini bertujuan membandingkan dua algoritma klasifikasi, yakni Naive Bayes dan Support Vector Machine (SVM), dalam menganalisis sentimen komentar pada platform Tiktok mengenai isu kesetaraan gender dari dua perspektif. Dataset yang digunakan berjumlah 1.600 komentar dikumpulkan dengan melakukan web scraping menggunakan web Apify. Data yang diambil mulai dari rentang waktu 2025 sampai 2026, dengan distribusi 289 komentar berlabel positif dan 1.284 komentar berlabel negatif. Penelitian ini dilakukan tahap preprocessing agar model nantinya bisa membaca dengan baik sentimennya. Tahapan dalam preprocessing terdapat cleaning, case fold, normalisasi, tokenizing, filtering dan stemming. Selanjutnya untuk pelabelan dalam penelitian ini dilakukan secara otomatis menggunakan metode berbasis lexicon dan menggunakan teknik SMOTE untuk mengatasi ketidakseimbangan data. Berikutnya hasil dari Naive Bayes dan SVM akan dievaluasi menggunakan confusion matrix yang berguna untuk mengetahui jumlah data yang berhasil dan gagal diprediksi secara benar oleh kedua model. Berdasarkan hasil pengujian, SVM lebih unggul dibandingkan dengan Naive Bayes dengan tingkat akurasi 86% berbanding 77%. Ini menandakan SVM lebih andal dalam tugas klasifikasi ini. Tingginya sentimen negatif menandakan bahwa topik kesetaraan gender di platform Tiktok masih menimbulkan perdebatan dan perbedaan pendapat di kalangan pengguna. Hasil penelitian menunjukkan bahwa SVM terbukti lebih unggul dan efektif dalam mengklasifikasikan sentimen yang berasal dari platform media social terutama Tiktok.

Kata Kunci: Sentimen, Kesetaraan Gender, Naive Bayes, SMOTE, Support Vector Machine.

1. Latar Belakang

Media sosial adalah salah satu tempat yang telah mengubah cara masyarakat dalam menyampaikan pendapat dan berpartisipasi dalam diskusi publik. Di antara berbagai platform media sosial yang tersedia, Tiktok mencatat pertumbuhan jumlah pengguna yang sangat pesat secara global. Tiktok bukan hanya digunakan sebagai hiburan semata, tetapi juga menjadi sarana komunikasi, edukasi bahkan pembelaan. Tiktok telah menjadi tempat berekspresi yang bermanfaat bagi berbagai kelompok, termasuk dalam menyuarakan pandangan seputar isu kesetaraan gender. Setiap harinya, ribuan hingga jutaan video diposting dan banyak komentar ditambahkan oleh pengguna Tiktok dari berbagai latar belakang dan usia [1].

Kesetaraan gender merupakan keadaan di mana pria maupun wanita memiliki derajat yang setara terhadap hak dan kewajiban yang ada [2]. Isu mengenai kesetaraan gender sejak dulu hingga sekarang mendapat banyak perhatian. Video terkait berupa edukasi, opini, maupun diskusi publik menghasilkan ribuan komentar setiap harinya. Komentar-komentar tersebut memperlihatkan bermacam sentimen publik, mulai dari dukungan hingga penolakan. Guna dari hasil analisis sentimen ini adalah memperoleh gambaran persepsi masyarakat secara lebih akurat [3].

Analisis sentimen merupakan bagian dari bidang Natural Language Processing (NLP) yang berfokus pada kemampuan komputer dalam memahami bahasa yang digunakan manusia secara alami [4]. Teknik ini telah banyak diterapkan dalam berbagai analisis ulasan produk, opini politik, hingga isu sosial. Naive Bayes dan Support Vector Machine merupakan dua algoritma yang kerap kali diterapkan dalam analisis sentimen berbasis machine learning [5]. Algoritme Naive Bayes dikenal sebagai metode yang populer karena kesederhanaan dan kecepatannya dalam memproses data teks. Sedangkan, Support Vector Machine karena mampu mengenali pola data yang berdimensi tinggi dengan baik dalam analisis sentimen [6], [7]. Beberapa penelitian telah membandingkan kedua algoritma ini dalam konteks serupa. Shabira dkk. (2023) membandingkan SVM dan Naive Bayes untuk deteksi seksisme online menggunakan dataset SemEval-2023 dengan penerapan SMOTE, dan menemukan bahwa SVM menghasilkan F1-Score 0,90 dibandingkan Naive Bayes yang hanya 0,79 [8]. Setiawan dkk. (2024) juga

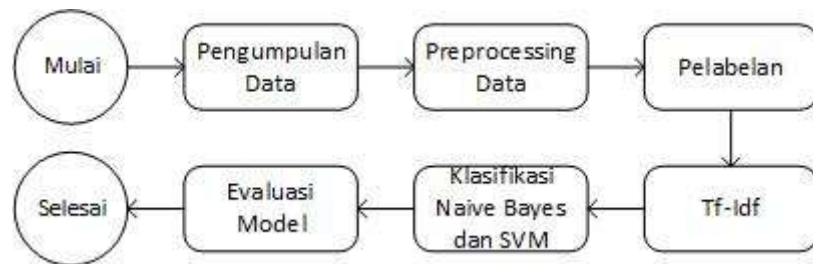
membandingkan kedua algoritma pada analisis sentimen komentar Tiktok produk skincare, dengan hasil akurasi SVM 59,43% lebih unggul dibanding Naive Bayes 47,65% [9].

Meskipun demikian, penelitian-penelitian tersebut belum mengkaji perbandingan Naive Bayes dan SVM yang membahas komentar Tiktok membahas isu kesetaraan gender. Karakteristik dari konten tersebut berbahasa informal dan campuran (Indonesia-Inggris) serta distribusi data yang tidak seimbang. Kebaruan penelitian ini terletak pada penggunaan dataset komentar Tiktok bertopik kesetaraan gender yang dikumpulkan pada rentang waktu 2025–2026, dengan pelabelan otomatis berbasis lexicon yaitu kamus kata yang telah memiliki nilai sentimen tertentu. Selain itu penanganan ketidakseimbangan data menggunakan SMOTE, serta representasi fitur TF-IDF dimana komentar diubah menjadi angka agar dapat diproses oleh algoritma klasifikasi. [10], [11].

Penelitian ini bertujuan merancang model analisis sentimen terhadap opini seputar kesetaraan gender dari komentar pengguna Tiktok menggunakan algoritma Naive Bayes dan SVM. Kedua algoritma tersebut diuji dan dibandingkan kinerjanya menggunakan metrik akurasi, presisi, recall, serta F1-Score untuk mengidentifikasi metode dengan performa terbaik. Hasil penelitian ini diharapkan menjadi acuan bagi kajian analisis sentimen berikutnya, sekaligus memberikan gambaran kecenderungan sikap masyarakat terhadap isu kesetaraan gender di Tiktok.

2. Metode Penelitian

Penelitian ini menerapkan pendekatan kuantitatif komparatif, di mana dua algoritma klasifikasi yakni Naive Bayes dan Support Vector Machine diuji dan dibandingkan performanya menggunakan dataset yang digunakan serta dievaluasi hasilnya. Alur tahapan penelitian secara keseluruhan dapat diamati pada Gambar 1 [12].



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Data yang diambil dalam penelitian ini berupa komentar pengguna Tiktok tentang video yang membahas topik kesetaraan gender. Terdapat beragam opini publik yang diutarakan terhadap video tersebut. Total data yang dikumpulkan berjumlah 1600 komentar yang diambil dari 2025-2026. Seluruh dataset dimuat dalam bentuk format CSV file [13].

2.2. Preprocessing Data

Preprocessing dilakukan sebagai langkah awal untuk menyiapkan data teks agar bersih dan seragam sebelum proses ekstraksi fitur berlangsung. Tahapan preprocessing yang dilakukan adalah sebagai berikut:

- Cleaning:** tahap ini bertujuan menghilangkan elemen-elemen yang tidak memberikan informasi bermakna dalam teks, seperti mention (@username), hashtag (#gender), tautan, angka, tanda baca, serta karakter di luar alfabet.
- Case Folding:** semua karakter teks diubah ke bentuk huruf kecil (lowercase) untuk menyeragamkan format penulisan dan mencegah perbedaan interpretasi yang disebabkan oleh variasi penggunaan huruf kapital [14].
- Normalisasi Teks:** kata-kata tidak baku maupun bentuk slang dikonversi ke padanan kata baku dengan merujuk pada kamus kamuskatabaku.xlsx. Misalnya, kata "gak" dikonversi menjadi "tidak" dan "yaa" menjadi "ya".
- Tokenizing:** kalimat dipecah menjadi unit-unit kata individual (token) menggunakan metode .split(), sehingga setiap kata dapat diproses dan dianalisis secara terpisah [15].
- Filtering Stopwords:** tahap ini mengeliminasi kata-kata yang dianggap tidak signifikan dalam membentuk makna teks, seperti "di", "dan", dan "yang". Daftar stopwords yang digunakan berasal dari NLTK dalam dua bahasa (Indonesia dan Inggris) yang kemudian diperkaya dengan tambahan kata kustom.
- Stemming:** kata-kata berimbuhan dikembalikan ke bentuk dasarnya dengan memanfaatkan library Sastrawi, misalnya kata "kesetaraan" menjadi "setara". [16].

2.3. Pelabelan

Pelabelan dilakukan secara otomatis menggunakan metode berbasis lexicon, yaitu dengan mencocokkan setiap kata pada komentar terhadap kamus yang memuat daftar kata bermuatan positif dan negatif beserta skor masing-masing. Setiap kata yang ditemukan dalam kamus dijumlahkan skornya. Jika total skor lebih besar dari nol, komentar diberi label positif sedangkan jika kurang dari atau sama dengan nol, komentar diberi label negatif. Pendekatan ini banyak digunakan dalam penelitian sentiment analisis karena lebih sederhana [17].

2.4. TF-IDF

Dalam penelitian ini, representasi fitur teks dilakukan menggunakan TF-IDF, sebuah metode pembobotan kata yang mempertimbangkan seberapa sering suatu kata muncul dalam dokumen tertentu sekaligus seberapa jarang kata tersebut dijumpai di seluruh koleksi dokumen. Kata yang memiliki frekuensi kemunculan tinggi dalam sebuah dokumen namun relatif jarang dijumpai pada dokumen-dokumen lainnya akan mendapatkan nilai bobot yang lebih besar, sehingga kata tersebut dianggap lebih merepresentasikan isi dari dokumen yang bersangkutan [18].

2.5. Klasifikasi Model

Setelah proses ekstraksi fitur menggunakan TF-IDF selesai dilakukan, dataset kemudian dibagi menggunakan rasio 80:20, dengan 80% data difungsikan sebagai data latih dan 20% sisanya sebagai data uji. Berikutnya akan dilakukan tahap klasifikasi dengan dua algoritma. Berikut dua model klasifikasinya:

2.5.1. Naive Bayes

Dilatih bekerja dengan menghitung kemungkinan suatu data masuk ke kategori tertentu berdasarkan pola dari data sebelumnya. Berikut adalah rumus teorema bayes [19]:

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

$P(C|X)$ menggambarkan peluang suatu data X diklasifikasikan ke dalam kelas C berdasarkan informasi yang dikandungnya, $P(X|C)$ menunjukkan peluang kemunculan data X apabila diketahui berasal dari kelas C , $P(C)$ adalah probabilitas awal (prior probability) suatu kelas sebelum data diobservasi, sedangkan $P(X)$ mencerminkan peluang keseluruhan kemunculan data X yang berfungsi sebagai faktor normalisasi dalam komputasi probabilitas.

2.5.2. Support Vector Machine

Proses algoritma dari SVM diawali dengan memvisualisasikan data untuk memahami distribusi sentimen yang ada. Langkah yang penting adalah mencari hyperplane, yaitu garis atau pemisah yang berfungsi membagi kelas data agar algoritma dapat membaca data dengan baik. Terdapat berbagai jenis kernel didalam SVM salah satunya yang digunakan adalah kernel linear. Kernel ini bekerja dengan menghitung hasil kali titik (dot product) dari dua vektor langsung pada ruang aslinya, tanpa perlu melakukan pemetaan ke dimensi yang lebih tinggi [20]. Berikut adalah rumus dari kernel linear:

$$K(x,y) = (x \cdot y) \quad (2)$$

$K(x,y)$ merupakan nilai kemiripan antara dua data atau vektor fitur, yaitu x dan y , yang dihitung menggunakan dot product ($x \cdot y$). Fungsi kernel K digunakan untuk mengukur hubungan atau tingkat kesamaan antara data pertama (x) dan data kedua (y) dengan cara mengalikan setiap elemen yang bersesuaian pada kedua vektor, kemudian menjumlahkan hasil perkalian tersebut sehingga diperoleh nilai kemiripan antar data.

2.6. Evaluasi Model

Untuk mengukur kualitas prediksi setiap model, digunakan confusion matrix sebagai alat evaluasi yang mampu memperlihatkan distribusi hasil klasifikasi secara rinci. Selanjutnya, nilai akurasi digunakan sebagai metrik evaluasi untuk mengukur tingkat ketepatan model dalam mengklasifikasikan data sesuai dengan kondisi sebenarnya [27].

3. Hasil dan Diskusi

Data yang digunakan dalam penelitian ini terdiri dari 1.600 komentar Tiktok yang bersumber dari satu konten tunggal, dan dimanfaatkan untuk mengukur serta membandingkan kinerja algoritma Naive Bayes dan SVM. Meskipun hasil pengambilan data menghasilkan beberapa kolom, penelitian ini hanya mengambil kolom komentar sebagai sumber data utama.

3.1. Pengumpulan Data

Pengambilan data dilakukan melalui penerapan teknik web scraping dengan memanfaatkan website Apify. Video yang diambil membahas kesetaraan gender yang memberikan informasi dari perspektif antara laki-laki dan perempuan. Total data yang dikumpulkan berjumlah 1600 komentar yang diambil dari rentang waktu 2025-2026. Seluruh dataset dimuat dalam bentuk format CSV file yang dapat dilihat pada Tabel 1.

Tabel 1. Data komentar Tiktok

No	Komentar
1.	men are so obsessed with "I can hit you because we are equal now" when it's about equality 🤔🤔🤔🤔🤔
2.	bahas kesetaraan dan di kepala langsung tertuju ke kekerasan aja udh aneh
3.	emang perempuan bisa jadi kepala keluarga? padahal di indo banyak yang ga dapat peran ayah ☐☐
4.	Man can protect Woman
5.	it's always that damn patriarchy man

3.2. Preprocessing Data

Tahapan dalam preprocessing data dimulai dari cleaning text, case folding, normalisasi, tokenizing, filtering, dan diakhiri dengan stemming. Pada tahapan preprocessing bertujuan agar model dapat melakukan klasifikasi dengan baik. Berikut adalah hasil preprocessing data dapat dilihat pada Tabel 2.

Tabel 2. Preprocessing Data

No	Cleansing	Case Folding	Normalisasi	Tokenizing	Filltering	Stemming	Hasil Akhir
1.	men are so obsessed with I can hit you becaus...	men are so obsessed with i can hit you because...	men are sok obsessed with i can hit you becaus...	['men', 'are', 'sok', 'obsessed', 'with', 'i',...]	['men', 'sok', 'obsessed', 'hit', 'equal', 'eq...]	men sok obsessed hit equal equality	men sok obsessed hit equal equality
2.	bahas kesetaraan dan di kepala langsung tertuj...	bahas kesetaraan dan di kepala langsung tertuj...	bahas kesetaraan dan di kepala langsung tertuj...	['bahas', 'kesetaraan', 'dan', 'di', 'kepala',...]	['bahas', 'kesetaraan', 'kepala', 'langsung', 'keras', 'aneh', 'pria pri...]	bahas tara kepala langsung keras aneh	bahas kesetaraan kepala langsung kekerasan ane...
3.	emang perempuan bisa jadi kepala keluarga pada...	emang perempuan bisa jadi kepala keluarga pada...	memang perempuan bisa jadi kepala keluarga pad...	['memang', 'perempuan', 'bisa', 'jadi', 'kepal...]	['perempuan', 'kepala', 'keluarga', 'indonesia...]	perempuan kepala keluarga indonesia peran ayah	perempuan kepala keluarga indonesia peran ayah
4.	Man can protect Woman protect from what	man can protect woman protect from what	man can protect woman protect from what	['man', 'can', 'protect', 'woman', 'protect', ...]	['man', 'protect', 'woman', 'protect']	man protect woman protect	man protect woman protect
5.	its always that damn patriarchy man	its always that damn patriarchy man	its always that damn patriarchy man	['its', 'always', 'that', 'damn', 'patriarchy'...]	['always', 'damn', 'patriarchy', 'man']	always damn patriarchy man	always damn patriarchy man

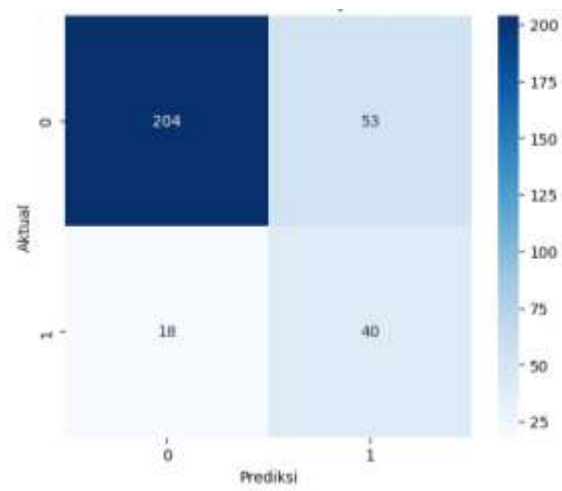
3.3. Pelabelan Data

Penelitian ini selanjutnya dilakukan pelabelan data dengan menerapkan metode lexicon. Metode ini memudahkan pelabelan pada teks review yang belum dilabeli. Pelabelan lexicon dilakukan dengan otomatis sehingga sentimen bernilai antara positif dan negatif. Setiap kata pada komentar akan dicocokkan dengan kamus lexicon. Berikut Tabel 3 adalah hasil dari sentimen dengan pelabelan lexicon.

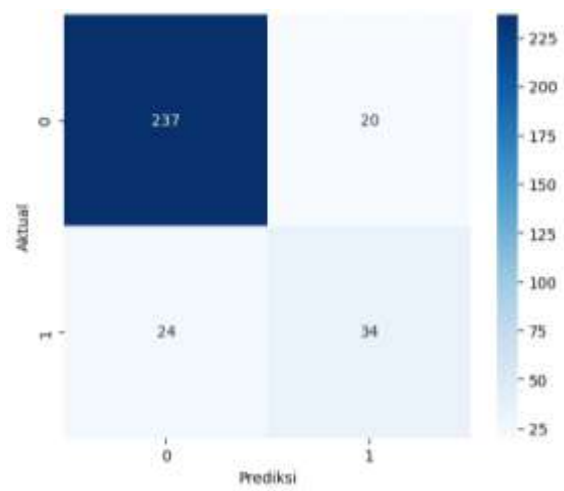
Pada hasil klasifikasi antara metode naive bayes dan svm diatas memperlihatkan bahwa hasil akurasi lumayan jauh perbedaanya. Klasifikasi dengan menggunakan naive bayes menunjukkan akurasi 77% yang berarti sudah cukup baik dalam mengenali teks. Selanjutnya hasil klasifikasi pada metode SVM menunjukkan angka 86%, dimana hasilnya lebih baik dibandingkan dengan naive bayes. Model Naive Bayes cukup baik dalam mendeteksi sentimen negatif dengan hasil f1-score sebesar 85%, namun kurang optimal untuk sentimen positif f1-score hasilnya 53%, dengan weighted average f1-score sebesar 79%. Model SVM menunjukkan performa lebih baik dengan akurasi 86%. Sentimen negatif terdeteksi sangat baik dengan f1-score bernilai 91%, sementara sentimen positif mencapai f1-score 60%, dengan weighted average f1-score 85%.

3.5. Evaluasi Model

Pada tahap Evaluasi model ini menggunakan confusion matrix. Confusion matrix adalah tabel yang digunakan dalam evaluasi model untuk memvisualisasikan jumlah data yang berhasil dan gagal diprediksi secara benar oleh model klasifikasi. Berikut adalah confusion matrix metode naive bayes pada Gambar 5, sedangkan metode svm pada Gambar 6.



Gambar 5. Wordcloud Positif



Gambar 6. Wordcloud Positif

Berdasarkan hasil confusion matrix, Model masih menghasilkan kesalahan pada 53 data berlabel negatif yang terklasifikasi sebagai positif, serta 18 data berlabel positif yang terklasifikasi sebagai negatif. Adapun model SVM berhasil mengklasifikasikan 237 data negatif dan 34 data positif secara tepat, dengan kesalahan berupa 20 data negatif yang terprediksi sebagai positif dan 24 data positif yang terprediksi sebagai negatif. Dari perbandingan tersebut, SVM memperlihatkan tingkat kesalahan yang lebih rendah dibandingkan Naive Bayes.

Perbandingan kedua model menunjukkan SVM lebih unggul dari Naive Bayes. Hal ini disebabkan Naive Bayes mengasumsikan setiap kata bersifat independen, padahal dalam teks nyata antar kata saling berkaitan. SVM mampu menangani kesusahan data teks dengan lebih baik karena bekerja berdasarkan batas pemisah antar kelas, bukan distribusi probabilitas. Naive Bayes tetap cenderung bias ke kelas negatif. SVM lebih berhasil mengatasinya, terbukti dari precision kelas positif SVM 63% yang jauh lebih baik dibanding Naive Bayes 43%.

Data komentar bersumber dari video TikTok yang memuat dua perspektif berbeda seputar kesetaraan gender. Sudut pandang laki-laki mengenai kesetaraan gender, serta harapan dan keinginan perempuan terkait isu yang sama. Perbedaan sudut pandang ini memicu banyak respons yang bersifat menolak, tidak setuju, atau mempermasalahkan perspektif yang ditampilkan, sehingga wajar jika komentar negatif mendominasi.

4. Kesimpulan

Hasil eksperimen menunjukkan bahwa SVM berhasil mencapai tingkat akurasi 86%, sementara Naive Bayes menghasilkan akurasi sebesar 77% pada dataset yang sama. Dengan demikian, tujuan penelitian untuk mengidentifikasi algoritma yang paling efektif telah tercapai. SVM sebagai metode terbaik pada klasifikasi sentimen data komentar TikTok tentang kesetaraan gender. Dominasi sentimen negatif ini erat kaitannya dengan topik yang dianalisis, yaitu video TikTok yang menampilkan perspektif laki-laki dan harapan perempuan seputar

kesetaraan gender. Perbedaan sudut pandang dalam konten yang dipilih memicu banyak respons ketidaksetujuan dari para pengguna, yang mencerminkan bahwa isu kesetaraan gender masih menjadi topik sensitif. Penelitian ini diharap bisa memberikan wawasan baru kepada peneliti yang ingin menggarap hal serupa. Dikarenakan masih banyaknya keterbatasan dalam penelitian ini diharapkan untuk peneliti selanjutnya dapat memperbanyak dataset yang digunakan untuk analisis. Selain itu peneliti berikutnya juga dapat menambahkan beberapa metode machine learning lain ataupun metode deep learning lainnya. Penambahan proses pelabelan manual juga dapat dilakukan untuk meningkatkan kualitas label sentimen dibandingkan hanya menggunakan metode lexicon otomatis. Dengan pengembangan tersebut, diharapkan hasil analisis sentimen menjadi lebih akurat, stabil, dan mampu memahami konteks bahasa dengan lebih baik.

Referensi

- [1] P. Deski Manalu et al., "Implementasi Algoritma Klasifikasi untuk Analisis Sentimen Media Sosial Tiktok Tahun 2025," *Jurnal Teknik Informatika dan Teknologi Informasi*, doi: 10.55606/jutiti.v5i1.5644.
- [2] S. F. L. Aniqurrohman, "Kesetaraan Gender Dan Nilai Nilai Yang Terkandung Di Dalamnya Menurut Hak Asasi Manusia," *Jurnal Dunia Ilmu Hukum (JURDIKUM)*, vol. 1, no. 2, pp. 50–56, Oct. 2023, doi: 10.59435/jurdikum.v1i2.170.
- [3] A. Anindya, A. Hanana, and N. Elian, "Upaya Meningkatkan Kesadaran Gender Kaum Muda melalui Pesan Kesetaraan Gender di Media Sosial Instagram," doi: 10.25077/rk.5.2.103-112.2021
- [4] P. D. P. Silitonga and M. K. M. Nasution, "Dari Sintaksis ke Pemahaman : Natural Language Processing (NLP) dalam Filsafat Ilmu Komputer," *Jurnal ICT : Information Communication & Technology*, vol. 25, pp. 181–186, doi: 10.36054/jict-ikmi.v25i2.336.
- [5] R. F. P. Pratama and W. Maharani, "Comparative Analysis of Naive Bayes and SVM for Improved Emotion Classification on Social Media," *Eumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 11–20, Apr. 2025, doi: 10.29408/edumatic.v9i1.29087.
- [6] D. Auliya Agustina, S. Subanti, E. Zukhronah, P. Studi Statistika, and U. Sebelas Maret, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine," 2020, doi: 10.20961/ijas.v3i2.44337
- [7] N. S. Wardani, A. Prahutama, and P. Kartikasari, "Analisis Sentimen Pemindahan Ibu Kota Negara dengan Klasifikasi Naïve Bayes untuk Model Bernoulli dan Multinomial," *Jurnal Gaussian*. vol. 9, 2020, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [8] D. Shabira et al., "MIND (Multimedia Artificial Intelligent Networking Database Deteksi Seksisme Online menggunakan Support Vector Machine dan Naïve Bayes," *Journal MIND Journal | ISSN*, vol. 8, no. 2, pp. 254–266, 2023, doi: 10.26760/mindjournal.v8i2.254-266.
- [9] T. Setiawan, S. Liem, and D. M. R. Pribadi, "Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Komentar Tiktok pada Produk Skincare," 2024. [Online]. Available: <https://jurnal.politap.ac.id/index.php/aicoms>
- [10] A. Alif Nurrahman, E. Moza, R. Md, R. Roshan, A. Rizky, and H. Irsyad, "Opini Publik terhadap Isu Pengoplosan Pertamax di Youtube Menggunakan Metode Naive Bayes," *AICOMS 2025*, vol. 4, no. 2, pp. 55–61, 2025, doi: 10.58466/aicoms.v4i1.1990.
- [11] Y. Asri, W. N. Suliyanti, D. Kuswardani, and M. Fajri, "Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile," *PETIR*, vol. 15, no. 2, pp. 264–275, Nov. 2022, doi: 10.33322/petir.v15i2.1733.
- [12] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [13] S. Rihastuti and A. Rosyidi, "Analisis Sentimen Pengguna TikTok tentang Progres Pembangunan IKN dengan Metode Random Forest" *Journal of Computer Science and Technology JCS-TECH*, vol. 5, no. 1, pp. 19–23, 2025, doi: <https://doi.org/10.54840/jcstech.v5i1.345>
- [14] A. Oktavia Praneswara and N. Cahyono, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, p. 3925, 2023, doi: <https://doi.org/10.33022/ijcs.v12i6.3473>
- [15] A. Septiani et al., "Analisis Sentimen Menggunakan Algoritma Naïve Bayes terhadap Regulasi TikTok Shop pada Media Sosial X (Twitter)" 2024, doi: 10.36040/jati.v8i4.9351.
- [16] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 1, pp. 84–93, Apr. 2023, doi: 10.34010/jamika.v13i1.9424.
- [17] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan Lexicon-Based pada Performa IndoBERT," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 3, Jul. 2025, doi: 10.57152/malcom.v5i3.2151.
- [18] N. F. Hilmi and F. Irwiensyah, "Analisis Sentimen Terhadap Aplikasi Tiktok Dari Ulasan Pada Google Playstore Menggunakan Metode Naïve Bayes," *SMATIKA JURNAL*, vol. 14, no. 01, pp. 146–156, Jul. 2024, doi: 10.32664/smatika.v14i01.1210.
- [19] A. Ridwan, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," 2020, doi: 10.47970/siskom-kb.v4i1.169.
- [20] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, Oct. 2023, doi: 10.57152/malcom.v3i2.897.