

Analisis Sentimen Pada ChatGPT Menggunakan Algoritma *Long Short Term Memory (LSTM)*

Sentiment Analysis on ChatGPT Using Long Short-Term Memory (LSTM) Algorithm

Sugeng Winardi^{1*}, Ajie W. Soejono², Dini Sari³

^{1,2,3}Program Studi Sistem Informasi, Universitas Respati Yogyakarta

*¹sugengw@respatu.ac.id, ²aji.ws@gmail.com, ³dinisari@gmail.com

Abstrak

Penelitian ini fokus pada analisis sentimen menggunakan ChatGPT dengan memanfaatkan metode LSTM dan memanfaatkan dataset dari Kaggle. Tujuan utama dari penelitian ini adalah untuk meningkatkan kinerja analisis sentimen teks dengan mengadopsi model LSTM yang lebih canggih dibandingkan dengan model berbasis aturan yang umumnya digunakan. Metode LSTM dipilih sebagai pendekatan untuk menganalisis sentimen pada dataset teks yang telah diolah dan diberi label sentimen. Tahapan penelitian dimulai dengan proses pengumpulan dan persiapan data, diikuti dengan pembuatan model rule-based. Selanjutnya, dilakukan pemrosesan dan pre-training metode LSTM, featurization, training metode LSTM, evaluasi model, dan analisis hasil. Pada tahap evaluasi, kinerja Metode LSTM akan dinilai menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Hasil evaluasi akan dibandingkan dengan model rule-based yang telah dibuat sebelumnya. Harapannya, hasil dari penelitian ini dapat memberikan wawasan baru terkait penggunaan Metode LSTM dalam analisis sentimen teks, sekaligus meningkatkan kinerja sentimen analisis secara keseluruhan.

Kata kunci: Sentimen Analisis; Analisis ChatGPT; Metode LSTM

Abstract

This research focuses on sentiment analysis using ChatGPT, leveraging the LSTM method and utilizing datasets from Kaggle. The primary goal of this study is to enhance the performance of text sentiment analysis by adopting more advanced LSTM based models compared to traditional rule-based models. The LSTM method was chosen to analyze sentiment in pre-processed and sentiment-labeled text datasets. The research stages begin with data collection and preparation, followed by the development of a rule-based model. Subsequently, steps include processing and pre-training the LSTM method, featurization, training the LSTM model, model evaluation, and result analysis. During the evaluation phase, the performance of the LSTM method will be assessed using evaluation metrics such as accuracy, precision, recall, and F1-score. The evaluation results will then be compared to the performance of the previously developed rule-based model. The findings of this study are hoped to provide new insights into the use of the LSTM method for text sentiment analysis while simultaneously improving overall sentiment analysis performance.

Keywords: Analysis Sentimen, ChatGPT Analysis, LSTM Metode

1. PENDAHULUAN

Analisis sentimen merupakan salah satu teknik pemrosesan bahasa alami (NLP) yang dapat mengidentifikasi serta mengekstrak sentimen atau perasaan dari teks. Hal ini menjadi krusial untuk memahami persepsi dan tanggapan pengguna terhadap produk atau layanan dalam dunia bisnis. Pada beberapa tahun terakhir, banyak perusahaan mengadopsi sentimen analisis guna meningkatkan kualitas produk dan layanan mereka. Meskipun terdapat berbagai teknik dan algoritma untuk melakukan sentimen analisis, salah satu yang paling populer adalah menggunakan *Long Short Term Memory (LSTM)*. Diperkenalkan oleh Google pada 2017, LSTM telah menjadi standar *de facto* dalam tugas pemrosesan bahasa alami, termasuk sentimen analisis. Contoh kemajuan terbaru dalam sentimen analisis mencakup penelitian menggunakan jaringan *multi-level matching and aggregation* untuk klasifikasi relasi antar teks secara *few-shot learning* (1) dan penggunaan jaringan embedding informasi heterogen pada analisis sentimen aspek (2).

LSTM sebagai model NLP, terbukti sangat akurat dan efektif dalam memproses teks panjang dan kompleks. Model ini mengatasi keterbatasan teknik sentimen analisis rule-based tradisional, seperti kesulitan dalam menangkap ironi atau nuansa bahasa. Penelitian tentang "A Survey on Deep Learning for Sentiment Analysis: Models, Techniques, and Challenges" merangkum kemajuan metode deep learning, termasuk LSTM seperti BERT dan GPT, sebagai metode yang sangat efektif untuk sentimen analisis (3).

Sumber dataset dari Kaggle memberikan kontribusi berharga dalam melakukan sentimen analisis, seringkali dibuat oleh perusahaan atau individu untuk memperbaiki produk atau layanan berdasarkan tanggapan pengguna. Kaggle sebagai platform menyediakan beragam dataset dan tantangan terkait analisis data, termasuk sentimen analisis, serta fitur kolaborasi dan diskusi dengan para ahli data. Penelitian lain juga menegaskan bahwa Kaggle adalah salah satu sumber daya terbaik untuk mengakses dataset dan algoritma machine learning yang digunakan dalam sentimen analisis (4).

Pemanfaatan ChatGPT sebagai alat untuk sentimen analisis diharapkan dapat memberikan hasil yang lebih baik, mengingat model ini merupakan NLP yang kuat dan terlatih dengan data yang sangat besar. Kemampuannya memahami konteks dan nuansa bahasa kompleks membuatnya dapat menghasilkan prediksi sentimen yang akurat dan beragam.

2. METODE/PERANCANGAN/MATERIAL

2.1 Deskripsi Dataset

Dataset yang digunakan untuk penelitian ini bersumber dari Kaggle, platform yang diakui secara luas yang dikenal sebagai hosting beragam kumpulan data untuk pembelajaran mesin dan aplikasi ilmu data. Kumpulan data memiliki 217.623 instans dan merupakan kumpulan signifikan dari tweet yang dibuat oleh ChatGPT, yang menyediakan sumber daya yang kaya untuk mengevaluasi teknik analisis sentimen. Atribut data mencakup *Tweet Content* dan label terkait yang menunjukkan sentimen. Label didefinisikan ke dalam tiga kategori: '*buruk*', terhitung 106.695 contoh '*baik*', mewakili 55.754 kasus; dan '*netral*', terdiri dari 55.174 contoh. ChatGPT, model bahasa canggih yang dikembangkan oleh OpenAI, menghasilkan tweet ini, yang mencerminkan berbagai skenario percakapan. Tweet diberi label berdasarkan algoritma analisis sentimen, dengan asumsi bahwa ChatGPT meniru percakapan mirip manusia dengan cukup akurat untuk mencerminkan sentimen asli di setiap tweet yang dihasilkannya. Dengan demikian, kumpulan data

menyajikan data yang cukup untuk penelitian dan representasi yang seimbang di berbagai kelas sentimen, meningkatkan generalisasi hasil.

2.2 Pre-Processing Data

Sebelum menerapkan algoritme LSTM pada dataset, satu hal penting adalah melakukan serangkaian langkah pemrosesan dan pembersihan data untuk memastikan data berada dalam kondisi optimal untuk analisis. Pertama, entri duplikat dalam kumpulan data diidentifikasi dan dihapus untuk mencegah bias dalam fase pelatihan model. Proses ini penting untuk mempertahankan integritas desain eksperimen kami, karena kejadian berulang secara tidak sengaja dapat memengaruhi pembelajaran model. Selain itu, contoh dengan nilai yang hilang diperiksa. Mengingat bahwa kumpulan data kami hanya terdiri dari data tekstual dan label, nilai yang hilang sangat minim. Namun, *instance* dengan label atau konten tweet yang hilang telah dihapus untuk mempertahankan struktur data yang konsisten. Karena kumpulan data bersifat tekstual dan kategorikal, konsep outlier tidak berlaku seperti pada data numerik.

Namun, selanjutnya meninjau kembali data untuk mengidentifikasi kejadian apa pun yang mungkin berada di luar kisaran yang diharapkan, seperti tweet yang terlalu panjang atau berisi karakter yang tidak standar. Entri ini ditangani dengan tepat untuk memastikan mereka tidak menyimpang dari proses pembelajaran model. Melalui langkah-langkah prapemrosesan ini, bertujuan untuk membuat kumpulan data yang kuat dan bersih yang secara akurat mencerminkan kinerja algoritme LSTM dalam analisis sentimen.

2.3 Data Split

Untuk memastikan kekokohan model kami dan memfasilitasi evaluasi yang adil atas kinerjanya, kami menerapkan pemisahan stratifikasi dataset ke dalam set pelatihan dan pengujian. Stratifikasi memastikan bahwa distribusi kelas sentimen (baik, buruk, dan netral) di pemisahan kami mencerminkan distribusi di kumpulan data asli, sehingga menghindari potensi bias. Dataset dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Rasio pemisahan ini adalah praktik yang diterima secara umum dalam pembelajaran mesin dan sering dipilih karena memberikan data yang cukup untuk melatih model sambil tetap menyisakan porsi yang memadai untuk pengujian. Set pelatihan, oleh karena itu, terdiri dari 85.356 contoh buruk, 44.604 contoh baik, dan 44.140 netral. Set pengujian terdiri dari 21.339 contoh buruk, 11.150 baik, dan 11.034 netral. Tahap ini memastikan bahwa model dilatih pada sebagian besar data, meningkatkan kemampuannya untuk menggeneralisasi sambil mempertahankan sejumlah besar contoh untuk evaluasi kinerjanya yang tidak bias pada data yang tidak terlihat.

2.4. Algoritma LSTM

Algoritma *Long Short-Term Memory (LSTM)* adalah jenis dari jaringan saraf rekuren (RNN) yang dirancang khusus untuk mengatasi masalah vanishing gradient dan long-term dependencies dalam pemrosesan urutan data. LSTM memungkinkan informasi yang relevan dari urutan masukan sebelumnya disimpan dan diakses dalam proses prediksi.

Berikut adalah beberapa konsep matematis yang mendasari algoritma LSTM:

a. *Input Gates (Forget, Input, Output):*

LSTM memiliki tiga jenis gerbang yang mengatur aliran informasi: *forget gate*, *input gate*, dan *output gate*. Setiap gerbang memiliki vektor bobot dan bias yang mengontrol berapa banyak

informasi yang akan dilewatkan.

b. *State Cell (C_t):*

State cell adalah komponen inti dari LSTM yang menyimpan informasi konteks dalam urutan data. Ini diatur oleh input dan forget gate.

c. *Hidden State (h_t):*

Hidden state adalah keluaran dari LSTM di setiap langkah waktu dan mengandung informasi yang relevan untuk langkah waktu tersebut.

d. *Sigmoid dan Tanh Aktivasi:*

LSTM menggunakan fungsi aktivasi sigmoid untuk mengatur gate (0 atau 1) dan fungsi aktivasi tanh untuk mengatur nilai yang akan disimpan dalam *state cell*.

Berikut adalah persamaan terkait

Forget Gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \dots \dots \dots 1)$$

Input Gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \dots \dots \dots 2)$$

Candidate Cell:

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \dots \dots \dots 3)$$

Update State Cell:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \dots \dots \dots 4)$$

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \dots \dots \dots 5)$$

Hidden State:

$$h_t = o_t \cdot \tanh(C_t) \dots \dots \dots 6)$$

Di sini, (x_t) adalah input pada langkah waktu (t) , (h_{t-1}) adalah hidden state dari langkah waktu sebelumnya, (σ) adalah fungsi sigmoid, (\tanh) adalah fungsi tangen hiperbolik, dan (W) dan (b) mewakili matriks bobot dan bias yang perlu diestimasi selama pelatihan.

3. HASIL DAN PEMBAHASAN

3.1 Kinerja Model

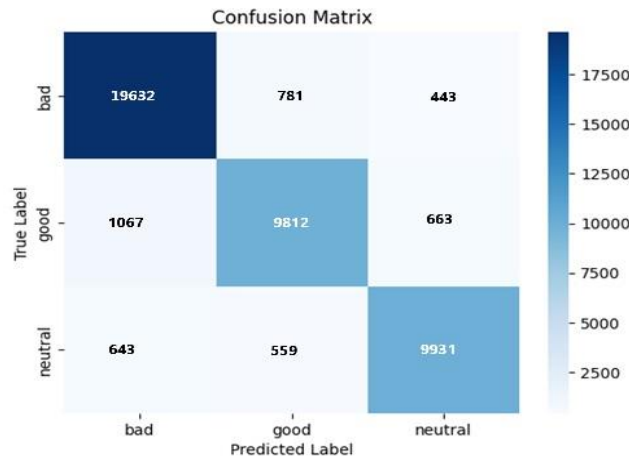
Pada Confusion Matrix, seperti yang ditunjukkan pada Gambar 1, memberikan gambaran rinci tentang kinerja model LSTM dalam mengklasifikasikan sentimen tweet ChatGPT. Ini terdiri dari tiga kelas sentimen: "Buruk", "Baik", dan "Netral", yang direpresentasikan dalam baris dan kolom.

Entri diagonal matriks, 19632 untuk "Buruk", 9812 untuk "Baik", dan 9931 untuk "Netral", menunjukkan tingkat positif yang sebenarnya, yang menunjukkan jumlah contoh di mana model mengidentifikasi sentimen ini dengan benar. Angka-angka ini cukup mengesankan, menunjukkan bahwa model tersebut dapat secara akurat mengklasifikasikan sentimen tweet di sebagian besar waktu.

Namun, ada juga entri diagonal bukan nol, yang mewakili kejadian di mana prediksi model tidak akurat. Misalnya, model salah memprediksi 781 tweet "Baik" dan 443 "Netral" sebagai "Buruk", sementara 1.067 tweet "Buruk" dan 663 "Netral" salah diklasifikasikan sebagai "Baik". Demikian pula, 643 tweet "Buruk" dan 559 "Baik" salah diprediksi sebagai "Netral".

Meskipun kesalahan klasifikasi ini secara signifikan lebih sedikit dari prediksi yang benar, mereka mengungkap tantangan tertentu yang membedakan antara kelas sentimen yang berbeda.

Secara khusus, mereka menyoroti sedikit kesulitan model dalam memprediksi sentimen "Baik" dan "Netral" secara akurat dibandingkan dengan sentimen "Buruk". Memahami dan mengatasi kesalahan klasifikasi ini bisa menjadi arahan yang bermanfaat untuk penyempurnaan dan penelitian model di masa depan.



Gambar 1. Confusion Matrix

Laporan klasifikasi, seperti yang ditunjukkan pada Tabel 3.1, memberikan tampilan yang lebih detail tentang performa model LSTM di seluruh kelas sentimen yang berbeda: "Buruk", "Baik", dan "Netral". Melihat skor presisi, model memiliki tingkat ketepatan yang tinggi di semua kelas sentimen dengan 0,93 untuk "Buruk", 0,89 untuk "Baik", dan 0,91 untuk "Netral". Skor ini menunjukkan bahwa sangat mungkin benar ketika model memprediksi sentimen tertentu.

Skor ingatan, yaitu 0,95 untuk "Buruk", 0,86 untuk "Baik", dan 0,90 untuk "Netral", memberikan wawasan tentang kelengkapan model. Kelas "Buruk" memiliki ingatan tertinggi, artinya model mengidentifikasi 94% tweet "Buruk" dengan benar. Namun, model tersebut sedikit kurang berhasil dalam mengidentifikasi tweet "Baik" dan "Netral", dengan penarikan kembali masing-masing 0,85 dan 0,89.

Skor F1, yang merupakan rata-rata harmonik dari presisi dan daya ingat, adalah ukuran performa seimbang model. Model ini memiliki skor F1 yang cukup tinggi, masing-masing 0,94, 0,88, dan 0,91 untuk "Buruk", "Baik", dan "Netral", yang menunjukkan keseimbangan yang baik antara presisi dan daya ingat.

Keakuratan keseluruhan model adalah 0,91, yang menunjukkan bahwa model memprediksi sentimen tweet dengan benar sebanyak 90% dari keseluruhan waktu. Skor rata-rata makro dan rata-rata tertimbang di seluruh presisi, daya ingat, dan skor F1 juga sejajar di sekitar 0,91, menunjukkan performa yang konsisten di seluruh kelas sentimen yang berbeda.

Tabel 1. Classification Report

| | precision | recall | f1-score |
|--------------|-----------|--------|----------|
| bad | 0,93 | 0,95 | 0,94 |
| good | 0,89 | 0,86 | 0,88 |
| neutral | 0,91 | 0,9 | 0,91 |
| accuracy | | | 0,91 |
| macro avg | 0,91 | 0,9 | 0,91 |
| weighted avg | 0,91 | 0,91 | 0,91 |

Hasil ini menunjukkan keefektifan model dalam klasifikasi sentimen. Namun, mereka menyoroti area di mana model dapat ditingkatkan, khususnya dalam memprediksi tweet "Baik" dan "Netral". Hal ini menggarisbawahi perlunya penyempurnaan dan penelitian lebih lanjut untuk mengoptimalkan performa model di semua kelas sentimen.

3.2 Analisis Spesifik Kelas

Dalam hal performa kelas individu, model LSTM menunjukkan keefektifan yang bervariasi. Kelas sentimen "Buruk" memiliki presisi, daya ingat, dan skor F1 tertinggi, yang menunjukkan bahwa model paling mahir dalam memprediksi kelas ini secara akurat. Dengan presisi 0,93, ini menyiratkan bahwa ketika model memperkirakan tweet memiliki sentimen "Buruk", 93% dari waktu itu benar. Selain itu, penarikan 0,95 menunjukkan bahwa dari semua tweet "Buruk" yang sebenarnya, model mengidentifikasi 95% di antaranya dengan benar. Statistik ini secara kolektif berkontribusi pada skor F1 yang tinggi sebesar 0,94, menunjukkan kekuatan prediksi yang seimbang untuk kelas sentimen "Buruk".

Kelas "Baik" dan "Netral" menghadirkan lebih banyak tantangan untuk model. Meskipun metrik presisi dan daya ingat untuk kelas-kelas ini masih relatif tinggi (0,89 dan 0,91 untuk "Baik"; 0,91 dan 0,9 untuk "Netral"), skor F1 mereka lebih rendah daripada kelas "Buruk", 0,86 dan 0,88, masing-masing. Ini menunjukkan bahwa model memiliki waktu yang sedikit lebih sulit untuk mengklasifikasikan kelas sentimen ini dengan benar. Hal ini bisa disebabkan oleh berbagai faktor. Misalnya, bahasa yang digunakan dalam tweet "Baik" dan "Netral" mungkin lebih bernuansa atau ambigu, atau mungkin ada masalah ketidakseimbangan kelas dalam kumpulan data pelatihan yang membuat model lebih selaras dengan sentimen "Buruk". Temuan ini menggarisbawahi pentingnya menyelidiki lebih lanjut dan mengatasi tantangan ini untuk meningkatkan kinerja model di semua kelas sentimen.

Confusion matrix secara visual merepresentasikan metrik performa ini, memungkinkan pemahaman intuitif tentang seberapa sering model membuat prediksi yang benar dan salah untuk setiap kelas sentimen. Misalnya, model mengklasifikasikan dengan benar 19.631 tweet "Buruk", tetapi salah mengklasifikasikan 780 sebagai "Baik" dan 442 sebagai "Netral". Demikian pula, 9812 tweet "Baik" diidentifikasi dengan benar, dengan 1066 dan 662 salah diklasifikasikan sebagai "Buruk" dan "Netral". Terakhir, model dengan benar memberi label 9930 tweet "Netral", sementara 642 dan 558 salah diberi label sebagai "Buruk" dan "Baik". Alat visual ini memungkinkan kita untuk dengan mudah menemukan di mana model unggul dan di mana membuat kesalahan, sehingga menginformasikan strategi potensial untuk perbaikan.

3.3. Interpretasi dan Implikasi

Hasil penelitian ini memiliki implikasi yang signifikan untuk bidang analisis sentimen, khususnya terkait data media sosial. Mencapai akurasi keseluruhan sebesar 90% dengan model LSTM menandakan tingkat kinerja yang tinggi dalam klasifikasi sentimen, sebuah pencapaian yang berkontribusi pada semakin banyaknya bukti yang mendukung metode berbasis pembelajaran mendalam untuk tugas ini. Yang penting, evaluasi mendetail tentang performa model di tiga kelas sentimen - "Baik", "Buruk", dan "Netral" - membantu mengidentifikasi area spesifik di mana model unggul dan di mana perbaikan lebih lanjut dapat dilakukan. Misalnya, skor F1 yang relatif lebih rendah untuk kelas "Baik" dan "Netral" menunjukkan bahwa penelitian lebih lanjut dan penyempurnaan model dapat menguntungkan area ini. Analisis kinerja bernuansa ini sangat penting

untuk mengembangkan model analisis sentimen yang lebih kuat dan serbaguna yang menangani berbagai data dan konteks.

Dari perspektif penerapan AI dalam analisis media sosial, temuan penelitian ini menggarisbawahi potensi model berbasis Transformer untuk klasifikasi sentimen yang efektif. Mengingat banyaknya konten buatan pengguna di platform media sosial, memiliki alat yang akurat dan efisien seperti model LSTM bisa sangat berharga. Misalnya, bisnis dapat menggunakan alat ini untuk memantau sentimen publik tentang produk atau layanan mereka, pembuat kebijakan dapat mengukur opini publik tentang berbagai masalah, dan peneliti dapat melacak tren dan pola masyarakat. Selain itu, penelitian ini menunjukkan bahwa model AI seperti ChatGPT dapat menjadi sumber data yang berharga untuk analisis sentimen, yang semakin memperluas cakupan aplikasi potensial dalam analisis media sosial. Namun, ini juga menekankan perlunya mempertimbangkan potensi tantangan dan bias dalam data yang dihasilkan AI, menggarisbawahi pentingnya transparansi dan pertimbangan etis dalam menggunakan model tersebut.

3.4. Tantangan dan Hipotesis

Sepanjang perjalanan studi ini, beberapa tantangan dan hasil tak terduga ditemui. Rintangan yang signifikan adalah berurusan dengan kompleksitas dan nuansa yang melekat pada bahasa manusia dalam tweet. Analisis sentimen bukanlah tugas yang mudah karena memerlukan pemahaman konteks, mendeteksi sarkasme, dan mengidentifikasi kehalusan dalam bahasa, yang dapat menjadi tantangan bagi model AI. Kompleksitas ini terutama terlihat pada kinerja model yang sedikit lebih rendah dalam memprediksi sentimen "Baik" dan "Netral", kemungkinan karena ambiguitas dan kehalusan yang lebih besar di kelas ini dibandingkan dengan bahasa yang lebih eksplisit yang biasanya dikaitkan dengan sentimen "Buruk".

Tantangan lain yang dihadapi terkait dengan kualitas data yang dihasilkan AI. Meskipun ChatGPT menjadi model bahasa yang kuat, ada beberapa contoh di mana tweet yang dihasilkan agak tidak masuk akal atau tidak memiliki sentimen yang jelas, yang mungkin menyebabkan gangguan pada data dan memengaruhi kinerja model. Selain itu, terlepas dari langkah-langkah pra-pemrosesan data kami yang ketat, mungkin ada beberapa bias atau ketidaksesuaian yang tidak terduga dalam data, yang dapat memengaruhi hasil.

Terakhir, memastikan ketangguhan dan generalisasi model LSTM merupakan sebuah tantangan. Sementara model menunjukkan kinerja yang kuat dalam rangkaian pengujian khusus kami, selalu ada kekhawatiran tentang bagaimana kinerjanya dengan distribusi data yang berbeda atau ketika berhadapan dengan data baru yang tidak terlihat. Terlepas dari tantangan ini, hasil dari penelitian ini memberikan wawasan dan panduan yang berharga untuk perbaikan lebih lanjut dalam menerapkan algoritma LSTM untuk tugas analisis sentimen.

4. KESIMPULAN DAN SARAN

Penelitian ini telah meneliti penggunaan model LSTM dalam analisis sentimen pada tweet yang dihasilkan oleh ChatGPT. Hasil penelitian menunjukkan bahwa model LSTM mampu mencapai akurasi tinggi dalam mengelompokkan tweet ke dalam kategori sentimen "Baik", "Buruk", dan "Netral". Namun, analisis mendalam menunjukkan bahwa model mengalami kesulitan dalam memprediksi secara akurat tweet dengan sentimen "Baik" dan "Netral", sehingga diperlukan penelitian lebih lanjut dan peningkatan model. Studi ini memberikan wawasan penting dalam analisis sentimen, khususnya dalam konteks data media sosial yang dihasilkan oleh AI. Selain itu, penelitian

ini menekankan potensi besar model pembelajaran mendalam seperti LSTM dalam menyelesaikan tugas pemrosesan bahasa alami yang kompleks, sekaligus menyoroti pentingnya evaluasi yang menyeluruh dan pengembangan berkelanjutan. Temuan ini diharapkan dapat mendorong perkembangan lebih lanjut dalam penerapan algoritme LSTM untuk analisis sentimen dan memperluas aplikasinya dalam analisis media sosial berbasis AI.

Berdasarkan temuan penelitian ini, ada beberapa arah yang dapat diambil untuk penelitian di masa depan. Pertama, penelitian lebih lanjut dapat difokuskan pada peningkatan dan pengoptimalan kinerja model LSTM dalam klasifikasi sentimen, terutama untuk kelas "Baik" dan "Netral." Upaya ini dapat mencakup penggunaan kumpulan data pelatihan yang lebih besar atau lebih beragam, integrasi fitur linguistik tambahan ke dalam model, atau penerapan metode seperti penyeimbangan kelas dan augmentasi data untuk mengatasi bias atau ketidakseimbangan dalam data pelatihan. Selain itu, meskipun model LSTM menunjukkan performa yang baik dalam penelitian ini, perbandingan kinerjanya dengan model pembelajaran mendalam canggih lainnya, seperti BERT atau GPT-3, dapat memberikan wawasan yang lebih mendalam. Potensi aplikasi model LSTM juga dapat diperluas ke domain pemrosesan bahasa alami lainnya, seperti peringkasan teks, terjemahan mesin, atau sistem tanya jawab. Penelitian mendatang juga dapat mengeksplorasi dampak etis dan sosial dari penggunaan model AI seperti ChatGPT dalam menghasilkan konten media sosial. Penelitian ini dapat memberikan wawasan penting tentang bagaimana teknologi tersebut dapat digunakan secara bertanggung jawab dan transparan dalam analisis sentimen maupun aplikasi AI lainnya.

DAFTAR PUSTAKA

- (1) Ahmed, S. A., MonaLisa, Hussain, M., & Khan, Z. U. (2022). Supervised machine learning for predicting shear sonic log (DTS) and volumes of petrophysical and elastic attributes, Kadanwari Gas Field, Pakistan. *Frontiers in Earth Science*, 10. <https://www.frontiersin.org/articles/10.3389/feart.2022.919130>
- (2) Alomari, E., Katib, I., Albeshri, A., & Mehmood, R. (2021). COVID-19: Detecting Government Pandemic Measures and Public Concerns from Twitter Arabic Data Using Distributed Machine Learning. *International Journal of Environmental Research and Public Health*, 18(1). <https://doi.org/10.3390/ijerph18010282>
- (3) Borgman, J., Stark, K., Carson, J., & Hauser, L. (2022). Deep Learning Encoding for Rapid Sequence Identification on Microbiome Data. *Frontiers in Bioinformatics*, 2. <https://www.frontiersin.org/articles/10.3389/fbinf.2022.871256>
- (4) Ganesh, P., Chen, Y., Lou, X., Khan, M. A., Yang, Y., Sajjad, H., Nakov, P., Chen, D., & Winslett, M. (2021). Compressing Large-Scale Transformer-Based Models: A Case Study on BERT. *Transactions of the Association for Computational Linguistics*, 9, 1061–1080. https://doi.org/10.1162/tacl_a_00413
- (5) Hertina, H., Nurwahid, M., Haswir, H., Sayuti, H., Darwis, A., Rahman, M., Yendra, R., & Hamzah, M. L. (2021). Data mining applied about polygamy using sentiment analysis on Twitters in Indonesian perception. *Bulletin of Electrical Engineering and Informatics*, 10(4), 2231–2236. <https://doi.org/10.11591/eei.v10i4.2325>
- (6) J. Zheng & L. Zheng. (2019). A Hybrid Bidirectional Recurrent Convolutional Neural Network Attention-Based Model for Text Classification. *IEEE Access*, 7, 106673–106685. <https://doi.org/10.1109/ACCESS.2019.2932619>

- (7) Kim, A. Y., Ha, J. G., Choi, H., & Moon, H. (2018). Automated Text Analysis Based on Skip-Gram Model for Food Evaluation in Predicting Consumer Acceptance. *Computational Intelligence and Neuroscience*, 2018, 9293437. <https://doi.org/10.1155/2018/9293437>
- (8) Kulkarni, A., Sthapit, A., Sedhain, A., Bhattarai, B., & Panthee, S. (2021). Texture Classification using Angular and Radial Bins in Transformed Domain. *International Journal of Advanced Computer Science and Applications*, 12(3). <https://doi.org/10.14569/IJACSA.2021.0120301>
- (9) Mudinas, A., Zhang, D., & Levene, M. (2018). Bootstrap Domain-Specific Sentiment Classifiers from Unlabeled Corpora. *Transactions of the Association for Computational Linguistics*, 6, 269–285. https://doi.org/10.1162/tacl_a_00020
- (10) Picozzi, M., & Iaccarino, A. G. (2021). Forecasting the Preparatory Phase of Induced Earthquakes by Recurrent Neural Network. *Forecasting*, 3(1), 17–36. <https://doi.org/10.3390/forecast3010002>
- (11) Pilař, L., Kvasničková Stanislavská, L., & Kvasnička, R. (2021). Healthy Food on the Twitter Social Network: Vegan, Homemade, and Organic Food. *International Journal of Environmental Research and Public Health*, 18(7). <https://doi.org/10.3390/ijerph18073815>
- (12) Preto, A. J., Matos-Filipe, P., Mourão, J., & Moreira, I. S. (2022). SYNPREDE: prediction of drug combination effects in cancer using different synergy metrics and ensemble learning. *GigaScience*, 11, giac087. <https://doi.org/10.1093/gigascience/giac087>
- (13) Shafique, M. A., & Marchán, S. S. (2022). Investigating the Impact of Information Sharing in Human Activity Recognition. *Sensors*, 22(6). <https://doi.org/10.3390/s22062280>
- (14) W. Li & B. Xu. (2020). Aspect-Based Fashion Recommendation With Attention Mechanism. *IEEE Access*, 8, 141814–141823. <https://doi.org/10.1109/ACCESS.2020.3013639>
- (15) Wang, Z., & Wu, Q. (2018). An Integrated Deep Generative Model for Text Classification and Generation. *Mathematical Problems in Engineering*, 2018, 7529286. <https://doi.org/10.1155/2018/7529286>
- (16) Wei, L., Song, Y., He, W., Chen, X., Ma, B., Lu, Y., & Zhu, X. (2020). Accuracy Improvement of IOL Power Prediction for Highly Myopic Eyes With an XGBoost Machine Learning-Based Calculator. *Frontiers in Medicine*, 7. <https://www.frontiersin.org/articles/10.3389/fmed.2020.592663>