



Performance Comparison of GRU and CNN-1D in Sentiment Analysis for the Minister of Finance Replacement on Social Media X

Perbandingan Kinerja Model GRU dan CNN-1D untuk Analisis Sentimen Pergantian Menteri Keuangan pada Media Sosial X

**Yoggy Montana Hendry^{1*}, Gerry Firmansyah²,
Hani Dewi Ariessanti³, Agus Herwanto⁴**

^{1,2,3,4}Fakultas Ilmu Komputer, Universitas Esa Unggul, Indonesia

E-Mail: ¹yoggymh26@student.esaunggul.ac.id, ²gerry@esaunggul.ac.id, ³hani.dewi@esaunggul.ac.id, ⁴agus.herwanto@esaunggul.ac.id

Received Jan 02nd 2026; Revised May 04th 2026; Accepted May 30th 2026; Available Online Jun 20th 2026

Corresponding Author: Yoggy Montana Hendry

Copyright © 2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

The social media platform X (formerly Twitter) has become an important source for understanding public opinion on social and political issues. The appointment of Purbaya Yudhi Sadewa as Minister of Finance elicited diverse public reactions, making it a relevant topic for sentiment analysis. This study compares the performance of two deep learning architectures, one-dimensional Convolutional Neural Network (CNN-1D) and Gated Recurrent Unit (GRU), in classifying Indonesian-language tweets into positive, neutral, and negative categories. A total of 3,567 tweets were collected via the official API between September 8 and October 8, 2025. The data were preprocessed through cleansing, case folding, normalization, and tokenization. Sentiment labeling was conducted using a lexicon-based approach (InSet), resulting in 1,957 positive (54.86%), 934 negative (26.18%), and 676 neutral (18.95%) sentiments. The results indicate that CNN-1D outperformed GRU, achieving 80% accuracy compared to 65%. CNN-1D was more effective in extracting local textual features, whereas GRU focused on sequential dependencies. These findings emphasize the importance of aligning model architecture with dataset characteristics in sentiment analysis of short social media texts.

Keywords: CNN-1D, GRU, Deep learning, Minister of Finance, Sentiment Analysis, Social Media X

Abstrak

Media sosial X (sebelumnya Twitter) telah menjadi sumber penting untuk memahami opini publik terhadap berbagai isu sosial dan politik. Salah satu isu yang memicu beragam tanggapan masyarakat adalah pergantian Menteri Keuangan Purbaya Yudhi Sadewa. Penelitian ini bertujuan membandingkan kinerja dua arsitektur *deep learning*, yaitu Convolutional Neural Network satu dimensi (CNN-1D) dan Gated Recurrent Unit (GRU), dalam mengklasifikasikan sentimen publik berbahasa Indonesia ke dalam kategori positif, netral, dan negatif. Sebanyak 3.567 tweet dikumpulkan melalui API resmi platform X dalam periode 8 September hingga 8 Oktober 2025. Data diproses melalui tahapan pembersihan (*cleansing data*), *case folding*, normalisasi, dan tokenisasi. Pelabelan sentimen dilakukan menggunakan pendekatan berbasis leksikon (InSet), menghasilkan distribusi positif sebesar 1.957 (54,86%), negatif sebesar 934 (26,18%), dan netral sebesar 676 (18,95%). Hasil evaluasi menunjukkan bahwa CNN-1D mengungguli GRU dengan akurasi 80%, dibandingkan dengan 65% pada GRU. CNN-1D lebih efektif dalam mengekstraksi fitur lokal pada teks pendek, sedangkan GRU berfokus pada pemodelan dependensi sekuensial. Temuan ini menegaskan pentingnya kesesuaian antara arsitektur model dan karakteristik data dalam analisis sentimen pada teks pendek.

Kata Kunci: Analisis Sentimen, CNN-1D, GRU, *Deep learning*, Menteri Keuangan, Sosial Media X

1. PENDAHULUAN

Media sosial kini menjadi kanal utama bagi masyarakat untuk menyampaikan opini terhadap isu-isu terkini, termasuk kebijakan ekonomi dan pergantian pejabat publik [1]. Pada 8 September 2025, Purbaya Yudhi Sadewa dilantik sebagai Menteri Keuangan, menggantikan Sri Mulyani Indrawati, sebuah peristiwa yang menarik perhatian luas dari publik serta diliput secara nasional dan internasional [2]. Pelantikan tersebut

memunculkan berbagai reaksi di platform digital, mulai dari komentar tokoh publik hingga perbincangan masyarakat mengenai arah kebijakan ekonomi baru. Oleh karena itu, analisis sentimen berbasis data menjadi penting untuk memahami kecenderungan opini publik terhadap kebijakan pemerintah di era digital.

Analisis sentimen umumnya dilakukan menggunakan pendekatan machine learning, khususnya *deep learning*. Berdasarkan kajian literatur terbaru, versi lain dari Recurrent Neural Network (RNN) dan Convolutional Neural Network (CNN) ditemukan sebagai metode populer dalam analisis sentimen berbasis *deep learning* [3]. Dua arsitektur yang banyak digunakan adalah CNN dan Gated Recurrent Unit (GRU). CNN dikenal unggul dalam mengekstraksi pola lokal (fitur n-gram) dari data teks [4], sedangkan GRU dirancang lebih sederhana (hanya dua gerbang) sehingga mengurangi jumlah parameter dan mempercepat komputasi tanpa mengorbankan kinerja [5]. Bahkan, penelitian yang dilakukan oleh Matobobo & Garidzira (2025) membandingkan CNN satu-dimensi (1D) dengan CNN dua-dimensi (2D) untuk analisis sentimen teks berurutan, menyoroti kelebihan CNN-1D dalam mengekstraksi pola teks [6].

Sejumlah penelitian terdahulu menunjukkan bahwa kedua model ini memiliki kinerja tinggi dalam analisis sentimen berbahasa Indonesia. Penelitian yang dilakukan oleh Angraini & Lhaksmana (2024) menemukan bahwa CNN mencapai akurasi sebesar 93,27% [7]. Sementara itu, Adam & Setiawan (2023) menemukan bahwa GRU mampu mencapai akurasi hingga 88,64% [8]. Hasil-hasil tersebut menunjukkan potensi besar dari kedua arsitektur dalam memahami opini publik melalui teks pendek seperti *tweet*.

Namun, berdasarkan penelitian sebelumnya yang secara langsung membandingkan kinerja model CNN-1D dengan GRU pada isu sosial-politik di Indonesia, khususnya terkait pergantian pejabat publik, penelitian tersebut masih terbatas. Kebanyakan penelitian sebelumnya menggunakan model gabungan atau salah satu dari *dataset* pembanding dari penelitian ini atau *dataset* dari lingkup yang berbeda. Selain itu, CNN-1D dan juga GRU memiliki keunggulan masing-masing dalam melakukan pekerjaan analisis sentimen.

Oleh sebab itu, penelitian ini berfokus untuk menutup celah penelitian tersebut dengan menganalisis dan membandingkan kinerja CNN-1D dan GRU dalam klasifikasi sentimen publik terhadap pergantian Menteri Keuangan di media sosial X (Twitter). Kajian ini diharapkan dapat memberikan kontribusi empiris terhadap pengembangan analisis sentimen berbahasa Indonesia serta memperkuat pemanfaatan *deep learning* dalam memahami dinamika opini masyarakat di ranah digital nasional.

2. TINJAUAN PUSTAKA

Sebelum melakukan penelitian, penting untuk meninjau jurnal yang membahas topik serupa. Jurnal-jurnal tersebut akan digunakan sebagai acuan atau referensi untuk memperkuat landasan teoritis dan metodologis dalam penulisan dan penyusunan penelitian ini. Berikut adalah beberapa jurnal yang peneliti gunakan selama penelitian.

Penelitian pertama yang dilakukan oleh Riyadi dkk. (2025) mengkaji analisis sentimen media sosial terkait calon presiden Indonesia dengan membandingkan beberapa model *deep learning*, yaitu CNN, RNN, CNN-RNN Hybrid, dan CNN-RNN Hybrid yang dimodifikasi. Penelitian ini bertujuan untuk meningkatkan performa klasifikasi sentimen melalui kombinasi arsitektur dan penyeimbangan data. Hasil penelitian menunjukkan bahwa model CNN memperoleh akurasi sebesar 90% dengan *F1-score* 0,88; RNN memperoleh akurasi sebesar 88% dengan *F1-score* 0,84; CNN-RNN Hybrid memperoleh akurasi sebesar 86% dengan *F1-score* 0,82; sedangkan CNN-RNN Hybrid yang dimodifikasi menghasilkan performa terbaik dengan akurasi sebesar 94% dan *F1-score* 0,95, yang menunjukkan peningkatan signifikan dibandingkan dengan model tunggal [9].

Penelitian kedua yang dilakukan oleh Irawan & Rochmah (2022) menerapkan algoritma Convolutional Neural Network (CNN) untuk mengetahui sentimen masyarakat terhadap kebijakan vaksin Covid-19 di Indonesia berdasarkan data Twitter. Penelitian ini bertujuan untuk membandingkan performa CNN dengan metode Naïve Bayes dalam mengklasifikasikan sentimen positif dan negatif. Metode CNN dan Naïve Bayes diuji menggunakan *dataset* yang sama setelah melalui tahapan *preprocessing*. Hasil penelitian menunjukkan bahwa CNN memperoleh rata-rata akurasi sebesar 98,66%, sedangkan Naïve Bayes memperoleh akurasi sebesar 94,66%, sehingga CNN terbukti lebih unggul dalam menangkap pola sentimen masyarakat pada data teks media sosial [10].

Penelitian ketiga yang dilakukan oleh Putri Inombi dkk. (2025) menganalisis sentimen pengguna media sosial X (Twitter) terhadap kebijakan Tabungan Perumahan Rakyat (TAPERA) di Indonesia dengan menerapkan metode CNN dan Bidirectional Encoder Representations from Transformers (BERT). Penelitian ini bertujuan untuk membandingkan performa kedua metode *deep learning* tersebut dalam mengklasifikasikan sentimen positif dan negatif masyarakat. *Dataset* diperoleh melalui proses *crawling* terhadap 1.790 tweet yang kemudian melalui tahapan *preprocessing*, meliputi *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. Hasil penelitian menunjukkan bahwa model BERT memperoleh performa yang sedikit lebih baik dibandingkan dengan CNN, dengan akurasi sebesar 86%, precision 78%, recall 61%, dan *F1-score* 68%, sedangkan CNN memperoleh akurasi 85%, precision 74%, recall 56%, dan *F1-score* 64%. Temuan ini BERT dan CNN sama-sama efektif dalam memahami konteks sentimen publik terhadap kebijakan TAPERA [11].

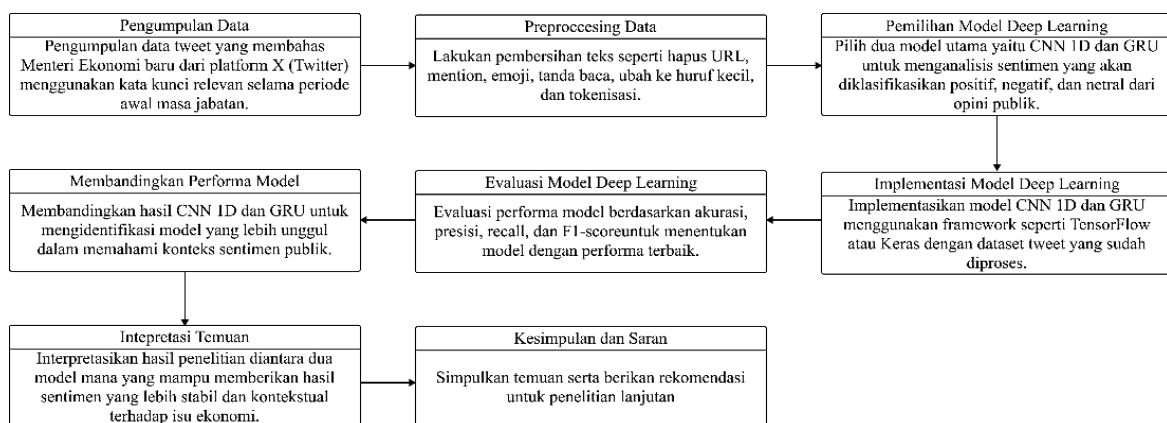
Penelitian keempat yang dilakukan oleh Febri Haerani (2026) mengkaji program MBG pada media sosial X, namun fokus pada analisis sentimen berbasis GRU. *Dataset* berjumlah 10.524 komentar/ulasan pengguna. Pasca praproses (pembersihan teks, *stopword removal*, dsb.), model GRU dilatih untuk membedakan sentimen positif dan negatif. Hasil menunjukkan bahwa mayoritas opini bersentimen positif terhadap program MBG, dengan model GRU menghasilkan *precision*, *recall*, dan *F1-score* sebesar 91% untuk sentimen positif dan 88% untuk sentimen negatif. Temuan ini menegaskan bahwa GRU efektif dalam memproses data besar dari media sosial untuk menangkap persepsi publik [12].

Penelitian kelima yang dilakukan oleh Dwi Jayanto dkk. (2025), melakukan pengujian terhadap model GRU yang menghasilkan akurasi tinggi sebesar 89% dalam mengklasifikasikan sentimen publik terhadap fatwa MUI terkait boikot produk pro-Israel. Keberhasilan ini menunjukkan bahwa GRU mampu menangkap dependensi sekuensial dan konteks kalimat panjang yang sering muncul dalam opini sosial di platform X. Namun, efektivitas GRU menurun ketika diterapkan pada *dataset* kecil dan tidak seimbang, dengan akurasi hanya 57%. Temuan tersebut menegaskan bahwa ukuran dan distribusi data berperan penting dalam menentukan kinerja model berbasis RNN seperti GRU [13].

Berdasarkan tinjauan terhadap penelitian-penelitian terdahulu, dapat diidentifikasi adanya celah penelitian yang masih belum banyak dikaji secara mendalam. Sebagian besar penelitian fokus pada perbandingan performa antarmodel *deep learning*, seperti CNN, GRU, dan BERT, dengan penekanan pada peningkatan akurasi klasifikasi sentimen. Namun demikian, masih terbatasnya penelitian yang secara khusus menganalisis perbandingan performa model berbasis sekuensial (RNN/GRU) dan model konvolusional (CNN 1D) pada konteks figur atau kebijakan publik tertentu dengan karakteristik data yang kompleks dan dinamis di platform X. Selain itu, aspek pengaruh kualitas data, proses pelabelan, serta potensi overfitting dan underfitting terhadap performa model belum dibahas secara eksplisit dalam penelitian-penelitian tersebut. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan melakukan perbandingan kinerja GRU dan CNN 1D pada analisis sentimen data media sosial X, disertai evaluasi terhadap kualitas data dan proses pelabelan guna memperoleh model yang lebih baik dan dapat digeneralisasi dengan baik.

3. METODOLOGI PENELITIAN

Metodologi yang digunakan dalam penelitian ini menggunakan pendekatan eksperimen komparatif, di mana dua arsitektur *deep learning* diuji untuk menilai kinerjanya dalam klasifikasi sentimen publik. Kemudian proses dimulai dengan pengumpulan data, *pre-processing*, pemilihan model, lalu model yang dipilih diimplementasikan dan dilatih menggunakan data yang telah diproses. Tahap terakhir adalah evaluasi model, di mana kinerja kedua model dibandingkan menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score* untuk menentukan model yang paling efektif dalam memahami opini publik terhadap Menteri Keuangan yang baru [14]. Pada Gambar 1 disajikan alur penelitian ini.



Gambar 1. Metodologi Penelitian

3.1. Objek Penelitian

Objek pada penelitian yang dilakukan adalah data teks berupa *tweets* yang berasal dari platform media sosial X yang membahas isu pergantian Menteri Keuangan Purbaya Yudhi Sadewa. Data yang digunakan mencakup opini publik berbahasa Indonesia yang diambil dalam rentang waktu 8 September 2025 hingga 8 Oktober 2025, saat isu tersebut ramai diperbincangkan di media sosial.

3.2. Teknik Pengumpulan Data

Pada penelitian ini, pengumpulan data pada penelitian ini dilakukan dengan metode *authenticated data crawling* menggunakan *library* Tweet Harvest yang dijalankan pada *environment* Node.js. Metode ini

digunakan untuk mengotomatisasi proses pengambilan data melalui API resmi X yang mendukung autentikasi token (*Bearer Token*).

Proses *crawling* dilakukan dengan menggunakan kata kunci pencarian yang digunakan meliputi “*purbaya*” dan “*menkeu*”, dengan periode pengambilan data mulai 8 September 2025 hingga 8 Oktober 2025. Berdasarkan hasil *crawling*, terkumpul sebanyak 3.567 *tweet*. Pada Tabel 1 adalah sebagian data yang sudah dikumpulkan. Pada Tabel 1 merupakan hasil *crawling*.

Tabel 1. Hasil *Crawling*

No	Tanggal	Waktu	Sentimen
1	2025-10-08	23:52:52	Kabar Gembira Menkeu Purbaya Ungkap Ada Stimulus Ekonomi Baru yang Diumumkan Pekan Depan lewat #zonauang #bantuansosial #bansos #purbaya #menkeu
2	2025-10-08	23:37:15	Menkeu Purbaya Mengajak Presiden Prabowo Untuk Patroli Anggaran
3	2025-10-08	23:36:40	Menkeu Purbaya Yudhi Sadewa mengatakan jika ekonomi tumbuh lebih cepat manfaatnya paling banyak dirasakan oleh kelas menengah yang selama beberapa tahun terakhir terpukul
...	...	...	...
3565	2025-09-08	16:36:03	Menkeu Purbaya Pede Ekonomi RI Tembus Persen dalam Dua Tahun
3566	2025-09-08	16:17:00	Di era Menkeu Purbaya Yudhi Sadewa yang pro pertumbuhan target GDP sektor saham potensial menarik meliputi Komoditas nikel batubara sawit Manfaat dari rupiah lemah ekspor USD Konsumen Dorong dari belanja pemerintah stimulus miliar Teknologi
3567	2025-09-08	15:54:13	TARGET GILA Menteri Keuangan Baru Ramal IHSG Tembus Blunder atau Visi Jenius GEGER Baru dilantik Menkeu Purbaya langsung pasang target yang bikin pasar saham heboh IHSG capai level di tahun investasi saham #financialfreedom #belajarinvestasi #cuan

3.3. Preprocessing Data

Pada tahap ini dilakukan proses *preprocessing data* untuk mempersiapkan *dataset* sebelum dilakukan analisis lebih lanjut. Tahapan ini bertujuan untuk memastikan data dalam kondisi bersih, terstruktur, dan siap diolah oleh model analisis sentimen [15]. Secara keseluruhan, tahapan *preprocessing* ini meliputi pembersihan data (*cleansing data*), *case folding*, normalisasi data, tokenisasi, *labeling* serta visualisasi awal data melalui grafik dan *word cloud* [16]. Proses ini menjadi langkah penting sebelum data digunakan dalam pelatihan dan evaluasi model analisis sentimen.

3.3.1. Cleansing data

Pada tahap *cleansing data* proses ini mencakup penghapusan data yang bernilai kosong (*missing values*) dan data yang terduplikasi agar tidak menimbulkan bias dalam hasil analisis. Selanjutnya, dilakukan pembersihan teks pada setiap *tweet* dengan menghapus elemen-elemen yang tidak relevan seperti *emoticon*, tautan URL, tag HTML, serta berbagai simbol atau ikon lain seperti tanda pagar (*hashtag*), tanda seru, dan karakter khusus lainnya [17]. Langkah ini bertujuan agar teks yang tersisa benar-benar merepresentasikan isi pesan yang akan dianalisis, serta meminimalkan gangguan dari elemen non-verbal yang tidak berkontribusi terhadap pemaknaan sentimen [18]. Tabel 2 terdapat perbandingan sebelum dan setelah *cleansing data*.

Tabel 2. Perbandingan *Cleansing data*

Sebelum Clenasing Data	Setelah <i>Cleansing data</i>
Keputusan Menteri Keuangan (Menkeu) Purbaya Yudhi Sadewa memperkecil transfer ke daerah (TKD) tahun 2026 sebesar Rp693 triliun atau terjun bebas ket... https://t.co/bAlZI3wNkM	Keputusan Menteri Keuangan Menkeu Purbaya Yudhi Sadewa memperkecil transfer ke daerah TKD tahun sebesar Rupiah triliun atau terjun bebas

3.3.2. Case folding

Pada tahap *case folding* yang dilakukan untuk menyeragamkan seluruh karakter dalam teks agar berada dalam bentuk huruf kecil (*lowercase*). Proses ini penting untuk menghindari perbedaan interpretasi yang disebabkan oleh variasi penulisan huruf besar dan huruf kecil serta membantu dalam meningkatkan konsistensi data teks dan serta memudahkan sistem analisis dalam mengenali kata secara akurat [15]. Pada Tabel 3 adalah perbandingan sebelum dan setelah dilakukan *Case folding*.

Tabel 3. Perbandingan *Case folding*

Sebelum <i>Case folding</i>	Setelah <i>Case folding</i>
Keputusan Menteri Keuangan Menkeu Purbaya Yudhi Sadewa memperkecil transfer ke daerah TKD tahun	keputusan menteri keuangan menkeu purbaya yudhi sadewa memperkecil transfer ke daerah tkd tahun sebesar

Sebelum <i>Case folding</i>	Setelah <i>Case folding</i>
sebesar Rupiah triliun atau terjun bebas	rupiah triliun atau terjun bebas

3.3.3. Normalisasi Kata

Pada tahap *normalisasi teks* dilakukan untuk menyeragamkan bentuk penulisan kata dalam setiap tweet agar memiliki makna yang jelas dan konsisten. Pada penelitian ini, proses normalisasi difokuskan pada penggantian kata-kata yang berbentuk singkatan, akronim, atau penyingkatan informal menjadi bentuk kata lengkapnya [15]. Tabel 4 adalah perbandingan sebelum dan setelah dilakukan normalisasi kata.

Tabel 4. Perbandingan Normalisasi Kata

Sebelum Normalisasi Kata	Setelah Normalisasi Kata
keputusan menteri keuangan menkeu purbaya yudhi sadewa memperkecil transfer ke daerah tkd tahun sebesar rupiah triliun atau terjun bebas	keputusan menteri keuangan menkeu purbaya yudhi sadewa memperkecil transfer ke daerah tkd tahun sebesar rupiah triliun atau terjun bebas

3.3.4. Tokenisasi Data

Setelah melalui proses pembersihan data, tahap berikutnya adalah tokenisasi. Tokenisasi merupakan proses memecah teks menjadi potongan-potongan kata yang disebut token. Setiap token mewakili satu unit kata yang akan menjadi dasar dalam analisis selanjutnya [15]. Tabel 5 adalah perbandingan sebelum dan setelah dilakukan tokenisasi data.

Tabel 5. Perbandingan Tokenisasi Data

Sebelum Tokenisasi Data	Setelah Tokenisasi Data
keputusan menteri keuangan menkeu purbaya yudhi sadewa memperkecil transfer ke daerah tkd tahun sebesar rupiah triliun atau terjun bebas	['keputusan', 'menteri', 'keuangan', 'menkeu', 'purbaya', 'yudhi', 'sadewa', 'memperkecil', 'transfer', 'ke', 'daerah', 'tkd', 'tahun', 'sebesar', 'rupiah', 'triliun', 'atau', 'terjun', 'bebas']

3.3.5. Klasifikasi Label

Proses klasifikasi label dilakukan secara semi-otomatis menggunakan pendekatan lexicon berbasis InSet (Indonesian Sentiment Lexicon) yang diperoleh dari GitHub. Pendekatan ini banyak digunakan dalam penelitian, di mana perbandingan jumlah kata bermuatan sentimen sering menjadi dasar pelabelan [19]. Setiap tweet yang telah melalui tahap *preprocessing* dianalisis dengan menghitung jumlah kata positif dan negatif. Tweet diberi label positif jika kata positif lebih dominan, label negatif jika kata negatif lebih banyak, dan label netral jika jumlahnya seimbang atau tidak mengandung kata bermuatan sentimen. Label positif mencerminkan emosi dukungan atau optimisme, label negatif menunjukkan kritik atau ketidakpuasan, sedangkan label netral bersifat informatif tanpa ekspresi emosional yang jelas. Berikut perhitungannya dijabarkan (1), (2), (3).

$$S_{positif} = \sum_{k=1}^n [w_k \in P] \quad (1)$$

$$S_{negatif} = \sum_{k=1}^n [w_k \in N] \quad (2)$$

Pada perhitungan (1) dan (2) misalkan setiap tweet terdiri dari kumpulan kata, kemudian daftar kata positif dan negatif dinyatakan sebagai P dan N, dan w_k adalah setiap kata pada data. Kemudian sentimen akan ditentukan dengan perbandingan skor sebagai berikut:

$$Label = \begin{cases} \text{Positif,} & \text{jika } S_{positif} > S_{negatif} \\ \text{Negatif,} & \text{jika } S_{negatif} > S_{positif} \\ \text{Netral,} & \text{jika } S_{positif} = S_{negatif} \end{cases} \quad (3)$$

Kemudian, setelah dilakukan klasifikasi label pada dataset, klasifikasi tersebut menghasilkan label positif sebanyak 54,86%, negatif sebanyak 26,18%, dan netral sebanyak 18,95%. Selanjutnya, dilakukan visualisasi untuk memudahkan pemahaman distribusi data. Gambar 2 adalah hasil klasifikasi label.

3.6. Konfigurasi Hyperparameter

Model *deep learning* baik CNN-1D ataupun GRU telah dilakukan konfigurasi *hyperparameter* yang telah ditentukan untuk memperoleh performa klasifikasi yang optimal dan stabil. Pada tabel 6, menunjukkan kedua model dilatih selama 25 *epoch* dengan *batch size* 32 dan *optimizer* Adam. Perbedaan utama terletak pada learning rate, di mana CNN-1D menggunakan 0.0003 dan GRU menggunakan 0.0005. Nilai dropout pada CNN-1D sebesar 0.2 dan 0.5, sedangkan pada GRU sebesar 0.3 dan 0.4. Jumlah unit/filter pada CNN-1D adalah 128, sedangkan GRU menggunakan 64 unit. Selain itu, *validation split* ditetapkan sebesar 0.2 pada CNN-1D dan 0.1 pada GRU. Konfigurasi ini dirancang untuk menyesuaikan karakteristik masing-masing arsitektur dalam menangani klasifikasi sentimen multi-kelas secara efektif.

Tabel 6. Konfigurasi Hyperparameter

List Hyperparameter		
Parameter	CNN-1D	GRU
Epoch	25	25
Batch Size	32	32
Optimizer	Adam	Adam
Learning Rate	0.0003	0.0005
Dropout	0.2 / 0.5	0.3 / 0.4
Units / Filters	128	64
Validation Split	0.2	0.1

3.7. Evaluasi Model

Evaluasi model merupakan tahap dalam penelitian untuk menilai sejauh mana model yang dibangun mampu menghasilkan prediksi yang akurat terhadap data uji. Pada penelitian ini, proses evaluasi dilakukan menggunakan empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *f1-score* [20]. Berikut perhitungannya dijabarkan pada persamaan (4), (5), (6), (7).

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Negative + False\ Positive} \quad (4)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (5)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (6)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

4. HASIL DAN PEMBAHASAN

4.1. Pemodelan CNN-1D

Proses pemodelan pertama yaitu arsitektur Convolutional Neural Network satu dimensi (CNN-1D) yang diterapkan untuk tugas analisis sentimen teks berbahasa Indonesia. Pemodelan dilakukan untuk mengevaluasi kemampuan CNN-1D dalam mempelajari pola lokal pada teks serta mengukur kinerja model melalui metrik pelatihan dan validasi. Evaluasi performa model disajikan dalam bentuk tabel ringkasan hasil pelatihan dan visualisasi grafik akurasi serta *loss* selama proses *training*. Berikut pada Tabel 6 adalah hasil pelatihan CNN-1D.

Tabel 7. Classification Report

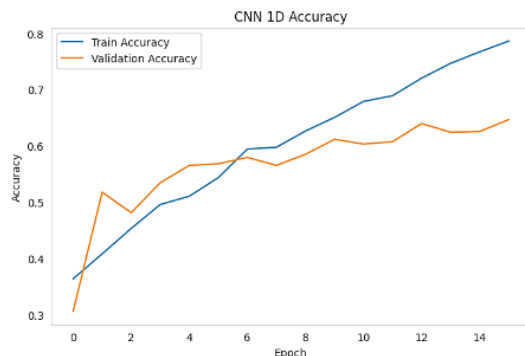
Training Model CNN-1D				
Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	
0.80	0.56	0.65	0.83	

Tabel 7 menunjukkan *Training Model* CNN-1D menyajikan ringkasan hasil pelatihan model yang mencakup *training accuracy*, *training loss*, *validation accuracy*, dan *validation loss*. Berdasarkan tabel tersebut, model CNN-1D memperoleh *training accuracy* sebesar 0,80 dengan *training loss* sebesar 0,56, yang menunjukkan bahwa model mampu mempelajari data latih dengan cukup baik. Sementara itu, *validation accuracy* sebesar 0,65 dan *validation loss* sebesar 0,83 mengindikasikan bahwa performa model pada data validasi masih lebih rendah dibandingkan dengan data latih. Perbedaan ini menunjukkan adanya indikasi overfitting ringan, di mana model cenderung lebih optimal pada data pelatihan dibandingkan dengan

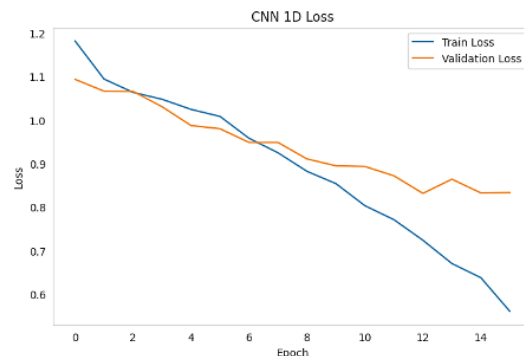
data yang belum pernah dilihat sebelumnya. Kemudian dari hasil yang sudah dijabarkan sebelumnya akan divisualisasikan pada Gambar 7 dan Gambar 8.

Pada Gambar 7 menunjukkan grafik CNN-1D *Accuracy* perkembangan nilai akurasi selama proses pelatihan. Garis akurasi pelatihan (*training accuracy*) menunjukkan peningkatan yang konsisten seiring bertambahnya *epoch*, menandakan bahwa model semakin baik dalam mengklasifikasikan data latih. Sementara itu, akurasi validasi (*validation accuracy*) juga mengalami peningkatan, namun cenderung lebih fluktuatif dan berada di bawah akurasi pelatihan. Pola ini mengindikasikan bahwa meskipun model mampu melakukan generalisasi, terdapat batasan dalam kemampuannya untuk mempertahankan performa yang stabil pada data validasi.

Kemudian pada Gambar 8 menunjukkan grafik CNN-1D *Loss* penurunan nilai *loss* pada data pelatihan secara bertahap, menandakan bahwa kesalahan prediksi model semakin berkurang selama proses *training*. Di sisi lain, *loss* validasi juga mengalami penurunan, namun dengan pola yang lebih lambat dan cenderung stabil di nilai yang lebih tinggi dibandingkan *loss* pelatihan. Kondisi ini memperkuat indikasi bahwa model CNN-1D mulai mengalami overfitting ringan, meskipun masih berada pada tingkat yang dapat diterima untuk klasifikasi sentimen.



Gambar 7. Plot Accuracy CNN-1D



Gambar 8. Plot Loss CNN-1D

4.2. Pemodelan GRU

Proses pemodelan yang terakhir adalah Gated Recurrent Unit (GRU). GRU merupakan salah satu arsitektur *recurrent neural network* yang dirancang untuk menangkap ketergantungan sekuensial dalam data teks, sehingga diharapkan mampu memahami konteks kalimat secara lebih menyeluruh. Evaluasi performa model GRU dilakukan melalui pengamatan metrik pelatihan dan validasi yang disajikan dalam bentuk tabel serta visualisasi grafik akurasi dan *loss* selama proses *training*. Berikut pada Tabel 7 adalah hasil pelatihan GRU.

Tabel 8. Classification Report

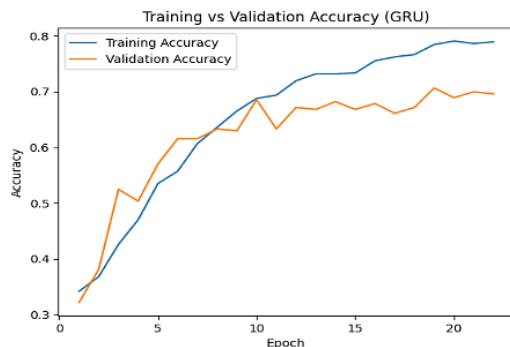
Training Model GRU			
Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
0.75	0.64	0.62	0.86

Tabel 8 menunjukkan bahwa Training Model GRU memperoleh *training accuracy* sebesar 0,75 dengan *training loss* sebesar 0,64. Sementara itu, nilai *validation accuracy* sebesar 0,62 dan *validation loss* sebesar 0,86 menunjukkan bahwa performa model pada data validasi masih berada di bawah performa pada data pelatihan. Perbedaan nilai ini mengindikasikan bahwa model GRU mampu mempelajari pola data latih dengan cukup baik, namun masih menghadapi tantangan untuk melakukan generalisasi pada data yang tidak terlihat sebelumnya. Kemudian, hasil yang telah dijabarkan sebelumnya akan divisualisasikan pada Gambar 9 dan Gambar 10.

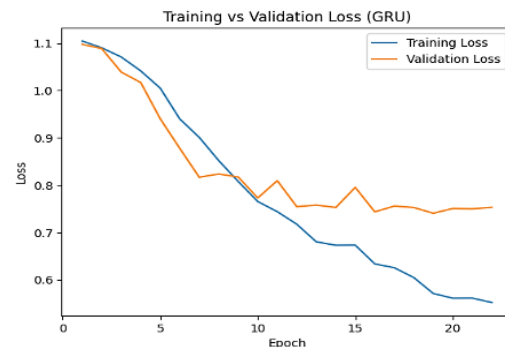
Gambar 9 menunjukkan grafik akurasi model GRU, yang menunjukkan bahwa akurasi pelatihan mengalami peningkatan yang relatif stabil seiring bertambahnya *epoch*, yang menandakan bahwa model GRU secara bertahap mampu menangkap pola sentimen dalam data latih. Akurasi validasi juga menunjukkan tren peningkatan, meskipun cenderung lebih fluktuatif dan berada di bawah akurasi pelatihan. Pola ini menunjukkan bahwa model GRU memiliki kemampuan generalisasi yang cukup, namun belum sepenuhnya optimal.

Selanjutnya, pada Gambar 10, grafik *Loss* model GRU menunjukkan penurunan nilai *loss* pada data pelatihan secara konsisten selama proses *training*, yang menandakan semakin berkurangnya kesalahan prediksi model. Di sisi lain, *loss* validasi mengalami penurunan pada awal *epoch*, kemudian cenderung stabil dengan fluktuasi kecil pada *epoch* berikutnya. Kondisi ini mengindikasikan bahwa model mulai mencapai

batas optimal pembelajaran dan berpotensi mengalami *overfitting* ringan, meskipun masih dalam batas yang dapat ditoleransi.



Gambar 9. Plot Accuracy GRU



Gambar 10. Plot Loss GRU

4.3. Evaluasi Model

Evaluasi model akan menghitung seberapa optimal masing-masing model berdasarkan metrik *precision*, *recall*, *F1-score*, *support*, serta akurasi secara keseluruhan. Evaluasi dilakukan untuk menilai kemampuan masing-masing model dalam melakukan klasifikasi sentimen positif, netral, dan negatif, serta membandingkan performa kedua model secara komprehensif. Berikut adalah Tabel 9 untuk model CNN-1D dan Tabel 10 untuk model GRU.

Tabel 9. Classification Report CNN-1D

Evaluasi Hasil Model CNN-1D				
Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Positif	90%	88%	89%	392
Netral	57%	48%	52%	135
Negatif	72%	85%	79%	187
Accuracy 80%				

Tabel 10. Classification Report GRU

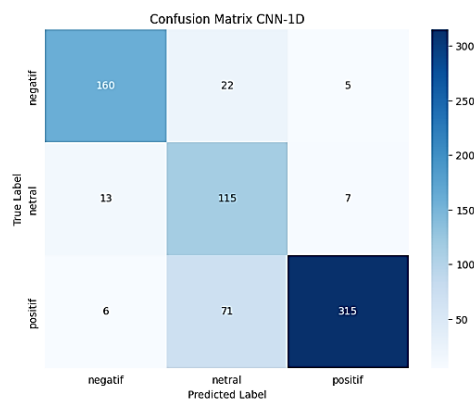
Evaluasi Hasil Model GRU				
Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Positif	77%	82%	79%	392
Netral	34%	39%	36%	135
Negatif	66%	50%	57%	187
Accuracy 65%				

Berdasarkan Tabel 9, model CNN-1D menunjukkan kinerja yang baik dengan akurasi keseluruhan sebesar 80%. Pada kelas positif, model mencapai *precision* 90%, *recall* 88%, dan *F1-score* 89%, yang menunjukkan kemampuan tinggi dalam mengenali sentimen positif secara konsisten. Pada kelas negatif, model juga menunjukkan performa yang cukup baik dengan *precision* 72%, *recall* 85%, dan *F1-score* 79%, menandakan bahwa sebagian besar data negatif berhasil diidentifikasi dengan benar. Namun, performa model pada kelas netral masih relatif rendah dengan *precision* 57%, *recall* 48%, dan *F1-score* 52%. Hal ini menunjukkan bahwa model CNN-1D masih mengalami kesulitan dalam membedakan sentimen netral dari sentimen positif maupun negatif, yang umumnya disebabkan oleh kemiripan pola linguistik antarkelas tersebut.

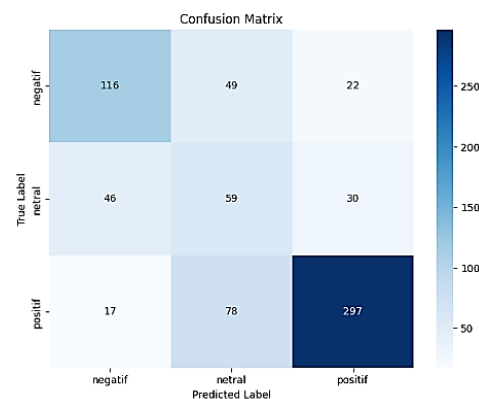
Sedangkan pada Tabel 10, model GRU menunjukkan akurasi keseluruhan sebesar 65%, yang lebih rendah dibandingkan dengan CNN-1D. Pada kelas positif, GRU memperoleh *precision* 77%, *recall* 82%, dan *f1-score* 79%, menandakan performa yang cukup baik namun masih berada di bawah CNN-1D. Pada kelas negatif, nilai *precision* 66%, *recall* 50%, dan *f1-score* 57% menunjukkan bahwa model GRU cenderung kurang konsisten dalam mengenali sentimen negatif. Performa terendah juga terjadi pada kelas netral, dengan *precision* 34%, *recall* 39%, dan *f1-score* 36%, yang mengindikasikan bahwa model GRU mengalami kesulitan yang lebih besar dalam mengklasifikasikan sentimen netral dibandingkan CNN-1D. Untuk memperjelas mengapa CNN-1D lebih unggul, akan divisualisasikan dalam bentuk *confusion Matrix* pada Gambar 11 dan Gambar 12.

Hasil confusion matrix ini menunjukkan mengapa CNN-1D lebih unggul dibandingkan GRU dalam klasifikasi sentimen. Pada CNN-1D, jumlah prediksi yang benar pada setiap kelas cenderung lebih tinggi dan kesalahan klasifikasi lebih rendah dibandingkan dengan GRU. Hal ini terlihat jelas pada kelas negatif dan

positif, di mana CNN-1D mampu mengklasifikasikan data dengan lebih akurat serta meminimalkan kesalahan ke kelas lain. Selain itu, distribusi nilai pada diagonal utama *confusion matrix* CNN-1D lebih dominan, yang mengindikasikan tingkat konsistensi dan akurasi model yang lebih baik secara keseluruhan.



Gambar 11. Confusion Matrix CNN-1D



Gambar 12. Confusion Matrix GRU

4.4. Evaluasi Hasil Akhir

Berdasarkan hasil evaluasi, CNN-1D menunjukkan performa lebih tinggi dibandingkan GRU dalam klasifikasi sentimen pada *dataset* penelitian ini. Keunggulan tersebut dipengaruhi oleh kesesuaian arsitektur model dengan karakteristik data. CNN-1D efektif dalam mengekstraksi fitur lokal seperti kata kunci dan pola n-gram yang secara eksplisit merepresentasikan sentimen pada teks pendek seperti tweet. Sebaliknya, GRU yang dirancang untuk menangkap dependensi jangka panjang kurang optimal pada *dataset* dengan ukuran terbatas dan struktur teks ringkas, sehingga proses generalisasi tidak maksimal. Struktur CNN-1D yang lebih sederhana dan fokus pada fitur diskriminatif membuatnya lebih adaptif dalam kondisi data terbatas.

Meskipun demikian, kesalahan klasifikasi masih terjadi, terutama pada kalimat ambigu dan sarkastik di kelas netral, yang menunjukkan keterbatasan model dalam memahami konteks pragmatik bahasa media sosial. Secara keseluruhan, perbedaan performa kedua model menegaskan bahwa efektivitas analisis sentimen sangat bergantung pada kesesuaian antara arsitektur model, panjang teks, dan ukuran *dataset*.

5. KESIMPULAN

Penelitian ini menunjukkan bahwa CNN-1D secara empiris mengungguli GRU dalam klasifikasi sentimen pada data teks media sosial X terkait isu sosial-politik ekonomi nasional. CNN-1D memperoleh akurasi lebih tinggi dibandingkan GRU, serta menunjukkan tingkat kesalahan yang lebih rendah dan prediksi yang lebih stabil. Hasil ini secara langsung menjawab tujuan penelitian mengenai perbandingan efektivitas kedua arsitektur pada teks pendek. Keunggulan CNN-1D terletak pada kemampuannya mengekstraksi fitur lokal seperti kata dan n-gram yang dominan pada struktur tweet, sementara GRU yang dirancang untuk dependensi jangka panjang tidak termanfaatkan secara optimal pada *dataset* dengan panjang teks terbatas. Temuan ini menegaskan bahwa kesesuaian antara karakteristik data dan arsitektur model menjadi faktor kunci dalam menentukan performa analisis sentimen, serta memberikan kontribusi empiris pada studi sentimen teks pendek berbahasa Indonesia.

REFERENSI

- [1] W. Nugraha, M. Tito Julianto, M. Khoirun Najib, dan E. Khatizah, "Public Sentiment Toward the Indonesian Capital Relocation Policy on X Using a BiLSTM-CNN Model," *Jurnal Telematika*, vol. 20, no. 2, hlm. 114–126, Jan 2026, doi: 10.61769/telematika.v20i2.796.
- [2] S. Sulaiman dan G. Suroyo, "Indonesia Replaces Respected Finance Minister With Economist Promising Rapid Growth," Reuters. Diakses: 25 Oktober 2025. [Daring]. Tersedia pada: <https://www.reuters.com/world/asia-pacific/indonesia-replaces-respected-finance-minister-with-economist-promising-rapid-2025-09-08/>
- [3] F. Fitroh dan F. Hudaya, "Systematic Literature Review: Analisis Sentimen Berbasis *Deep learning*," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 2, hlm. 132–140, Agu 2023, doi: 10.25077/teknosi.v9i2.2023.132-140.
- [4] F. Fahry, T. C. Miswaty, dan H. Harun, "Analyzing Marketplace Reviews Using Word2Vec, CNN, and Deep K-Means with Sociolinguistic Approaches," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 6, hlm. 5489–5502, Des 2025, doi: 10.52436/1.jutif.2025.6.6.5340.
- [5] M. N. A. P. Khano, D. R. S. Saputro, Sutanto, dan A. Wibowo, "Sentiment Analysis With Long-Short Term Memory (LSTM) And Gated Recurrent Unit (GRU) Algorithms," *Barekeng*, vol. 17, no. 4, hlm. 2235–2242, Des 2023, doi: 10.30598/barekengvol17iss4pp2235-2242.

- [6] C. Matobobo dan T. C. Garidzira, "Examining 1D and 2D CNN Architectures in Comparison for Sentiment Analysis in Sequential Data: A Case Study of Spotify Music Reviews," *The Indonesian Journal of Computer Science*, vol. 14, no. 2, Apr 2025, doi: 10.33022/ijcs.v14i2.4740.
- [7] N. A. Angraini dan K. M. Lhaksmana, "Sentiment Analysis About Legislative Elections using *Deep learning* with LSTM and CNN Models," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 1, Jun 2024, doi: 10.47065/bits.v6i1.5283.
- [8] A. Z. R. Adam dan E. B. Setiawan, "Social Media Sentiment Analysis Using Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU)," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 1, hlm. 119–131, Feb 2023, doi: 10.26555/jiteki.v9i1.25813.
- [9] S. Riyadi, F. D. Daffa, C. D. Damarjati, dan M. S. A. M. A. Ali, "Sentiment Analysis on Social Media Uusing CNN-RNN Hybrid: A Case Study of Indonesian Presidential Candidate," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, Mei 2025, doi: 10.22219/kinetik.v10i2.2125.
- [10] F. A. Irawan dan D. A. Rochmah, "Penerapan Algoritma CNN Untuk Mengetahui Sentimen Masyarakat Terhadap Kebijakan Vaksin Covid-19," *JURNAL INFORMATIKA*, vol. 9, no. 2, 2022, [Daring]. Tersedia pada: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [11] S. Putri Inombi, D. Rahmawaty Isa, S. K. Nasib, I. K. Hasan, dan S. Nurmardia Abdussamad, "Analisis Sentimen Pengguna X (Twitter) Terhadap Kebijakan Tapera Di Indonesia Menggunakan Metode CNN Dan BERT," *Jurnal Matematika dan Ilmu Pengetahuan Alam*, vol. 6, 2025, doi: 10.3483/trigonometri.v1i1.800.
- [12] F. Haerani, "Analisis Sentimen Program Makan Bergizi Gratis Berdasarkan Komentar Di Media Sosial X Menggunakan GRU," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 1, Jan 2026, doi: 10.23960/jitet.v14i1.8754.
- [13] W. Dwi Jayanto, T. Feri Efendi, dan S. Rokhmah, "Analisis Sentimen Publik Pada Aplikasi X Terhadap Fatwa Mui Produk Pro Israel Dengan Metode Gated Recurrent Unit," Jun 2025, doi: 10.8734/Kohesi.v1i2.36.
- [14] R. K. Das, M. Islam, M. M. Hasan, S. Razia, M. Hassan, dan S. A. Khushbu, "Sentiment analysis in multilingual context: Comparative analysis of machine learning and hybrid *deep learning* models," *Heliyon*, vol. 9, no. 9, Sep 2023, doi: 10.1016/j.heliyon.2023.e20281.
- [15] A. F. Aufar, Mochamad Alfian Rosid, A. Eviyanti, dan I. R. I. Astutik, "Optimizing Text Preprocessing for Accurate Sentiment Analysis on E-Wallet Reviews," *JICTE (Journal of Information and Computer Technology Education)*, vol. 7, no. 2, hlm. 42–50, Okt 2023, doi: 10.21070/jicte.v7i2.1650.
- [16] K. K. A. G. R. Atmaja dan I. G. I. Sudipa, "Sentiment Analysis of Public Opinion towards the Crypto Academy Class Using the Convolutional Neural Network Method," 2025. [Daring]. Tersedia pada: <https://isarpublisher.com/journal/isarjst>
- [17] J. Homepage dkk., "MALCOM: Indonesian Journal of Machine Learning and Computer Science Sentiment Review Analysis on the BSB Mobile Application Using Convolutional Neural Networks and Support Vector Machine Algorithm Analisis Ulasan Sentimen pada Aplikasi Mobile BSB Menggunakan Algoritma Convolutional Neural Networks dan Support Vector Machine," vol. 6, no. 2, hlm. 829–838, 2026, doi: 10.57152/malcom.v6i2.2637.
- [18] Y. Fauziah, B. Yuwono, dan A. S. Aribowo, "Lexicon Based Sentiment Analysis in Indonesia Languages : A Systematic Literature Review," *RSF Conference Series: Engineering and Technology*, vol. 1, no. 1, hlm. 363–367, Des 2021, doi: 10.31098/cset.v1i1.397.
- [19] Y. Asri, D. Kuswardani, W. N. Suliyanti, Y. O. Manullang, dan A. R. Ansyari, "Sentiment analysis based on Indonesian language lexicon and IndoBERT on user reviews PLN mobile application," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 38, no. 1, hlm. 677, Apr 2025, doi: 10.11591/ijeecs.v38.i1.pp677-688.
- [20] S. Riyadi, F. D. Daffa, C. D. Damarjati, dan M. S. A. M. A. Ali, "Sentiment Analysis on Social Media Uusing CNN-RNN Hybrid: A Case Study of Indonesian Presidential Candidate," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, Mei 2025, doi: 10.22219/kinetik.v10i2.2125.