PROSIDING SEMINAR NASIONAL LINGKUNGAN LAHAN BASAH
"Meningkatkan Peran Teknologi dan Kecerdasan Buatan dalam Implementasi Ekonomi Sirkuler untuk Mencapai Tujuan Pembangunan Berkelanjutan"
Volume 11 No. 2 – Juni 2026

EVALUASI CONVOLUTIONAL NEURAL NETWORK DALAM ANALISIS SENTIMEN TEKS MEDIA SOSIAL DENGAN PEMBELAJARAN TERAWASI

Muhammad Bahit[a], Yuslena Sari[b], Andreyan Rizky Baskara[c], Eka Setya Wijaya[d], Harry Pratama Yunus[e] & Alysa Armelia[f]

[a-f]Universitas Lambung Mangkurat (ULM) Banjarmasin

Email korespondensi: muhammadbahit@ulm.ac.id

ABSTRAK

Perkembangan teknologi informasi dan komunikasi mendorong meningkatnya penggunaan media sosial sebagai sarana berbagi informasi, opini, maupun pengalaman sehari-hari. Analisis sentimen pada teks media sosial menghadapi sejumlah tantangan, seperti panjang teks yang relatif singkat, penggunaan bahasa tidak baku, singkatan, serta keberadaan slang atau bahasa gaul. Oleh karena itu, diperlukan metode yang mampu menangani keragaman Bahasa. Penelitian ini bertujuan untuk mengevaluasi kinerja Convolutional Neural Network (CNN) dalam klasifikasi sentimen berbasis pembelajaran terawasi. Arsitektur CNN yang digunakan berbasis 1D convolution layer dengan input berupa representasi word embedding hasil transformasi teks menjadi vektor numerik berdimensi tetap. Proses konvolusi dilakukan untuk mengekstraksi fitur semantik dari urutan kata dengan lapisan max pooling untuk mereduksi dimensi fitur dan mempertahankan informasi paling representatif. Hasil ekstraksi fitur kemudian diteruskan ke lapisan fully connected dengan aktivasi Softmax untuk klasifikasi sentimen positif dan negatif. Evaluasi performa model dilakukan menggunakan kurva AUC-ROC dan Confusion Matrix. Hasil penelitian menunjukkan bahwa model CNN mampu memberikan performa klasifikasi yang baik dengan nilai AUC mendekati 1 dan tingkat kesalahan yang rendah pada prediksi false negative. Hasil penelitian ini menunjukkan bahwa CNN dengan 1D convolution efektif dalam menangani data teks tidak terstruktur dan menjadi alternatif untuk analisis sentimen berbasis deep learning.

Kata kunci: *Analisis Sentimen, Convolutional Neural Network, Pembelajaran Terawasi*

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang sangat pesat dalam beberapa tahun terakhir telah membawa perubahan signifikan dalam berbagai aspek kehidupan manusia (Wu, 2025). Salah satu wujud perkembangan tersebut adalah muncul dan meluasnya penggunaan media sosial sebagai salah satu kanal utama komunikasi, penyebaran informasi, dan interaksi masyarakat di era digital (Grover et al., 2022). Media sosial seperti Twitter, Facebook, Instagram, TikTok, dan YouTube tidak hanya menjadi ruang hiburan atau sarana pertemanan, tetapi juga berkembang sebagai medium penting dalam menyalurkan aspirasi, opini, serta pandangan masyarakat mengenai berbagai isu, baik yang bersifat pribadi, sosial, politik, ekonomi, maupun budaya (Xu et al., 2022).

Jumlah pengguna media sosial yang masif menyebabkan volume data teks yang dihasilkan meningkat secara eksponensial. Data yang muncul berupa status, komentar, tweet, caption, maupun ulasan, membentuk sebuah ekosistem informasi yang beragam (Hou et al., 2021). Menurut laporan

dari *We Are Social* tahun 2023, jumlah pengguna aktif media sosial di dunia telah lebih dari 4,7 miliar orang atau lebih dari setengah populasi global. Setiap harinya, jutaan hingga miliaran posting teks diunggah dan disebarkan melalui berbagai platform, yang menjadikan media sosial sebagai salah satu sumber data terbesar dan paling dinamis di era big data (Al-Zobbi et al., 2018).

Jumlah data teks yang besar dari media sosial menjadi potensi yang sangat berharga, informasi yang terkandung di dalamnya dapat dimanfaatkan untuk berbagai kepentingan, seperti analisis tren pasar, evaluasi kebijakan publik, deteksi isu sosial, strategi pemasaran, hingga pengembangan produk (Bazzaz Abkenar et al., 2021). Di sisi lain, besarnya volume dan keragaman data teks juga dapat menimbulkan tantangan yaitu bagaimana mengekstrak informasi opini publik yang tersirat di balik teks-teks tersebut (Ghani et al., 2019).

Analisis sentimen merupakan cabang dari pengolahan bahasa alami Natural Language Processing (NLP) yang berfokus pada upaya mengidentifikasi, mengekstrak, dan mengklasifikasikan pendapat yang terkandung dalam sebuah teks (Sharma et al., 2025). Analisis sentimen bertujuan untuk mengkategorikan teks ke dalam label tertentu, seperti positif, negatif, atau netral (Mao et al., 2024). Penerapan analisis sentimen sangat luas, dalam dunia bisnis, perusahaan dapat memanfaatkannya untuk menilai kepuasan pelanggan, memahami persepsi konsumen terhadap produk, atau merespons keluhan secara cepat (Kumar et al., 2025). Dalam ranah politik, analisis sentimen berguna untuk mengukur opini publik terhadap kandidat, kebijakan, atau isu tertentu, sementara bidang kesehatan, analisis sentimen dapat digunakan untuk memantau persepsi masyarakat terhadap vaksinasi, gaya hidup sehat, atau pelayanan medis (Raghupathi et al., 2020).

Namun demikian, analisis sentimen bukanlah tugas yang sederhana. Kompleksitas bahasa alami menjadi salah satu faktor utama yang mempersulit proses klasifikasi sentimen. Bahasa teks memiliki ambiguitas makna yang sangat tinggi yang sering dijumpai dalam percakapan sehari-hari, khususnya di media sosial (Yadav et al., 2021). Pada tahap awal, analisis sentimen banyak dilakukan dengan menggunakan metode *machine learning* seperti Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, dan k-Nearest Neighbor (k-NN). Metode-metode ini bergantung pada representasi teks berbasis fitur sederhana, seperti bag-of-words atau term frequency-inverse document frequency (TF-IDF). Meskipun metode tersebut cukup berhasil dalam beberapa kasus, pendekatan ini memiliki keterbatasan karena kurang mampu hubungan semantik antar kata dan konteks dalam kalimat.

Seiring dengan berkembang teknologi kecerdasan buatan, khususnya deep learning, pendekatan baru dengan kinerja lebih baik dalam pengolahan bahasa alami. Salah satu arsitektur yang terbukti efektif adalah Convolutional Neural Network (CNN). CNN dapat diterapkan dalam pengolahan teks karena kemampuannya dalam mengekstraksi fitur penting dari data sekuensial dan CNN bekerja dengan memanfaatkan filter atau kernel yang bergerak melintasi representasi teks

untuk menangkap pola lokal seperti n-gram yang bermakna. Hal ini membuat CNN mampu mengidentifikasi kombinasi kata yang penting dalam menentukan sentimen. Sebagai contoh, frase "tidak puas" atau "sangat baik" dapat terdeteksi sebagai pola yang berkontribusi pada klasifikasi sentimen meskipun konteks kalimat berbeda. Dibandingkan dengan metode lain, CNN lebih unggul karena dapat menangkap fitur kontekstual sekaligus mengurangi kebutuhan akan rekayasa fitur manual.

Meskipun CNN memiliki keunggulan, efektivitas penerapannya dalam analisis sentimen teks media sosial tidaklah absolut. Beberapa faktor teknis dapat memengaruhi performa CNN, seperti Proses pra-pemrosesan (preprocessing) termasuk normalisasi teks, penghapusan kata tidak penting (*stopwords*), penanganan emotikon, atau deteksi bahasa campuran. Jenis representasi kata (word embedding) apakah menggunakan *Word2Vec*, *GloVe*, atau representasi berbasis *contextual embeddings* seperti BERT. Parameter pelatihan (hyperparameter tuning) termasuk ukuran filter, jumlah lapisan konvolusi, ukuran batch, serta jumlah epoch. Kualitas dan kuantitas data latih karena CNN membutuhkan data dalam jumlah besar agar mampu mempelajari pola secara optimal.

Dalam pembelajaran terawasi (supervised learning) berperan penting karena CNN dilatih berdasarkan pasangan input-output, yakni teks dengan label sentimen yang sudah ditentukan. Melalui pendekatan ini, model dapat belajar membedakan pola teks positif, negatif, maupun netral. Namun, keberhasilan supervised learning sangat bergantung pada kualitas data berlabel yang digunakan. Jika data pelatihan tidak representatif, maka hasil evaluasi CNN juga tidak akan akurat.

Beberapa penelitian telah menemukan metode CNN dapat meningkatkan akurasi analisis sentimen dibandingkan metode lain. Namun, hasil tersebut tidak selalu konsisten. Oleh sebab itu, diperlukan evaluasi sistematis mengenai bagaimana CNN dapat melakukan analisis sentimen teks media sosial, khususnya dengan kerangka supervised learning. Evaluasi tersebut penting tidak hanya untuk mengetahui seberapa baik akurasi CNN, tetapi juga untuk memahami faktor-faktor apa saja yang memengaruhi kinerjanya.

Berdasarkan latar belakang yang dijelaskan, dapat disimpulkan bahwa analisis sentimen teks media sosial merupakan tantangan sekaligus peluang di era big data. CNN sebagai salah satu metode deep learning dapat digunakan untuk meningkatkan akurasi analisis sentimen. Namun, untuk mengetahui apakah akurasi metode CNN dapat digunakan, maka diperlukan evaluasi terhadap kinerja CNN dengan mempertimbangkan aspek preprocessing, representasi kata, parameter pelatihan, dan perbandingan dengan metode lain. Hasil evaluasi ini diharapkan dapat memberikan pengembangan sistem analisis sentimen yang lebih akurat, efisien.

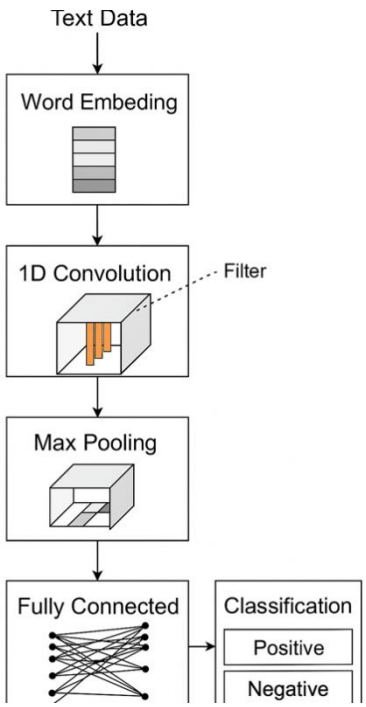
METODE PENELITIAN

Penelitian ini menggunakan arsitektur Convolutional Neural Network (CNN) dengan pendekatan 1D Convolution Layer yang dirancang untuk memproses data berurutan satu dimensi. Data yang digunakan merupakan teks hasil word embedding yang direpresentasikan dalam bentuk deretan vektor numerik yang menggambarkan hubungan semantik antar kata.

Lapisan 1D convolution digunakan untuk mengenali pola dalam urutan kata. Filter pada lapisan konvolusi dengan urutan teks untuk mengekstraksi fitur penting yang mewakili sentimen. Setiap filter memiliki peran dalam mendeteksi pola sentimen positif maupun negatif.

Selanjutnya, lapisan max pooling digunakan untuk mereduksi dimensi fitur dengan memilih nilai maksimum dari hasil konvolusi, sehingga hanya fitur paling representatif yang dipertahankan. Hasil dari proses konvolusi dan pooling kemudian diteruskan ke lapisan fully connected untuk melakukan klasifikasi akhir menggunakan fungsi aktivasi Softmax. Fungsi ini menghasilkan probabilitas dari masing-masing kelas, yaitu positif dan negatif.

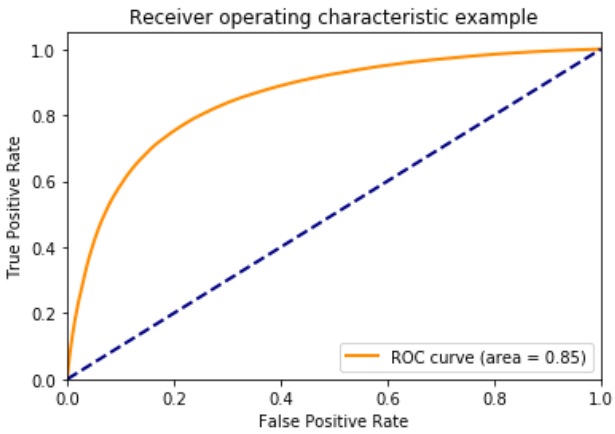
Untuk melatih model, dataset teks dibagi menjadi data latih dan data uji. Proses pelatihan dilakukan dengan menggunakan optimizer Adam dan fungsi loss kategori categorical cross-entropy. Evaluasi performa model dilakukan menggunakan AUC-ROC curve dan Confusion Matrix untuk mengukur kemampuan model dalam mengklasifikasikan sentimen secara akurat.



Gambar 1 Diagram arsitektur CNN

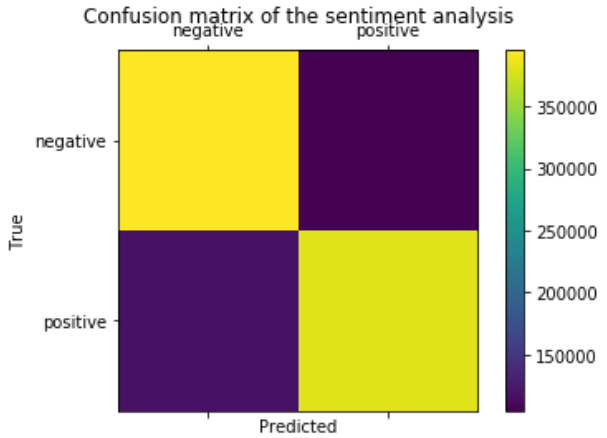
HASIL DAN PEMBAHASAN

Pada penelitian ini arsitektur Convolutional Neural Network (CNN) yang digunakan dengan pendekatan 1D convolution layer dan dirancang untuk memproses data berurutan satu dimensi. Representasi teks hasil word embedding berupa deretan vektor numerik yang menggambarkan semantik antar kata. Oleh karena itu, penggunaan 1D convolution menjadi tepat karena lapisan ini dapat mengenali pola dalam urutan kata. Lapisan konvolusi satu dimensi bekerja dengan cara menggeser filter di sepanjang urutan teks untuk mengekstraksi fitur penting dari setiap kombinasi kata. Setiap filter bertugas mendeteksi pola positif atau negative sedangkan lapisan max pooling memilih nilai tertinggi dari hasil konvolusi untuk mereduksi dimensi dan mempertahankan fitur paling representatif. Hasil dari proses ini kemudian diteruskan ke lapisan fully connected untuk klasifikasi akhir menggunakan fungsi aktivasi softmax yang menentukan probabilitas kelas sentimen positif dan negatif.



Gambar 2 Hasil performa klasifikasi CNN Kurva AUC-ROC

Berdasarkan Gambar 2 model CNN menunjukkan hasil klasifikasi yang cukup baik pada data teks media sosial. Hasil tersebut dapat dilihat dari Kurva AUC-ROC dimana area di bawah kurva (AUC) yang mendekati 1 menunjukkan model dapat membedakan kelas positif dan negatif secara konsisten.



Gambar 3 Hasil performa klasifikasi CNN Confusion Matrix

Dari gambar 3 dapat disimpulkan bawah confusion matrix menunjukkan model mampu mengklasifikasikan sebagian besar data secara benar, dengan tingkat kesalahan yang relatif kecil pada prediksi false negative.

Hasil penelitian ini dapat disimpulkan bawah CNN dengan 1D convolution layer terbukti efektif untuk menganalisis teks berurutan satu dimensi. Model ini memberikan alternatif yang kuat untuk analisis sentimen berbasis deep learning, khususnya dalam menangani data tidak terstruktur dan penuh variasi seperti teks media sosial. Namun, peningkatan lebih lanjut dapat dilakukan dengan mengoptimalkan representasi embedding kontekstual serta memperdalam arsitektur CNN untuk menghasilkan pola semantik yang lebih kompleks.

KESIMPULAN

Berdasarkan hasil penelitian dan evaluasi yang telah dilakukan, disimpulkan bahwa penerapan Convolutional Neural Network (CNN) dengan pendekatan 1D convolution layer terbukti efektif dalam melakukan analisis sentimen pada teks media sosial. Model CNN mampu mengekstraksi pola semantik dari urutan kata melalui proses konvolusi dan pooling, sehingga dapat mengenali perbedaan antara sentimen positif dan negatif dengan baik. Hasil pengujian menggunakan kurva AUC-ROC menunjukkan nilai AUC yang mendekati 1, menandakan kemampuan model dalam membedakan kelas secara konsisten. Selain itu, Confusion Matrix memperlihatkan tingkat akurasi yang tinggi dengan kesalahan prediksi yang relatif kecil, terutama pada kelas false negative. Hal ini membuktikan bahwa CNN memiliki performa yang andal dalam mengolah data teks yang tidak terstruktur dan penuh variasi, seperti unggahan media sosial.

UCAPAN TERIMA KASIH

Terima kasih kepada Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) Universitas Lambung Mangkurat yang telah memberikan dana penelitian tahun 2025 dengan no kontrak 1910/UN8/LT/2025.

DAFTAR PUSTAKA

Al-Zobbi, M., Shahrestani, S., & Ruan, C. (2018). Experimenting sensitivity-based anonymization framework in apache spark. *Journal of Big Data*, *5*(1), 38. https://doi.org/10.1186/s40537-018-0149-0

Bazzaz Abkenar, S., Haghi Kashani, M., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics*, *57*, 101517. https://doi.org/10.1016/j.tele.2020.101517

Ghani, N. A., Hamid, S., Targio Hashem, I. A., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, *101*, 417–428. https://doi.org/10.1016/j.chb.2018.08.039

Grover, P., Kar, A. K., & Dwivedi, Y. (2022). The evolution of social media influence—A literature review and research agenda. *International Journal of Information Management Data Insights*, *2*(2), 100116. https://doi.org/10.1016/j.jjimei.2022.100116

Hou, Q., Han, M., Qu, F., & He, J. S. (2021). Understanding social media beyond text: A reliable practice on Twitter. *Computational Social Networks*, *8*(1), 4. https://doi.org/10.1186/s40649-021-00088-x

Kumar, A., Bhatia, A., Kumar, S., Kumar, R., & Yadav, V. (2025). Sentimental analysis using natural language processing. *AIP Conference Proceedings*, *3224*(1), 020054. https://doi.org/10.1063/5.0246883

Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University - Computer and Information Sciences*, *36*(4), 102048. https://doi.org/10.1016/j.jksuci.2024.102048

Raghupathi, V., Ren, J., & Raghupathi, W. (2020). Studying Public Perception about Vaccination: A Sentiment Analysis of Tweets. *International Journal of Environmental Research and Public Health*, *17*(10), 3464. https://doi.org/10.3390/ijerph17103464

Sharma, N. A., Ali, A. B. M. S., & Kabir, M. A. (2025). A review of sentiment analysis: Tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics*, *19*(3), 351–388. https://doi.org/10.1007/s41060-024-00594-x

Wu, M.-Y. (2025). Examining the impacts of information and communication technology (ICT) on national development and wellbeing: A global perspective. *Journal of Economy and Technology*, *3*, 190–201. https://doi.org/10.1016/j.ject.2024.11.006

Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, *3*, 100073. https://doi.org/10.1016/j.dajour.2022.100073

Yadav, A., Patel, A., & Shah, M. (2021). A comprehensive review on resolving ambiguities in natural language processing. *AI Open*, *2*, 85–92. https://doi.org/10.1016/j.aiopen.2021.05.001