

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI TIKTOK PADA GOOGLE PLAY STORE MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK

SENTIMENT ANALYSIS OF TIKTOK APPLICATION USER REVIEWS ON THE GOOGLE PLAY STORE USING CONVOLUTIONAL NEURAL NETWORK

Samuel Arie Rasmadi¹, Heribertus Yulianton²

Universitas Stikubank Semarang^{1,2}

samuelarie24@gmail.com¹

ABSTRAK

TikTok is a short-video-based social media platform with extremely rapid user growth, generating millions of reviews on the Google Play Store that contain valuable information regarding user experiences. The massive volume of this review data makes manual analysis inefficient, thus requiring an automated computational approach. This study aims to implement the Convolutional Neural Network (CNN) method to classify the sentiments of TikTok user reviews into positive, negative, and neutral categories. The main challenge in this study is the characteristics of the review texts, which frequently use mixed languages (code-mixing). The CNN model is implemented by utilizing word embedding techniques to extract in-depth textual features. The testing results show that the CNN model successfully achieves a very high accuracy rate of 94.97% with a loss value of 0.3440. The model's performance in the positive and negative classes is highly robust, with F1-scores of 0.97 and 0.95 respectively, while the neutral class scored 0.84. This research proves that the CNN architecture is highly effective and robust in recognizing complex user opinion patterns, thereby providing accurate recommendations for application developers to automatically understand public preferences.

Keywords: Sentiment Analysis, Convolutional Neural Network, Google Play Store, Natural Language Processing, TikTok.

ABSTRAK

TikTok merupakan platform media sosial berbasis video pendek dengan pertumbuhan pengguna yang sangat pesat, menghasilkan jutaan ulasan di Google Play Store yang mengandung informasi berharga mengenai pengalaman pengguna. Besarnya volume data ulasan tersebut menyebabkan analisis manual menjadi tidak efisien, sehingga diperlukan pendekatan komputasional otomatis. Penelitian ini bertujuan untuk mengimplementasikan metode Convolutional Neural Network (CNN) dalam mengklasifikasikan sentimen ulasan pengguna TikTok ke dalam kategori positif, negatif, dan netral. Tantangan utama dalam penelitian ini adalah karakteristik teks ulasan yang sering menggunakan bahasa campuran (code-mixing). Model CNN diimplementasikan dengan memanfaatkan teknik word embedding untuk mengekstraksi fitur-fitur tekstual secara mendalam. Hasil pengujian menunjukkan bahwa model CNN berhasil mencapai tingkat akurasi yang sangat tinggi sebesar 94,97% dengan nilai loss 0,3440. Performa model pada kelas positif dan negatif sangat kuat dengan F1-score masing-masing 0,97 dan 0,95, sementara pada kelas netral sebesar 0,84. Penelitian ini membuktikan bahwa arsitektur CNN sangat efektif dan tangguh dalam mengenali pola opini pengguna yang kompleks, sehingga dapat memberikan rekomendasi yang akurat bagi pengembang aplikasi dalam memahami preferensi publik secara otomatis.

Kata Kunci: Analisis Sentimen, Convolutional Neural Network, Google Play Store, Natural Language Processing, TikTok.

PENDAHULUAN

Kemajuan pesat teknologi informasi telah memicu transformasi dramatis dalam berbagai aspek kehidupan manusia. Salah satu ciri menonjol dari era digital kontemporer adalah prevalensi platform media sosial berbasis video pendek, yang dicontohkan oleh TikTok, yang telah menjadi bagian integral dari kehidupan sehari-hari di seluruh dunia. Aplikasi Tik

Tok adalah sebuah jejaring sosial dan platform video musik asal negeri Tiongkok yang diluncurkan pada awal September 2016 (Bersastra, 2020), platform ini telah meningkatkan kreativitas pengguna dan mengalami peningkatan popularitas yang signifikan, terutama selama pandemi (Husain et al., 2024). TikTok diproyeksikan akan melampaui 1,9 miliar pengguna aktif secara global pada tahun 2025, menjadikan

Indonesia sebagai pasar yang signifikan, sehingga secara langsung memengaruhi dinamika koneksi digital. Interaksi ini salah satunya terekam dengan jelas melalui pemberian rating dan ulasan pengguna pada platform distribusi aplikasi Google Play Store.

Hingga tahun 2025, TikTok tercatat telah diunduh lebih dari 1 miliar kali dengan akumulasi lebih dari 21 juta ulasan di Google Play Store. Platform ini telah berevolusi dari sekadar sarana hiburan menjadi instrumen edukasi dan pemasaran yang bernilai ekonomi tinggi. Ulasan yang ditinggalkan oleh pengguna pada dasarnya mampu memberikan wawasan yang bermanfaat serta dapat dipercaya terkait pengalaman penggunaan aplikasi (Rizaldi et al., 2023). Walaupun demikian, tidak sedikit pengguna yang menyampaikan sentimen negatif mengenai masalah privasi, perubahan algoritma, hingga kendala teknis aplikasi. Karena volume ulasan yang terus meningkat, yang mencapai jutaan titik data setiap bulannya, proses evaluasi manual menjadi sangat tidak efisien dan rentan terhadap bias subjektif. Akibatnya, metodologi komputasi sangat penting untuk ekstraksi dan analisis opini publik yang tepat.

Metode yang sangat baik untuk mengatasi masalah data teks yang luas ini adalah implementasi analisis sentimen. Sebagai bagian dari cabang Natural Language Processing (NLP), analisis sentimen merupakan sekumpulan metode, alat, maupun pendekatan yang diimplementasikan guna menemukan serta mengekstrak informasi subjektif, termasuk opini dari suatu dokumen teks (Humam, 2023). Penelitian ini memungkinkan algoritma untuk secara otomatis mengkategorikan opini pengguna di aplikasi TikTok ke dalam klasifikasi baik, negatif, atau netral. Otomatisasi ini memungkinkan pemrosesan data menjadi lebih terstruktur, sehingga pola preferensi pengguna dapat diidentifikasi dengan cepat.

Meskipun analisis sentimen menawarkan keunggulan dalam menangani big data, terdapat tantangan utama dalam klasifikasi sentimen pada konteks pengguna di Indonesia, yaitu karakteristik teks ulasan yang kerap kali menggunakan bahasa campuran (code-mixing), seperti paduan antara bahasa Indonesia dan Inggris. Kondisi ini menuntut penggunaan model machine learning yang mutakhir untuk mempertahankan tingkat akurasi. Untuk menjawab tantangan tersebut, penelitian ini menggunakan Convolutional Neural Network (CNN). Convolutional Neural Network (CNN) adalah metode pembelajaran mendalam yang sering digunakan untuk klasifikasi gambar, mampu mengenali wajah, rambu lalu lintas, tumor, dan berbagai elemen data visual lainnya (Pangaribuan et al., 2021). Awalnya dirancang untuk pemrosesan gambar digital, arsitektur CNN baru-baru ini telah diadaptasi secara efektif untuk mengatasi tantangan NLP. Dengan memanfaatkan teknik representasi word embedding dan ekstraksi fitur melalui convolutional layers, model CNN diharapkan mampu mengenali pola kalimat yang kompleks dan meningkatkan akurasi klasifikasi sentimen pada ulasan aplikasi TikTok secara signifikan.

METODE

Tempat dan Waktu

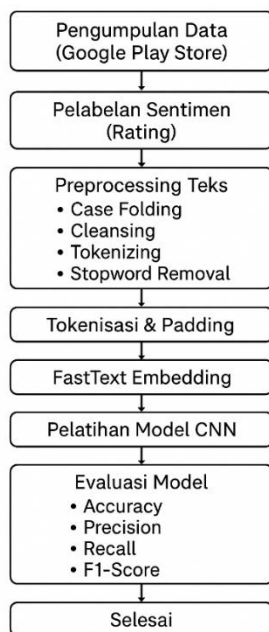
Penelitian dilakukan secara daring (online). Pengambilan data dilakukan pada bulan September 2025 dari Google Play Store. Implementasi dan pelatihan model dilaksanakan menggunakan Google Colaboratory.

Alat dan Bahan

Perangkat keras: laptop (Asus X441N, RAM 4 GB). Perangkat lunak dan pustaka utama: Python, Google Colab, Google Drive, TensorFlow/Keras, Scikit-learn, Gensim (fastText), NLTK, Pandas, dan utilitas google_play_scraper. Data: kumpulan ulasan TikTok berbahasa Indonesia sebanyak 10.000 record.

Prosedur Penelitian

Prosedur penelitian meliputi: (1) pengumpulan ulasan dari Google Play Store melalui web scraping; (2) pelabelan awal berdasarkan skor rating; (3) pra-pemrosesan teks; (4) transformasi teks menjadi representasi numerik menggunakan fastText; (5) pembangunan dan pelatihan model CNN; dan (6) evaluasi kinerja model klasifikasi. Prosedur penelitian meliputi pengumpulan data, pelabelan sentimen berdasarkan rating, pra-pemrosesan teks, transformasi data menjadi representasi numerik, pelatihan model CNN, dan evaluasi kinerja model. Alur keseluruhan tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Diagram alur metodologi penelitian

Teknik Pengumpulan Data

Data dikumpulkan menggunakan pustaka `google_play_scraper` dengan parameter `lang='id'` dan `country='id'`. Pengambilan dilakukan bertahap (10 batch $\times$ 1.000 ulasan) untuk menghindari *rate limiting*. Hanya atribut content (teks) dan score (rating) yang digunakan. Data mentah disimpan dalam format CSV dan dianonimkan sebelum diproses.

Teknik Analisis Data

Pra-pemrosesan

Pra-pemrosesan meliputi case folding, pembersihan (menghapus URL, mention, angka, tanda baca dan karakter khusus), tokenisasi (ToksTokTokenizer), dan penghapusan stopword bahasa Indonesia dengan pengecualian kata negasi (mis. “tidak”, “bukan”). Stemming dan normalisasi slang **tidak** dilakukan untuk mempertahankan bentuk kata asli sehingga fastText dapat memanfaatkan informasi subword.

Tokenisasi dan Padding

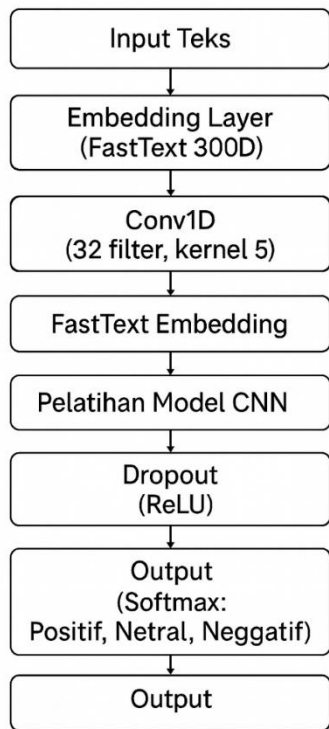
Teks diubah menjadi urutan indeks menggunakan Tokenizer Keras dengan batas kosakata `num_words=5000`. Panjang urutan diseragamkan dengan padding hingga `MAX_LEN = 100` kata.

Representasi Kata (fastText)

Embedding menggunakan pretrained FastText bahasa Indonesia (cc.id.300.vec, dimensi 300). Matriks embedding dibuat sesuai kosakata; bila kata tidak ditemukan (OOV) diisi nol dan disesuaikan selama pelatihan karena lapisan embedding dikonfigurasi `trainable=True` untuk memungkinkan *fine-tuning* pada domain ulasan TikTok.

Arsitektur Model CNN

Model CNN berisi lapisan Embedding $\rightarrow$ Conv1D (32 filter, `kernel_size=5`) $\rightarrow$ GlobalMaxPooling $\rightarrow$ Dropout (`rate=0.5`) $\rightarrow$ Dense (16, ReLU) $\rightarrow$ Output Dense (3, Softmax). Loss function menggunakan `categorical_crossentropy` dan optimizer Adam (`lr=0.001`). Model yang digunakan adalah Convolutional Neural Network (CNN) yang terdiri dari lapisan embedding, convolutional 1D, global max pooling, dropout, dan dense layer. Struktur arsitektur model CNN yang digunakan dalam penelitian ini ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur model cnn

Teknik Pengujian dan Evaluasi

Dataset dibagi 70% training dan 30% testing; dari data latih digunakan 20% sebagai validation split selama training. Pelatihan dilakukan sampai maksimal 18 epoch dengan batch_size=32. Untuk menghindari overfitting diterapkan EarlyStopping (monitor val_loss) dan ReduceLROnPlateau. Karena distribusi kelas tidak seimbang, digunakan *class weighting* (menghitung bobot dengan `compute_class_weight` skema balanced) saat `model.fit()` untuk memberi penalti lebih besar pada kelas minoritas. Evaluasi dilakukan menggunakan metrik akurasi, precision, recall, F1-score per kelas, serta confusion matrix. Model dan artefak (tokenizer, embedding matrix, model .h5) disimpan untuk reproduktibilitas.

HASIL DAN PEMBAHASAN

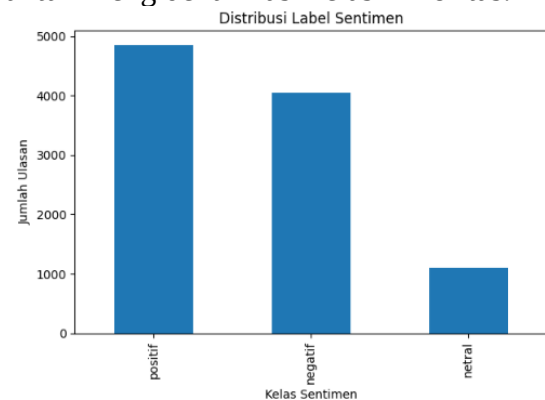
Hasil Pengumpulan dan Distribusi Data

Studi ini menggunakan dataset yang mencakup 10.000 peringkat pengguna aplikasi TikTok, yang dikumpulkan dari Google Play Store pada September 2025. Klasifikasi sentimen dilakukan secara otomatis berdasarkan skor peringkat

pengguna: rating 4–5 menunjukkan sentimen positif, rating 3 menandakan sentimen netral, dan rating 1–2 menunjukkan sentimen negatif.

Hasil distribusi data menunjukkan bahwa **kelas sentimen positif mendominasi**, diikuti oleh sentimen negatif dan netral. Dominansi ulasan positif mengindikasikan bahwa secara umum pengguna memberikan penilaian yang baik terhadap aplikasi TikTok. Namun demikian, jumlah ulasan negatif yang masih cukup signifikan menunjukkan adanya permasalahan yang dialami oleh sebagian pengguna dan perlu mendapatkan perhatian lebih lanjut.

Distribusi kelas sentimen yang tidak seimbang ini berpotensi memengaruhi kinerja model klasifikasi. Oleh karena itu, selama fase pelatihan model, strategi pembobotan kelas diterapkan untuk mengurangi bias terhadap kelas mayoritas dan meningkatkan kemampuan model untuk mengidentifikasi kelas minoritas.



Gambar 3. Diagram batang distribusi sentimen

Hasil Pra-pemrosesan dan Representasi Data Teks

Hasil Pra-pemrosesan Teks

Fase pra-pemrosesan dilakukan untuk meningkatkan kualitas data teks sebelum diproses oleh model Convolutional Neural Network (CNN). Pra-pemrosesan dilakukan secara bertahap dan berurutan, di mana hasil dari satu tahap digunakan sebagai masukan untuk tahap berikutnya.

Studi ini menggunakan proses pra-pemrosesan seperti pengubahan huruf

besar-kecil, pembersihan data, tokenisasi, dan penghapusan kata-kata penghenti (stopwords). Pada proses penghapusan stopword, kata negasi seperti “*tidak*”, “*tak*”, dan “*bukan*” tetap dipertahankan karena memiliki peran penting dalam menentukan polaritas sentimen. Selain itu, ditambahkan stopword khusus yang sering muncul pada ulasan informal pengguna.

Penelitian ini tidak menerapkan stemming maupun normalisasi kata tidak baku. Keputusan ini diambil karena penggunaan pre-trained fastText berbasis subword memungkinkan model tetap menangkap informasi semantik dari variasi morfologi dan ejaan kata yang umum ditemukan pada ulasan pengguna.

No	Original (content)	Case Folding
1	saya baru update software android 16 pada perangkat galaxy A53 5G akan tetapi tik tok menjadi mono. tolong software tik tok segera diupdate agar support ke android 16 sehingga saat live suara di perangkat saya jadi stereo lagi.terimakasih🙏	saya baru update software android 16 pada perangkat galaxy a53 5g akan tetapi tik tok menjadi mono. tolong software tik tok segera diupdate agar support ke android 16 sehingga saat live suara di perangkat saya jadi stereo lagi.terimakasih🙏
2	memberikan visibilitas yang tidak jelas,, tiba tiba keluar notif visibilitas. padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas,, makin kesini apk tiktok makin tidak baik baik saja	memberikan visibilitas yang tidak jelas,, tiba tiba keluar notif visibilitas. padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas,, makin kesini apk tiktok makin tidak baik baik saja
3	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang 3 bisa klaim hadiah lagi apa gak makasih🙏	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang 3 bisa klaim hadiah lagi apa gak makasih🙏

Gambar 4. Preprocessing case folding

No	Original (content)	Cleansed
1	saya baru update software android 16 pada perangkat galaxy A53 5G akan tetapi tik tok menjadi mono. tolong software tik tok segera diupdate agar support ke android 16 sehingga saat live suara di perangkat saya jadi stereo lagi.terimakasih🙏	saya baru update software android pada perangkat galaxy a g akan tetapi tik tok menjadi mono tolong software tik tok segera diupdate agar support ke android sehingga saat live suara di perangkat saya jadi stereo lagi terimakasih
2	memberikan visibilitas yang tidak jelas,, tiba tiba keluar notif visibilitas. padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas,, makin kesini apk tiktok makin tidak baik baik saja	memberikan visibilitas yang tidak jelas tiba tiba keluar notif visibilitas padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas makin kesini apk tiktok makin tidak baik baik saja
3	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang 3 bisa klaim hadiah lagi apa gak makasih🙏	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang bisa klaim hadiah lagi apa gak makasih

Gambar 5. Preprocessing cleansed

No	Original (content)	Tokenized
1	saya baru update software android 16 pada perangkat galaxy A53 5G akan tetapi tik tok menjadi mono. tolong software tik tok segera diupdate agar support ke android 16 sehingga saat live suara di perangkat saya jadi stereo lagi.terimakasih🙏	[saya, baru, update, software, android, pada, perangkat, galaxy, a, g, akan, tetapi, tik, tok, menjadi, mono, tolong, software, tik, tok, segera, diupdate, agar, support, ke, android, sehingga, saat, live, suara, di, perangkat, saya, jadi, stereo, lagi, terimakasih]
2	memberikan visibilitas yang tidak jelas,, tiba tiba keluar notif visibilitas. padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas,, makin kesini apk tiktok makin tidak baik baik saja	[memberikan, visibilitas, yang, tidak, jelas, tiba, tiba, keluar, notif, visibilitas, padahal, tidak, melakukan, apa, apa, kena, pelanggaran, yang, tidak, jelas, makin, kesini, apk, tiktok, makin, tidak, baik, baik, saja]
3	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang 3 bisa klaim hadiah lagi apa gak makasih🙏	[saya, suka, main, tiktok, dulunya, bisa, ambil, hadiahnya, sekarang, malah, gak, bisa, karena, akun, saya, dibatasi, saya, udah, coba, pake, cara, apa, pun, ttp, aja, gak, bisa, coba, saya, akan, beri, bintang, bisa, klaim, hadiah, lagi, apa, gak, makasih]

Gambar 6. Preprocessing tokenized

No	Original (content)	Cleaned Text
1	saya baru update software android 16 pada perangkat galaxy A53 5G akan tetapi tik tok menjadi mono. tolong software tik tok segera diupdate agar support ke android 16 sehingga saat live suara di perangkat saya jadi stereo lagi.terimakasih🙏	update software android perangkat galaxy a g tik tok mono tolong software tik tok diupdate support android live suara perangkat stereo terimakasih
2	memberikan visibilitas yang tidak jelas,, tiba tiba keluar notif visibilitas. padahal tidak melakukan apa apa kena pelanggaran yang tidak jelas,, makin kesini apk tiktok makin tidak baik baik saja	visibilitas tidak notif visibilitas tidak kena pelanggaran tidak kesini apk tiktok tidak
3	saya suka main tiktok dulunya bisa ambil hadiahnya sekarang malah gak bisa karena akun saya dibatasi saya udah coba pake cara apa pun ttp aja gak bisa coba saya akan beri bintang 3 bisa klaim hadiah lagi apa gak makasih🙏	suka main tiktok dulunya ambil hadiahnya gak akun dibatasi udah coba pake ttp aja gak coba bintang klaim hadiah gak makasih
4	Saat Buat Postingan Dalam Bentuk Teks Atau Tulisan, Jangan Terlalu Sedikit Batas Huruf Nya. Buat Jadi Minimal 2500 & Maksimal 5000 Huruf Begitu Begitu. Saya Jadi	postingan bentuk teks tulisan batas huruf minimal maksimal huruf susah

Gambar 7. Preprocessing cleansed text

Representasi Data Teks

Setelah tahap pra-pemrosesan, teks ulasan diubah menjadi format numerik dengan Tokeniser dari Keras, dengan batasan kosakata sebanyak 5.000 kata. Panjang setiap ulasan diseragamkan melalui proses padding hingga maksimum 100 token.

Representasi kata menggunakan pre-trained FastText bahasa Indonesia dengan dimensi vektor 300. Lapisan embedding diatur sebagai trainable, sehingga bobot embedding dapat disesuaikan selama proses pelatihan agar lebih representatif terhadap karakteristik domain ulasan aplikasi TikTok.

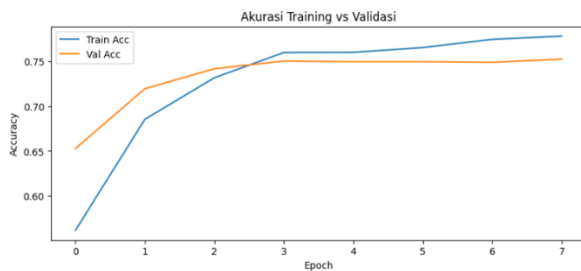
Hasil Pelatihan dan Evaluasi Model CNN

Evaluasi Proses Pelatihan Model (Training dan Validation)

Proses pelatihan model Convolutional Neural Network (CNN) dievaluasi dengan memantau pergerakan nilai metrik akurasi dan loss pada data training maupun data validation selama beberapa epoch. Grafik Akurasi Pelatihan versus Validasi menunjukkan bahwa model tersebut menunjukkan pembelajaran yang efektif, dibuktikan dengan peningkatan substansial dalam akurasi pelatihan dari awal epoch hingga mencapai stabilitas. Akurasi validasi juga mengalami peningkatan dan mulai stabil (konvergen) pada epoch ke-3 di kisaran nilai 0.75.

Pada grafik Loss Training versus Validation, training loss secara konsisten menurun seiring dengan peningkatan jumlah epoch, yang menunjukkan bahwa

model tersebut meningkatkan kemampuannya untuk meminimalkan kesalahan prediksi pada data pelatihan. Nilai validation loss mengalami penurunan di awal, namun cenderung stagnan dan sedikit meningkat setelah epoch ke-3. Hal ini merupakan dinamika yang wajar dalam pelatihan deep learning, namun secara keseluruhan model tetap mampu mempertahankan performa generalisasinya sebelum terjadi overfitting yang ekstrem.



Gambar 8. Grafik training dan validation accuracy



Gambar 9. Grafik training dan validation loss

Hasil Pengujian Model

Model Convolutional Neural Network (CNN) dilatih menggunakan 70% data, dengan 30% dialokasikan untuk pengujian, dan 20% data pelatihan dialokasikan untuk validasi. Prosedur pelatihan dilakukan maksimal 10 epoch dengan ukuran batch 32, menggunakan strategi Early Stopping dan Reduce Learning Rate on Plateau untuk mengurangi overfitting.

Berdasarkan hasil evaluasi pada data uji, **model CNN yang diusulkan mampu mencapai nilai akurasi sebesar 94%.** Hasil ini menunjukkan bahwa model tersebut memiliki kemampuan yang patut dipuji dalam mengkategorikan evaluasi pengguna aplikasi TikTok ke dalam tiga

klasifikasi sentimen: positif, netral, dan negative..

🌟 Akurasi pada data uji: 94.97% | Loss: 0.3440
94/94 ————— 1s 5ms/step

```

=== Classification Report ===

```

	precision	recall	f1-score	support
negatif	0.96	0.94	0.95	1224
netral	0.85	0.84	0.84	321
positif	0.96	0.98	0.97	1455
accuracy			0.95	3000
macro avg	0.92	0.92	0.92	3000
weighted avg	0.95	0.95	0.95	3000

Gambar 10. Classification report (precision, recall, dan F1-score)

Selain akurasi, penilaian kinerja model dilakukan dengan menggunakan metrik presisi, recall, dan F1-score untuk setiap kategori sentimen. Model menunjukkan efikasi klasifikasi yang luar biasa untuk sikap positif dan negatif. Sentimen Positif: Mencapai akurasi 0,96, recall 0,98, dan F1-score optimal 0,97, didukung oleh dataset terbesar dengan 1455 entri. Sentimen Negatif: Mencapai presisi 0,96, recall 0,94, dan F1-score 0,95 dari dataset yang mencakup 1224 entri. Sentimen Netral: Menunjukkan kinerja yang sedikit lebih rendah dibandingkan dengan dua kelas lainnya, dengan presisi 0,85, recall 0,84, dan F1-score 0,84 (diperoleh dari 321 titik data). Penurunan ini terutama disebabkan oleh karakteristik ulasan netral yang kerap kali ambigu atau mengandung campuran kata bersentimen lemah, sehingga lebih sulit diekstraksi fitur tekstualnya oleh model.

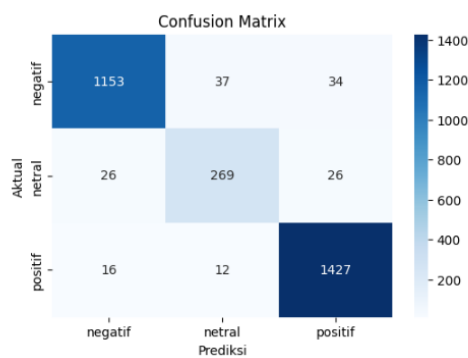
Hasil evaluasi menunjukkan bahwa model tersebut menunjukkan kinerja yang unggul baik dalam kategori sentimen positif maupun negatif. Sebaliknya, kinerja dalam kategori emosi netral relatif menurun, yang disebabkan oleh singkatnya dan ambiguitas penilaian netral.

Pemeriksaan matriks kebingungan menunjukkan bahwa sebagian besar kesalahan klasifikasi muncul di dalam kelas netral, meskipun prediksi untuk kelas positif dan negatif tetap relatif stabil. Penerapan class weighting terbukti membantu meningkatkan sensitivitas model terhadap kelas minoritas tanpa mengubah struktur data asli. Kinerja model

dalam membedakan kelas aktual dan kelas prediksi divisualisasikan melalui Confusion Matrix. Matriks ini mengonfirmasi temuan pada classification report dengan rincian prediksi benar (diagonal utama) sebagai berikut:

- Dari 1455 ulasan bersentimen positif, model berhasil menebak dengan benar sebanyak 1427 ulasan (hanya 28 ulasan yang meleset ke kelas lain).
- Dari 1224 ulasan bersentimen negatif, model menebak dengan benar sebanyak 1153 ulasan.
- Dari 321 ulasan bersentimen netral, model menebak dengan benar sebanyak 269 ulasan.

Tingkat kesalahan prediksi terbanyak terjadi pada ulasan aktual negatif yang diprediksi sebagai netral (37 data) dan positif (34 data). Namun, rasio kesalahan ini sangat kecil bila dibandingkan dengan total data uji. Dominasi nilai tinggi pada garis diagonal matriks memberikan bukti empiris bahwa arsitektur CNN yang diterapkan sangat tangguh dalam mengekstraksi dan mengklasifikasikan pola bahasa alami pada ulasan pengguna.



Gambar 11. Confusion matrix

SIMPULAN

Hasil pengujian dan penilaian menunjukkan bahwa metode Convolutional Neural Network (CNN) sangat berhasil untuk analisis sentimen ulasan pengguna aplikasi TikTok. Model yang dibangun mampu menghasilkan tingkat akurasi pengujian yang sangat tinggi, yakni mencapai 94.97%, dengan tingkat kesalahan (loss) sebesar 0.3440. Model ini menunjukkan akurasi prediksi yang luar

biasa untuk kelas mayoritas, seperti yang ditunjukkan oleh nilai F1-Score sebesar 0,97 untuk sentimen positif dan 0,95 untuk sentimen negatif. Meskipun klasifikasi pada sentimen netral (F1-score 0.84) menunjukkan performa yang sedikit lebih rendah dikarenakan ambiguitas karakteristik teks ulasan netral serta proporsi data yang lebih minoritas, secara keseluruhan model CNN mampu melakukan generalisasi data dengan sangat baik. Untuk pengembangan penelitian selanjutnya, disarankan untuk melakukan penyeimbangan data (data balancing) yang lebih ketat pada kelas minoritas atau mengombinasikan CNN dengan arsitektur lain seperti LSTM (Long Short-Term Memory) untuk menangkap konteks semantik bahasa campuran yang lebih panjang, sehingga akurasi pada kelas yang ambigu dapat semakin ditingkatkan.

DAFTAR PUSTAKA

- Bersastra, K. (2020). *Key words: VI(2)*, 147–157.
- Humam, M. N. (2023). Perbandingan Kinerja CNN dan Naïve Bayes pada Analisis Sentimen Performa Manchester United di Twitter. *Journal of Information Engineering and Educational Technology*, 7(2), 83–91.
<https://doi.org/10.26740/jieet.v7n2.p83-91>
- Husain, N. P., Sukirman, S., & SAJIAH, S. (2024). Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine. *Journal of System and Computer Engineering (JSCE)*, 5(1), 91–102.
<https://doi.org/10.61628/jsce.v5i1.1105>
- Pangaribuan, N. I., Simbolon, A. S., Aruan, N. M., Toba, K., & Utara, S. (2021). *SENTIMENT ANALYSIS FOR E-LEARNING APPLICATION USING*. 16–25.
<https://doi.org/10.47002/seminastika.v3i1.236>

Rizaldi, S. A. R., Alam, S., & Kurniawan, I. (2023). Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 109–117. <https://doi.org/10.55123/storage.v2i3.2334>

Buku

Emanuel, A. (1986). *Introduction of information system*. Jakarta: PT Grafika.

Jurnal Ilmiah

Munaf, D. R. (2012). Perancangan system informasi berbasis mobile. *Jurnal JNTETI*, 1(2): 12-17.

Website

Saidi, A. I. (2014, Juli 2). forumnfc.com .diakses pada 15 November 2014, pukul 15.00 dari www.warungnarasi.com: <http://www.warungnarasi.com/nalar-memilih-presiden-dalam-perspektif-semiotika>