JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Analisis Sentimen Komentar Cyberbullying Terhadap Fenomena Flexing di Tiktok Menggunakan Artificial Neural Network

Lailatul Qodriyah*, Ifnu Wisma Dwi Prastya, Guruh Purbo Dirgontoro

Fakultas Sains Teknologi, Teknik Informatika, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia
Email: [1*]l4ilatulqodriyah11@gmail.com, [2]ifnuprastya@unugiri.ac.id, [3]gputrad@gmail.com
Email Penulis Korespondensi: l4ilatulqodriyah11@gmail.com
Submitted **12-01-2026;** Accepted **24-02-2026;** Published **28-02-2026**

Abstrak

Fenomena flexing pada platform TikTok memunculkan dinamika interaksi digital yang kompleks dan berpotensi memicu komentar bernada negatif maupun indikasi cyberbullying. Penelitian ini bertujuan untuk menganalisis kecenderungan sentimen pengguna terhadap konten flexing serta mengevaluasi performa algoritma Artificial Neural Network (ANN) dalam mengklasifikasikan komentar tersebut. Sebanyak 4.013 komentar dikumpulkan melalui proses scraping pada akun TikTok Miechel Halim dan diberi label secara otomatis menggunakan pendekatan lexicon-based. Teks selanjutnya diproses melalui pembersihan data dan representasi Term Frequency Inverse Document Frequency (TF-IDF), kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20. Model ANN dilatih dalam dua skenario, yaitu sebelum dan sesudah penerapan Synthetic Minority Oversampling Technique (SMOTE) untuk mengatasi ketidakseimbangan label. Hasil pengujian menunjukkan akurasi sebesar 89,79% pada model awal dan meningkat menjadi 90,04% setelah penerapan SMOTE, dengan perbaikan pada recall kelas negatif yang sebelumnya sulit dikenali. Temuan ini menunjukkan bahwa ANN mampu mengklasifikasikan sentimen komentar TikTok secara efektif, meskipun konteks bahasa informal dan distribusi kelas yang tidak merata tetap menjadi tantangan dalam deteksi komentar negatif terkait fenomena cyberbullying.

Kata Kunci: Analisis Sentimen; Komentar TikTok; Deteksi Cyberbullying; Konten Flexing; Artificial Neural Network; TF-IDF

Abstract

The rising trend of flexing on TikTok has created a dynamic digital space that often triggers varied user reactions, including subtle forms of cyberbullying. This study aims to analyze public sentiment toward flexing content and evaluate the performance of the Artificial Neural Network (ANN) algorithm in classifying user comments. A total of 4,013 comments were collected through a scraping process on the TikTok account of Miechel Halim and automatically labeled using a lexicon-based approach. The comments were then pre-processed and transformed into Term Frequency–Inverse Document Frequency (TF-IDF) representations before being split into training and testing datasets with an 80:20 ratio. The ANN model was trained under two scenarios before and after the application of the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance. Experimental results show that the initial model achieved an accuracy of 89.79%, which increased to 90.04% after SMOTE, accompanied by an improvement in the recall of the negative class. These findings indicate that ANN is effective for sentiment classification of TikTok comments, although informal language patterns and highly imbalanced labels remain challenges in identifying negative or potentially harmful remarks related to cyberbullying.

Keywords: Sentiment Analysis; TikTok Comments; Cyberbullying Detection; Flexing Content; Artificial Neural Network; TF-IDF

1. PENDAHULUAN

Perkembangan platform video pendek seperti TikTok telah membentuk pola komunikasi digital baru yang cepat, padat, dan berbasis interaksi komentar, terutama di kalangan generasi muda [1]. Sejumlah penelitian menunjukkan bahwa konten *flexing* di media sosial telah menjadi objek kajian dalam analisis sentimen, khususnya pada generasi Z yang aktif berinteraksi melalui kolom komentar [2]. Salah satu tren yang menonjol adalah *flexing*, yakni praktik menampilkan kemewahan, barang bermerek, atau gaya hidup glamor sebagai strategi ekspresi diri dan pembentukan citra di ruang digital [3]. Dalam konteks pemasaran dan pengaruh kreator, paparan konten dari *social media influencers* dapat mendorong perilaku konsumsi yang menonjol melalui mekanisme perbandingan sosial, *fear of missing out* (FOMO), dan kecenderungan materialisme [4]. Namun konten *flexing* tidak selalu diterima sebagai hiburan, sejumlah kajian menunjukkan bahwa paparan kemewahan dapat memicu kecemburuan sosial serta respons negatif berupa komentar sinis hingga indikasi *cyberbullying* [5].

Komentar pada TikTok memiliki karakteristik linguistik yang unik, seperti singkat, spontan, sarat bahasa gaul, hiperbola, dan emoji, serta sangat dipengaruhi oleh konteks visual konten [6]. Penelitian terkait *cyberbullying* pada media sosial juga menunjukkan bahwa ujaran agresif sering kali muncul dalam bentuk tidak langsung, seperti sarkasme atau candaan yang bernada menyerang [7][8]. Selain itu, distribusi sentimen pada komentar TikTok cenderung tidak seimbang, di mana komentar netral mendominasi sebagian besar interaksi pengguna [9]. Ketidakseimbangan ini berpotensi menurunkan kemampuan model klasifikasi dalam mengenali kelas minoritas, terutama komentar negatif yang relevan dengan perundungan daring. Pada level praktik berbasis data, laporan uji coba TikTok *Research API* juga menunjukkan bahwa perundungan di kolom komentar dapat muncul dalam berbagai bentuk, seperti ujaran kebencian, pelecehan, maupun ancaman, sehingga menegaskan perlunya pendekatan analitik yang lebih sistematis dalam menganalisis komentar TikTok [10].

Permasalahan utama dalam penelitian ini adalah bagaimana mengklasifikasikan sentimen komentar terhadap konten *flexing* secara akurat pada teks yang bersifat informal, kontekstual, dan memiliki distribusi kelas yang tidak seimbang sehingga deteksi komentar negatif sebagai indikasi *cyberbullying* dapat dilakukan secara lebih andal. Sejalan dengan kebutuhan analisis teks yang adaptif, metode kecerdasan buatan seperti *Artificial Neural Network* (ANN) banyak

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

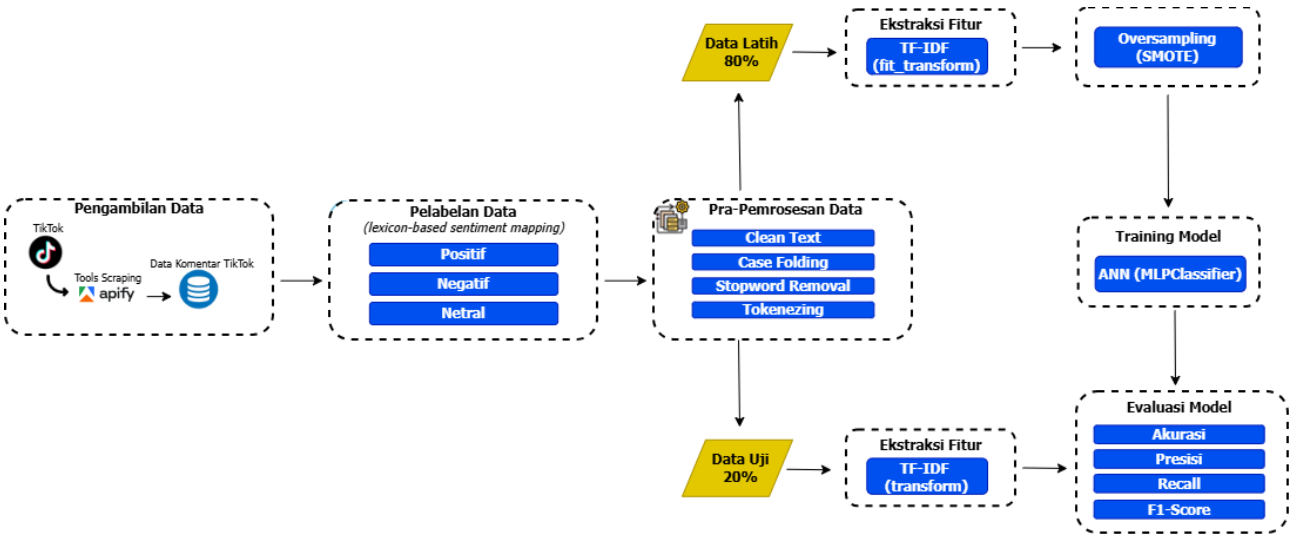
digunakan karena kemampuannya dalam menangkap pola non-linear pada data berdimensi tinggi [11][12]. Dalam konteks analisis sentimen, ANN dilaporkan memiliki performa yang kompetitif dibandingkan metode tradisional seperti *Naïve Bayes* [13] *dan Support Vector Machine* (SVM) [14]. Untuk mengatasi keterbatasan anotasi manual pada data berskala besar, pendekatan *lexicon-based* sering digunakan sebagai metode pelabelan awal pada teks informal [15]. Meski demikian, tinjauan sistematis terbaru menunjukkan bahwa deteksi *cyberbullying* masih menghadapi tantangan pada keragaman bahasa, konteks, serta ketimpangan distribusi data, sehingga diperlukan perancangan pipeline analisis yang lebih terstruktur dan komprehensif [16].

Meskipun berbagai penelitian telah membahas ujaran kebencian umum atau analisis sentimen pada platform lain, kajian yang secara spesifik memfokuskan komentar terhadap konten *flexing* sebagai pemicu interaksi negatif di TikTok masih terbatas. Berbeda dengan penelitian sebelumnya yang berfokus pada analisis sentimen umum atau perbandingan algoritma klasifikasi tertentu, penelitian ini secara khusus mengkaji komentar terhadap konten *flexing* sebagai pemicu interaksi negatif serta mengintegrasikan pelabelan berbasis leksikon, pembobotan TF-IDF, dan penanganan ketidakseimbangan kelas dalam kerangka *Artificial Neural Network* (ANN) secara terpadu. Terlebih lagi, integrasi antara pelabelan berbasis leksikon, pembobotan TF-IDF, serta penyeimbangan kelas menggunakan teknik *oversampling* dalam satu kerangka eksperimen berbasis ANN belum banyak dikaji secara sistematis. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna TikTok terhadap konten *flexing* serta mengevaluasi efektivitas *Artificial Neural Network* (ANN) dalam mengklasifikasikan komentar ke dalam tiga kelas, yaitu positif, netral, dan negatif, guna memperkaya bukti empiris serta meningkatkan sensitivitas model dalam mendeteksi komentar negatif pada konteks *cyberbullying* [17].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini merupakan penelitian pengembangan yang berfokus pada perancangan dan evaluasi model *machine learning* untuk menganalisis sentimen komentar *cyberbullying* terhadap fenomena *flexing* di media sosial TikTok. Penelitian ini memanfaatkan teknik *scraping* berbasis *Apify* untuk memperoleh data komentar, yang selanjutnya diproses dengan pelabelan sentimen otomatis menggunakan pendekatan *lexicon-based*. Setelah itu, komentar melalui tahap pra-pemrosesan yang mencakup pembersihan teks, *case folding*, penghapusan *stopword*, dan tokenisasi. Pembagian dataset dilakukan setelah tahap prapemrosesan, dengan proporsi 80% sebagai data latih dan 20% sebagai data uji. Pada tahap pelatihan, fitur teks diekstraksi menggunakan metode TF-IDF, kemudian dilakukan penerapan SMOTE untuk menyeimbangkan jumlah data pada setiap kelas. Model kemudian dilatih menggunakan algoritma *Artificial Neural Network* (ANN) dan dievaluasi menggunakan akurasi, *presisi*, *recall*, dan F1-*score*. Tahapan ini dirancang untuk menghasilkan model klasifikasi sentimen yang mampu mendeteksi komentar negatif secara akurat pada konteks *cyberbullying* di platform *TikTok* [18]. Alur tahapan penelitian yang dilakukan dalam studi ini disajikan pada Gambar 1. Diagram tersebut menggambarkan proses mulai dari pengambilan data, pelabelan sentimen, prapemrosesan teks, ekstraksi fitur menggunakan TF-IDF, penanganan ketidakseimbangan kelas dengan SMOTE, hingga proses pelatihan dan evaluasi model ANN.



Gambar 1. Alur proses penelitian

Berdasarkan Gambar 1, proses penelitian dirancang secara sistematis dengan memisahkan data latih dan data uji sebelum tahap pelatihan model. Penerapan SMOTE dilakukan hanya pada data latih untuk menghindari kebocoran data (data *leakage*) serta menjaga objektivitas evaluasi model.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.2 Proses Pengumpulan Data (Scraping)

Data penelitian dikumpulkan melalui teknik scraping dengan memanfaatkan platform *Apify*, yaitu sebuah *framework web-scraping* yang memungkinkan pengambilan data media sosial secara terstruktur. Peneliti memfokuskan pengambilan data pada akun TikTok milik Miechel Halim, seorang *content creator* yang dikenal dengan konten bertema *flexing* seperti *outfit* mewah, barang *branded*, serta gaya hidup glamor. Dari hasil *scraping*, diperoleh sebanyak 4.013 komentar publik dengan dua kolom utama yaitu *No* dan *Comment*. Proses scraping dilakukan dengan tetap mempertahankan keaslian data tanpa mengubah makna komentar, serta mengikuti ketentuan etika dalam pengumpulan data digital [19]. Pendekatan serupa juga digunakan dalam penelitian *cyberbullying detection* pada konten TikTok oleh Roy et al. [20] dan Prameswari et al. [21] untuk memperoleh data komentar secara real-time.

2.3 Pelabelan Data (Labeling)

Tahap *labeling* dilakukan secara automatic labeling menggunakan pendekatan *lexicon-based sentiment mapping*, di mana sistem secara otomatis menentukan label sentimen (positif, negatif, atau netral) berdasarkan kamus kata kunci yang telah disusun sebelumnya. Kamus sentimen negatif mencakup kata-kata kasar, makian, penghinaan sosial, serta ekspresi kebencian dan ketidaksukaan yang umum digunakan dalam konteks cyberbullying dan bahasa informal TikTok, seperti anj, anjing, goblok, tolol, idiot, bacot, kampungan, norak, alay, hina, sampah, brengsek, bego, parah, menjijikkan, hujat, hate, ilfil, aneh, ribet, sinting, dan stress. Sementara itu, kamus sentimen positif memuat kata-kata pujian, dukungan, serta ekspresi apresiatif yang sering muncul pada komentar TikTok, meliputi kata seperti keren, mantap, bagus, hebat, cakep, cantik, terbaik, wow, salut, amazing, menyala, glowing, outfit, body goals, elegan, gokil, inspiratif, bahagia, lucu, manis, dan panutan. Proses pelabelan menghasilkan distribusi sentimen sebesar 1.056 positif**,** 169 negatif, dan 2.788 netral dari total 4.013 komentar. Pendekatan ini merujuk pada metode *lexicon-based sentiment analysis* sebagaimana digunakan oleh Wirawan et al. (2023) [22] dan Pujari & Susmitha (2024) [23] yang menekankan efektivitas sistem otomatis berbasis kamus dalam analisis teks pendek di media sosial.

2.4 Prapemrosesan Data (Preprocessing)

Tahap ini bertujuan untuk membersihkan teks sebelum diubah menjadi representasi numerik. Proses *preprocessing* dilakukan melalui beberapa langkah, yaitu:

a. Pembersihan teks
b. Mengubah seluruh karakter menjadi huruf kecil (*case folding*)
c. Menghapus *stopword* Bahasa Indonesia
d. Melakukan tokenisasi menggunakan *word_tokenize*

Hasil akhir dari tahap ini berupa kolom *Clean_Comment*, yaitu teks komentar yang telah dinormalisasi. Langkah ini mengikuti pendekatan *preprocessing* yang digunakan pada penelitian Dianda Rifaldi, Abdul Fadhil, & Herman (2023) [24] yang menekankan pentingnya normalisasi bahasa dalam analisis teks media sosial.

2.5 Pembagian Data (Train–Test Split)

Dataset yang telah melalui tahap prapemrosesan selanjutnya dipisahkan menjadi data latih dan data uji dengan perbandingan 80% dan 20%. Untuk menjaga konsistensi hasil, pembagian data dilakukan secara acak dengan menetapkan parameter *random_state* tertentu. Praktik pembagian data ini umum digunakan dalam penelitian analisis sentimen berbasis *machine learning* dan *Artificial Neural Network* [25]. Pada tahap ini SMOTE belum diterapkan karena oversampling hanya boleh dilakukan pada data latih untuk menghindari *data leakage* dan menjaga validitas evaluasi model.

2.6 Pembobotan Fitur Teks menggunakan TF-IDF

Pada tahap ini, setiap komentar diubah ke dalam bentuk numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini bekerja dengan memberikan bobot yang lebih besar pada kata-kata yang jarang muncul tetapi memiliki peran penting dalam membedakan isi antar dokumen. Dengan pendekatan ini, model lebih mampu menangkap kata-kata kunci yang berpengaruh terhadap penentuan kelas sentimen.

Dalam penelitian ini, *TfidfVectorizer* dari *scikit-learn* digunakan untuk membentuk representasi fitur teks dengan batas maksimum 5.000 fitur. Penggunaan TF-IDF telah banyak diterapkan dalam penelitian analisis sentimen dan terbukti mampu meningkatkan performa model klasifikasi teks. Efektivitas metode TF-IDF dalam meningkatkan akurasi model *Support Vector Machine* pada klasifikasi ulasan berbahasa Indonesia telah dibuktikan dalam penelitian Westley, Thomas, dan Rumaisa (2023) [26]. Temuan-temuan tersebut menjadi dasar kuat penggunaan TF-IDF dalam penelitian ini.

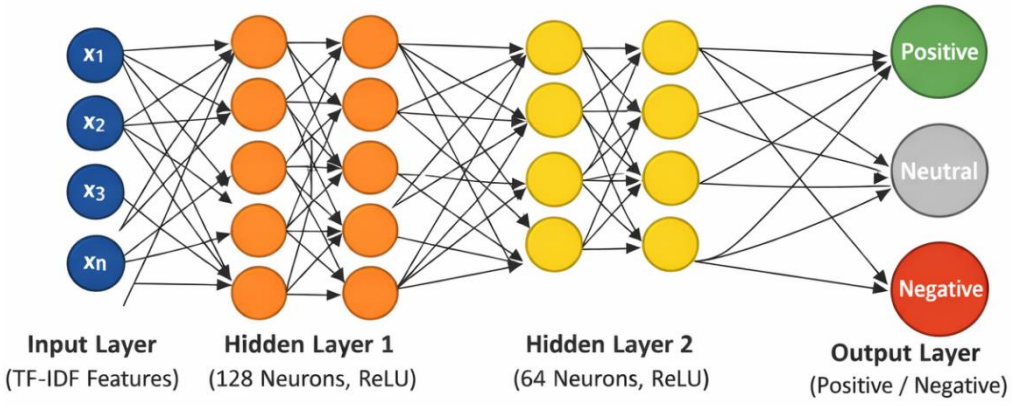
2.7 Model Artificial Neural Network (ANN)

Artificial Neural Network (ANN) merupakan metode pembelajaran mesin yang umum digunakan dalam klasifikasi teks karena kemampuannya memodelkan hubungan nonlinier pada data berdimensi tinggi. ANN bekerja melalui struktur berlapis yang terdiri atas neuron-neuron saling terhubung, di mana bobot koneksi diperbarui secara iteratif selama proses pelatihan sehingga mampu menangkap pola kompleks pada data teks [27].

Dalam penelitian ini, ANN diterapkan untuk mengklasifikasikan sentimen komentar berdasarkan fitur teks hasil pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF). Model yang digunakan berupa jaringan

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

feedforward multilayer perceptron dengan fungsi aktivasi *Rectified Linear Unit* (ReLU) dan proses optimasi menggunakan Adam optimizer, yang efektif dan banyak digunakan dalam klasifikasi sentimen berbahasa Indonesia [28]. Arsitektur ANN yang digunakan terdiri atas input layer, dua hidden layer, dan output layer, Arsitektur model *Artificial Neural Network* yang digunakan dalam penelitian ini ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur ANN

Gambar 2 menunjukkan bahwa model terdiri dari tiga lapisan utama, yaitu input layer yang menerima fitur TF-IDF, dua hidden layer dengan fungsi aktivasi ReLU untuk menangkap pola non-linear, serta output layer yang menghasilkan klasifikasi sentimen. Penggunaan dua hidden layer bertujuan meningkatkan kemampuan representasi model dalam mengenali pola kompleks pada teks komentar yang bersifat informal.

Proses komputasi pada setiap neuron dilakukan dengan menghitung kombinasi linier antara input dan bobot, yang dirumuskan sebagai berikut:

$$z = \sum_{i=1}^{n} w_i \, x_i + b \quad (1)$$

Dengan keterangan sebagai berikut:
x_i: nilai input ke-i, yaitu bobot fitur hasil pembobotan TF-IDF
w_i: bobot koneksi antara input ke-i dan neuron
b: bias neuron
n: jumlah fitur input
z: nilai hasil kombinasi linier sebelum fungsi aktivasi

Nilai z selanjutnya diproses menggunakan fungsi aktivasi ReLU yang dirumuskan sebagai berikut:

$$f(z) = \max(0, z) \quad (2)$$

Dengan keterangan:
z: nilai input bersih neuron
$f(z)$: nilai keluaran neuron setelah fungsi aktivasi

Untuk mengukur kesalahan antara hasil prediksi model dan label aktual, digunakan *fungsi loss categorical cross-entropy,* yang umum digunakan pada permasalahan klasifikasi multikelas. Fungsi loss tersebut dirumuskan sebagai berikut:

$$L = -\sum_{j=1}^{k} y_j \log(\hat{y}_j) \quad (3)$$

Dengan keterangan sebagai berikut:
L: nilai loss atau kesalahan prediksi model
k: jumlah kelas sentimen (positif, netral, dan negatif)
y_j: label aktual untuk kelas ke-j
$\hat{y}_j$: probabilitas hasil prediksi model untuk kelas ke-j
$\log(\hat{y}_j)$: logaritma natural dari probabilitas prediksi

2.8 Penerapan Synthetic Minority Over-sampling Technique (SMOTE)

Synthetic Minority Over-sampling Technique (SMOTE) diterapkan dalam penelitian ini sebagai strategi penanganan ketidakseimbangan kelas pada data sentimen. SMOTE merupakan metode oversampling yang membangkitkan data sintetis untuk kelas minoritas berdasarkan kedekatan antar sampel dalam ruang fitur sehingga dapat memperkaya variasi data tanpa duplikasi langsung [29]. Penerapan SMOTE dilakukan setelah tahap ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF) dan hanya diterapkan pada data latih agar distribusi kelas lebih seimbang sebelum model dilatih ulang. Pendekatan ini telah banyak digunakan dalam penelitian analisis sentimen berbahasa Indonesia dan terbukti membantu model klasifikasi dalam mempelajari pola linguistik di kelas minoritas dengan lebih efektif [30].

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.9 Metode Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja *Artificial Neural Network* (ANN) dalam mengklasifikasikan sentimen komentar menggunakan data uji yang tidak terlibat dalam proses pelatihan. Kinerja model diukur menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-*score*, yang secara luas digunakan dalam penelitian klasifikasi teks dan analisis sentimen berbahasa Indonesia karena mampu merepresentasikan performa model secara komprehensif, khususnya pada data dengan distribusi kelas yang tidak seimbang [31]. Selain itu, *confusion matrix* digunakan untuk memberikan gambaran detail mengenai kesalahan prediksi model pada setiap kelas sentimen.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian serta interpretasi yang diperoleh dari proses analisis sentimen terhadap komentar TikTok terkait fenomena *flexing*. Tahapan hasil yang disajikan meliputi distribusi data, prapemrosesan, pembobotan TF-IDF, pelatihan model *Artificial Neural Network* (ANN) sebelum dan sesudah SMOTE, serta implementasi dan pengujian model. Selain itu, pembahasan juga mengaitkan temuan teknis dengan fenomena sosial *flexing* di Indonesia untuk memberikan gambaran yang lebih komprehensif.

3.1 Hasil Penelitian

a. Distribusi Data Komentar

Penelitian ini menggunakan komentar pengguna TikTok pada konten kreator Miechel Halim sebagai sumber data, dengan karakteristik konten yang menonjolkan gaya hidup mewah dan praktik *flexing*. Data dikumpulkan dari beberapa video yang diunggah pada tanggal 13 April 2025, 25 April 2025, 27 April 2025, 8 Mei 2025, 22 Mei 2025, 28 Juni 2025, 10 Juli 2025, 29 Juli 2025, 11 Agustus 2025, dan 9 September 2025 sehingga merepresentasikan respons pengguna dalam rentang waktu yang berbeda. Tabel 1 menampilkan distribusi label tiktok.

Tabel 1. Distribusi Label Komentar TikTok

Label	Jumlah	Persentase
Positif	1.056	26.31%
Negatif	169	4.21%
Netral	2.788	69.47%
Total	**4.013**	**100%**

Data hasil pelabelan otomatis terbagi ke dalam tiga kelas sentimen, yakni positif, negatif, dan netral. Dari total 4.013 komentar, sentimen netral mendominasi dengan 2.788 komentar, sementara sentimen positif dan negatif masing-masing berjumlah 1.056 dan 169 komentar. Dominasi sentimen netral ini mengindikasikan bahwa sebagian besar pengguna TikTok cenderung memberikan respons pasif terhadap konten *flexing* yang ditampilkan.

b. Hasil Prapemrosesan Teks

Prapemrosesan dilakukan untuk membersihkan teks dari berbagai elemen yang tidak relevan, termasuk emoji, tanda baca, URL, dan angka, sekaligus menormalkan teks ke dalam bentuk yang lebih konsisten. Hasil prapemrosesan menghasilkan teks dengan struktur yang lebih sederhana sehingga mudah diolah menggunakan TF-IDF. Contoh perubahan teks sebelum dan sesudah proses pembersihan data ditampilkan pada Tabel 2.

Tabel 2. Sebelum dan sesudah pembersihan data

Sebelum Pembersihan Data	Setelah Pembersihan Data
MAAAAUUU	MAAAAUUU
mau dear... akuuuu Xtra Small 😻	mau dear akuuuu Xtra Small
Bajunya bisa buat hijab ga? Pake manset kulit gtu	Bajunya bisa buat hijab ga Pake manset kulit gtu

Pada tahap pembersihan data yang dipaparkan pada Tabel 2, elemen-elemen non-teks seperti emoji, tanda baca, dan simbol dihilangkan agar tidak mengganggu proses analisis. Penghapusan elemen tersebut dilakukan tanpa mengubah isi utama komentar, sehingga informasi linguistik yang berkaitan dengan sentimen tetap terjaga.

Tahapan normalisasi huruf melalui proses case folding ditunjukkan pada Tabel 3.

Tabel 3. Sebelum dan sesudah case folding

Sebelum Case Folding	Setelah Case Folding
MAAAAUUU	maaaauuu
mau dear... akuuuu Xtra Small 😻	mau dear akuuuu xtra small
Bajunya bisa buat hijab ga? Pake manset kulit gtu	bajunya bisa buat hijab ga pake manset kulit gtu

Proses selanjutnya adalah c*ase folding* yang ditunjukkan pada Tabel 3, di mana seluruh teks diubah ke dalam bentuk huruf kecil. Dengan penerapan *case folding*, perbedaan penulisan huruf kapital dapat dihilangkan sehingga

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

representasi kata menjadi lebih konsisten. Dengan penyeragaman ini, sistem dapat mengenali kata secara konsisten dan mengurangi redundansi fitur pada tahap pembobotan.

Dampak penghapusan kata umum (stopword removal) terhadap struktur kalimat disajikan pada Tabel 4.

Tabel 4. Sebelum dan sesudah stopword removal

Sebelum Stopword Removal	Setelah Stopword Removal
MAAAAUUU	maaaauuu
mau dear... akuuuu Xtra Small 😻	dear akuuuu xtra small
Bajunya bisa buat hijab ga? Pake manset kulit gtu	bajunya hijab ga pake manset kulit gtu

Hasil penerapan *stopword removal* yang ditampilkan pada Tabel 4 menunjukkan penghapusan kata-kata umum yang tidak memberikan kontribusi berarti dalam menentukan sentimen. Tahap ini membantu memperkecil kompleksitas data dengan mempertahankan kata-kata yang lebih bermakna secara emosional maupun kontekstual, sehingga informasi penting dalam komentar menjadi lebih menonjol.

Proses tokenisasi yang memecah teks menjadi unit kata ditampilkan pada Tabel 5.

Tabel 5. Sebelum dan sesudah tokenisasi

Sebelum Tokenisasi	Setelah Tokenisasi
MAAAAUUU	['maaaauuu']
mau dear... akuuuu Xtra Small 😻	['dear', 'akuuuu', 'xtra', 'small']
Bajunya bisa buat hijab ga? Pake manset kulit gtu	['bajunya', 'hijab', 'ga', 'pake', 'manset', 'kulit', 'gtu']

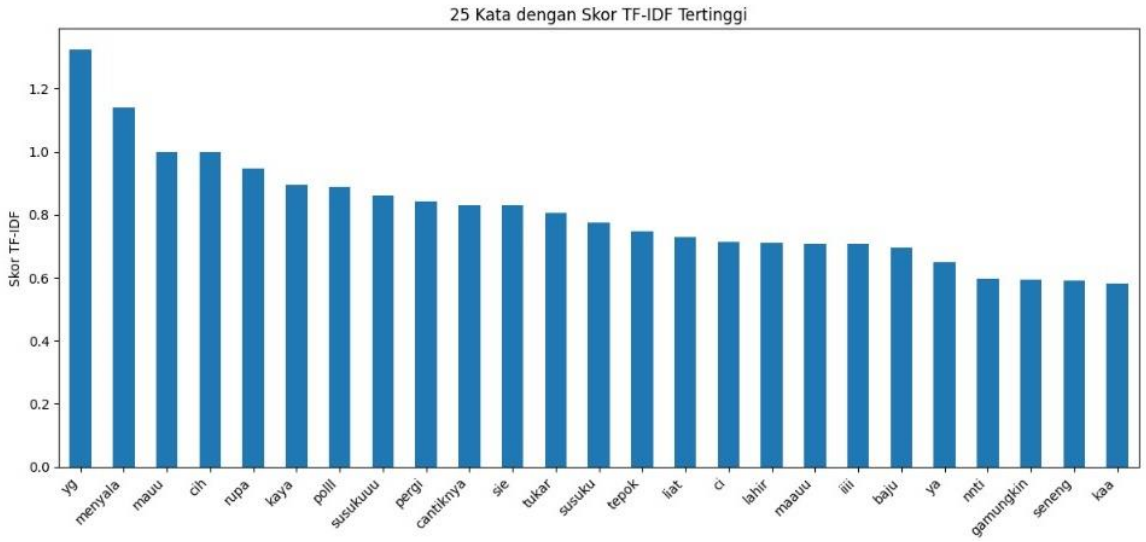
Tahap tokenisasi yang disajikan pada Tabel 5 memecah teks yang telah dinormalisasi menjadi kumpulan kata tunggal. Proses ini menjadi tahap penting karena mengubah teks mentah menjadi format yang dapat dihitung secara matematis. Token hasil pemrosesan ini kemudian digunakan sebagai dasar dalam pembentukan fitur menggunakan metode TF-IDF.

Secara keseluruhan, prapemrosesan yang dilakukan berfungsi sebagai fondasi dalam membangun kualitas dataset. Teks yang telah melalui proses pembersihan, normalisasi, dan tokenisasi menghasilkan representasi data yang lebih terstruktur dan relevan. Kondisi ini memungkinkan metode TF-IDF dan model *Artificial Neural Network* (ANN) bekerja secara lebih optimal dalam mempelajari pola bahasa pada komentar TikTok yang bersifat singkat, informal, dan bervariasi.

c. Hasil Pembobotan TF-IDF

Agar dapat diolah oleh model *Artificial Neural Network* (ANN), data teks komentar terlebih dahulu dikonversi ke dalam bentuk numerik menggunakan metode TF-IDF. Metode ini bekerja dengan menghitung tingkat kepentingan suatu kata berdasarkan frekuensi kemunculannya dalam satu komentar serta tingkat kelangkaannya di seluruh dataset. Melalui mekanisme ini, TF-IDF memberikan bobot yang lebih besar pada kata-kata yang bersifat spesifik dan informatif, sementara kata-kata yang sering muncul dan kurang diskriminatif memperoleh bobot yang lebih rendah.

Penerapan TF-IDF pada komentar TikTok sangat relevan karena karakteristik bahasa yang digunakan cenderung singkat, spontan, dan dipenuhi kata-kata informal atau slang. Melalui pembobotan ini, sistem dapat menangkap kata-kata yang mencerminkan ekspresi, emosi, maupun konteks tertentu dari pengguna, sehingga representasi teks menjadi lebih bermakna dibandingkan sekadar menghitung frekuensi kata secara sederhana. Visualisasi kata dengan skor TF-IDF tertinggi disajikan pada Gambar 3.



Gambar 3. Visualisasi Fitur TF-IDF Tertinggi

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Berdasarkan Gambar 3, terlihat bahwa sejumlah kata memiliki skor TF-IDF yang lebih tinggi dibandingkan kata lainnya. Kata-kata dengan bobot tertinggi tersebut mencerminkan pola bahasa yang sering muncul pada komentar TikTok, khususnya kata-kata spontan, ekspresif, dan bersifat kontekstual. Tingginya skor TF-IDF menunjukkan bahwa kata-kata tersebut tidak hanya sering muncul pada komentar tertentu, tetapi juga relatif jarang digunakan secara merata di seluruh dataset, sehingga memiliki daya pembeda yang kuat.

Visualisasi ini menunjukkan bahwa komentar pengguna TikTok terhadap konten *flexing* didominasi oleh kata-kata yang bersifat informal dan reaktif, seperti ekspresi spontan maupun istilah gaul. Hal ini mengindikasikan bahwa respons pengguna lebih banyak berupa reaksi langsung terhadap konten dibandingkan opini panjang yang terstruktur. Dari sisi pemodelan, fitur-fitur dengan bobot TF-IDF tinggi ini berperan penting dalam membantu model ANN mengenali pola sentimen, karena kata-kata tersebut menjadi sinyal utama dalam membedakan komentar positif, netral, maupun negatif.

Secara keseluruhan, hasil pembobotan TF-IDF membuktikan bahwa metode ini efektif dalam mengekstraksi kata-kata kunci yang relevan dari komentar TikTok. Dengan representasi numerik yang dihasilkan, model ANN mampu memanfaatkan informasi frekuensi dan konteks kata secara lebih optimal dalam mengklasifikasikan sentimen pada data teks yang informal dan dinamis.

d. Hasil Model ANN Sebelum SMOTE

Model *Artificial Neural Network* (ANN) pada tahap awal dilatih dengan data asli, tanpa penerapan metode penanganan ketidakseimbangan kelas. Evaluasi dilakukan pada data uji untuk memperoleh gambaran awal mengenai kemampuan model dalam mengklasifikasikan sentimen komentar sebelum diterapkan teknik penyeimbangan data.

Model *Artificial Neural Network* (ANN) diimplementasikan menggunakan *algoritma Multi-Layer Perceptron (MLPClassifier)* dari *Scikit-learn*. Arsitektur jaringan terdiri dari dua hidden layer dengan 128 dan 64 neuron serta menggunakan fungsi aktivasi ReLU. Proses optimasi dilakukan dengan algoritma Adam, dan model dilatih hingga maksimal 120 iterasi. Parameter *random_state* sebesar 42 digunakan untuk memastikan konsistensi hasil eksperimen.

Hasil evaluasi model ANN sebelum SMOTE ditunjukkan pada Tabel 6, yang memuat nilai *precision*, *recall*, F1-*score*, serta jumlah data *(support)* pada masing-masing kelas sentimen. Berdasarkan hasil pengujian, model ANN sebelum SMOTE menghasilkan akurasi sebesar 89,79%, yang menunjukkan bahwa model telah mampu melakukan klasifikasi sentimen komentar dengan cukup baik pada data uji.

Tabel 6. Hasil Model ANN sebelum SMOTE

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,88	0,41	0,56	34
Netral	0,90	0,97	0,93	558
Positif	0,89	0,79	0,84	211
Accuracy			0,90	803
Macro Average	0,89	0,72	0,78	803
Weighted Average	0,90	0,90	0,89	803

Berdasarkan Tabel 6, model ANN menunjukkan performa yang sangat baik pada kelas netral dan positif. Namun demikian, performa model pada kelas negatif masih terbatas terutama pada nilai *recall* yang mengindikasikan bahwa banyak komentar negatif belum berhasil terklasifikasi secara tepat. Hal ini disebabkan oleh ketidakseimbangan jumlah data, di mana kelas negatif memiliki proporsi yang jauh lebih kecil dibandingkan kelas lainnya.

e. Hasil Model ANN Setelah SMOTE

Untuk menangani permasalahan ketidakseimbangan kelas, metode *Synthetic Minority Over-sampling Technique* (SMOTE) diterapkan pada data latih. Penerapan metode ini bertujuan untuk meningkatkan proporsi kelas minoritas, khususnya kelas negatif, sehingga distribusi data menjadi lebih seimbang sebelum proses pelatihan model dilakukan kembali.

Hasil evaluasi model ANN setelah penerapan SMOTE disajikan pada Tabel 7. Berdasarkan hasil pengujian, model ANN setelah penerapan SMOTE memperoleh akurasi sebesar 90,04%, yang menunjukkan adanya peningkatan kinerja model dibandingkan dengan model tanpa penyeimbangan data.

Tabel 7. Model ANN setelah SMOTE

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,79	0,44	0,57	34
Netral	0,90	0,97	0,94	558
Positif	0,89	0,79	0,84	211
Accuracy			0,90	803
Macro Average	0,86	0,73	0,78	803
Weighted Average	0,90	0,90	0,89	803

Berdasarkan Tabel 7, penerapan SMOTE menunjukkan adanya peningkatan pada performa klasifikasi, khususnya pada metrik *recall* dan F1-*score* untuk kelas negatif. Hal ini mengindikasikan bahwa teknik *oversampling* berhasil

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9494
Hal 367-375
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

membantu model dalam mengenali pola pada kelas minoritas secara lebih efektif dibandingkan sebelum dilakukan penyeimbangan data.

f. Evaluasi Performa Model

Perbandingan kinerja model ANN sebelum dan sesudah penerapan SMOTE dilakukan untuk menilai dampak metode tersebut terhadap hasil klasifikasi. Ringkasan perbandingan tersebut disajikan pada Tabel 8.

Tabel 8. Perbandingan Performa Model ANN Sebelum dan Sesudah SMOTE

Model	Akurasi	Precision Negatif	Recall Negatif	F1 Negatif	F1 Weighted
ANN sebelum SMOTE	89.79%	0.88	0.41	0.56	0.89
ANN ssetelah SMOTE	90.04%	0.79	0.44	0.57	0.90

Berdasarkan Tabel 8, penerapan SMOTE menghasilkan peningkatan performa model ANN, terutama pada kelas negatif yang sebelumnya sulit dikenali. Peningkatan terlihat pada nilai recall dan F1-score kelas negatif, meskipun dalam skala yang relatif kecil.

Hasil ini menunjukkan bahwa ketidakseimbangan distribusi data berpengaruh terhadap kemampuan model dalam mendeteksi komentar negatif. Selain itu, karakteristik bahasa komentar TikTok yang singkat, informal, dan kontekstual juga menjadi tantangan bagi model berbasis TF-IDF. Secara keseluruhan, ANN menunjukkan performa yang stabil, sementara penerapan SMOTE membantu meningkatkan sensitivitas model terhadap kelas minoritas

3.2 Pembahasan

Hasil penelitian menunjukkan bahwa sentimen netral mendominasi komentar pengguna TikTok pada seluruh konten yang dianalisis. Dari total 4.013 komentar, sebanyak 2.788 komentar (69,47%) bersentimen netral, 1.056 komentar (26,31%) bersentimen positif, dan hanya 169 komentar (4,21%) bersentimen negatif. Dominasi sentimen netral ini konsisten pada setiap video yang diambil sebagai data penelitian.

Pada video tanggal 13 April 2025 (*Get Ready With Me* menggunakan kemeja) dan 25 April 2025 (menampilkan ekspresi senyum kreator), komentar didominasi oleh respons netral berupa interaksi ringan dan pengamatan singkat. Video tanggal 27 April 2025 yang mengangkat tren kesenjangan sosial dengan menampilkan ruang karaoke pribadi juga menunjukkan pola serupa, di mana mayoritas komentar tetap bersentimen netral meskipun topik berpotensi memicu respons emosional.

Konten tanggal 8 Mei 2025 yang berisi rekomendasi produk lipstik serta video tanggal 22 Mei 2025 yang menampilkan kreator sebagai model memperlihatkan peningkatan komentar positif, namun tetap tidak menggeser dominasi komentar netral. Komentar negatif pada kedua konten ini muncul dalam proporsi yang sangat kecil.

Selanjutnya, pada rangkaian video *Get Ready With Me* dengan outfit mahal dan unik yang diunggah pada 28 Juni 2025, 10 Juli 2025, 29 Juli 2025, 11 Agustus 2025, dan 9 September 2025, komentar netral secara konsisten menjadi sentimen dominan. Meskipun konten menampilkan unsur *flexing* secara eksplisit, sebagian besar pengguna merespons secara pasif melalui candaan ringan, observasi singkat, atau interaksi sosial, sementara komentar negatif tetap minimal.

Temuan ini menunjukkan bahwa fenomena *flexing* di Indonesia cenderung dipersepsikan sebagai hiburan visual, bukan sebagai pemicu respons emosional yang kuat, baik dalam bentuk dukungan berlebihan maupun *cyberbullying*. Kondisi ini berdampak pada performa model *Artificial Neural Network* (ANN), di mana dominasi kelas netral menyebabkan model lebih optimal mengenali pola netral dibandingkan kelas negatif. Hal ini tercermin dari nilai F1-score yang tinggi pada kelas netral dan rendahnya recall pada kelas negatif.

Penerapan *Synthetic Minority Over-sampling Technique* (SMOTE) meningkatkan sensitivitas model terhadap kelas negatif, ditunjukkan oleh peningkatan recall dari 0,41 menjadi 0,44. Walaupun peningkatan performa relatif kecil, penyeimbangan data tetap memberikan dampak positif terhadap kemampuan model dalam mengidentifikasi kelas minoritas.

4. KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna TikTok terhadap fenomena *flexing* dengan menerapkan metode *Artificial Neural Network* (ANN) dan pembobotan TF-IDF. Hasil implementasi dan pengujian menunjukkan bahwa fenomena *flexing* di media sosial tidak menimbulkan respons emosional yang kuat dari pengguna, yang tercermin dari dominasi komentar dengan sentimen netral. Kondisi ini menunjukkan bahwa masyarakat Indonesia cenderung menganggap *flexing* sebagai hal yang biasa sehingga tidak menimbulkan reaksi positif atau negatif yang signifikan. Model ANN yang digunakan mampu mengelola data teks dengan baik, terutama dalam mengenali komentar netral dan positif, yang ditunjukkan dengan nilai F1-*score* yang tinggi pada kedua kelas tersebut. Namun, performa model pada kelas negatif masih rendah karena jumlah data minoritas yang sangat sedikit dan karakteristik komentar negatif yang cenderung bersifat ambigu, sarkastik, atau memiliki konteks yang sulit ditangkap oleh model berbasis TF-IDF. Penerapan SMOTE berhasil meningkatkan performa pada kelas minoritas meskipun peningkatannya tidak terlalu besar. Secara keseluruhan, metode ANN dengan TF-IDF cukup efektif untuk analisis sentimen komentar TikTok, tetapi penelitian ini memiliki keterbatasan pada ketidakseimbangan data dan kompleksitas bahasa pada komentar negatif. Penelitian

selanjutnya dapat mempertimbangkan penambahan sampel, pendekatan embedding berbasis deep learning, serta algoritma klasifikasi yang lebih adaptif untuk meningkatkan performa model pada kelas minoritas.

REFERENCES

[1] A. F. Maharani and C. Gusnita, "Analisis Cyberbullying : Komentar Kebencian Terhadap Pembuat Konten Beauty Influencer di Media Sosial Tiktok," vol. 6, no. 4, pp. 519–527, 2024.
[2] A. H. Firstiyanti, M. Wijaya, S. Subiyantoro, and S. Kusumo, *Sentiment Analysis of Comments on Flexing Content Posted by Gen Z Celebrities : Does Gender Matter ?*, no. ICCoLliC. Atlantis Press SARL, 2024. doi: 10.2991/978-2-38476-321-4.
[3] R. Pakpahan and D. Yoesgiantoro, "ANALYSIS OF THE INFLUENCE OF FLEXING IN Analisa Pengaruh Flexing Di Media Sosial Terhadap Kehidupan Masyarakat," vol. 7, no. 1, pp. 173–178, 2023, doi: 10.52362/jisicom.v7i1.1093.
[4] T. Cam, T. Dinh, and Y. Lee, "Heliyon Social media influencers and followers ' conspicuous consumption : The mediation of fear of missing out and materialism," *Heliyon*, vol. 10, no. 16, p. e36387, 2024, doi: 10.1016/j.heliyon.2024.e36387.
[5] O. O. Amusan and A. M. Udefi, "AI-powered sentiment analysis for classifying harmful content on social media : A case study with ChatGPT Integration," 2024.
[6] M. Agbaje and O. Afolabi, "Revue d ' Intelligence Artificielle Neural Network-Based Cyber-Bullying and Cyber-Aggression Detection Using Twitter (X) Text," vol. 38, no. 3, pp. 837–846, 2024.
[7] A. Andira *et al.*, "Analisis Sentimen Cyberbullying pada 'Tiktok' Menggunakan Metode Long Short Term Memory," vol. XVI, no. 1, pp. 1–10, 2023.
[8] G. Ramos *et al.*, "A comprehensive review on automatic hate speech detection in the age of the transformer," *Soc. Netw. Anal. Min.*, vol. 14, no. 1, pp. 1–25, 2024, doi: 10.1007/s13278-024-01361-3.
[9] N. Bayes, C. To, and M. Public, "Analisis Sentimen Komentar Tiktok Tentang Kasus TPPO TKI Di Kamboja Menggunakan Klasifikasi Naive Bayes Untuk Mengukur Opini Publik," vol. 24, no. 2, pp. 602–616, 2025.
[10] S. Hinduja and J. W. Patchin, "HARASSMENT IN TIKTOK COMMENTS A Pilot Test of the TikTok Research API," 2023.
[11] L. Sandra, "Lockdown Countdown : Lockdown Sentiment Analysis on Twitter Using Artificial Neural Network," pp. 198–202, 2022.
[12] S. Yulian, F. Noorihsan, N. Y. Setiawan, and M. Chandra, "Analisis Sentimen Ulasan Google Review New Star Cineplex Pasuruan menggunakan Artificial Neural Network (ANN)," vol. 7, no. 2, 2023.
[13] V. No, L. H. Sarumpaet, and R. R. Suryono, "Edumatic : Jurnal Pendidikan Informatika Analisis Sentimen Publik Program PPPK di Media Sosial X menggunakan Naïve Bayes dan SVM," vol. 9, no. 2, pp. 362–371, 2025, doi: 10.29408/edumatic.v9i2.30065.
[14] V. No, K. Miranda, and R. R. Suryono, "Edumatic : Jurnal Pendidikan Informatika Analisis Sentimen Pinjaman Online : Studi Komparatif Algoritma Naïve Bayes , Decision Tree , dan KNN," vol. 9, no. 2, pp. 372–381, 2025, doi: 10.29408/edumatic.v9i2.30142.
[15] A. H. Pratama, M. Hayaty, U. A. Yogyakarta, and R. R. Utara, "Performance of Lexical Resource and Manual Labeling on Long Short-Term Memory Model for Text Classification," vol. 9, no. 1, pp. 74–84, 2023, doi: 10.26555/jiteki.v9i1.25375.
[16] A. G. Philipo, "Cyberbullying Detection : Exploring Datasets , Technologies , and Approaches on Social Media Platforms".
[17] S. Dharmawan, V. Christanti, M. Novario, and J. Perdana, "Klasifikasi Ujaran Kebencian Menggunakan Metode FeedForward Neural Network (IndoBERT)," pp. 1–6.
[18] S. Sifath, T. Islam, S. Kumar, and M. Ul, "Recurrent neural network based multiclass cyber bullying classification," *Nat. Lang. Process. J.*, vol. 9, no. July, p. 100111, 2024, doi: 10.1016/j.nlp.2024.100111.
[19] M. A. Jabbar, *Proceeding of 2021 International Conference on Wireless Communications , Networking and Applications*. 2021.
[20] A. C. Roy and T. Mahmud, "A multi-class cyberbullying classification on image and text in code-mixed Bangla-English social media content," *Nat. Lang. Process. J.*, vol. 13, no. August, p. 100191, 2025, doi: 10.1016/j.nlp.2025.100191.
[21] B. A. Prameswari, H. S. Oktaviani, and ..., "Building prediction model for detecting cyberbullying using tiktok comments," *2023 IEEE 8th …*, 2023.
[22] A. Wirawan, H. D. Cahyono, and Winarno, "Easy Data Augmentation in Sentiment Analysis of Cyberbullying," *2023 6th Int. Conf. Inf. Commun. Technol. ICOIACT 2023*, pp. 443–447, 2023, doi: 10.1109/ICOIACT59844.2023.10455817.
[23] P. Pujari, "Sentiment Analysis of Cyberbullying Data in Social Media," no. Ml.
[24] D. Rifaldi, Abdul Fadlil, and Herman, "Teknik Preprocessing Pada Text Mining," *J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, pp. 161–171, 2023.
[25] A. C. N. Rosyadi, K. R. Leonida, M. Athoillah, and H. B. Rohmanto, "Analisis Sentimen di Twitter terhadap Kenaikan Uang Kuliah Tunggal Menggunakan Artificial Neural Network (ANN)," 2024, *Universitas PGRI Adi Buana Surabaya*.
[26] V. Westley, D. Thomas, and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," vol. 6, pp. 1767–1774, 2022, doi: 10.30865/mib.v6i3.4218.
[27] I. Najiyah, "Analisis Sentimen Tanggapan Masyarakat Indonesia Tentang Kenaikan BBM Menggunakan Metode Artificial Neural Network," *J. Responsif Ris. Sains Dan Inform.*, 2023.
[28] A. Arasy, S. Agustian, L. Handayani, and I. Iskandar, "Klasifikasi Sentimen Menggunakan Metode Multilayer Perceptron dengan Fitur TF-IDF," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3, pp. 908–919, 2025, doi: 10.57152/malcom.v5i3.2052.
[29] A. Bijaksana and P. Negara, "The Influence Of Applying Stopword Removal And Smote On Indonesian Sentiment Classification," vol. 14, no. 3, pp. 172–185, 2023.
[30] K. W. Chandra and H. Irsyad, "Efektifitas SMOTE dalam Mengatasi Imbalanced Class Algoritma K-Nearest Neighbors pada Analisis Sentimen terhadap Starlink," vol. 4, no. 1, pp. 31–42, 2024.
[31] R. E. Pambudi, H. Purnomo, and A. Aglasia, "Analisis Klasifikasi Sentimen Pengguna MyPertamina Menggunakan Metode Evaluasi Precision , Recall , dan," vol. 07, no. 02, pp. 17–22, 2025.