

Deep Learning Neural Dalam Analisis Sentimen : Sebuah Studi Literatur

Bayu Yanuargi[#], Ema Utami[#], Kusrini[#], Arli Aditya Parikesit^{*}

[#] Doctoral Program of Informatic, Faculty of Informatic Tech, Universitas Amikom, Indonesia

^{*} Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Science, Indonesia

E-mail: bayu.yanuargi[at]students.amikom.ac.id, ema.u[at]amikom.ac.id, kusrini[at]amikom.ac.id, arli.parikesit[at]i3l.ac.id

ABSTRACTS

This study provides a thorough review of the literature on the application of deep learning in sentiment categorization. The major purpose is to collect statistical data on deep learning research for sentiment analysis and hybrid model construction. The research discovered that the most often utilized deep learning algorithms in 2022-2024 were BERT, LSTM, and GRU, each with varied degrees of accuracy. Specifically, GRU had the highest accuracy (98%), followed by LSTM (93.58%) and BERT (91.37%). In addition, 31% of the examined publications modified these methods to create new hybrid models. Among them, the RoBERTa and LSTM hybrid models achieved the highest accuracy (91.01%). This systematic review examines the changing landscape of sentiment analysis using deep learning, focusing on the efficacy of hybrid models in boosting classification accuracy.

Manuscript received Sep 8, 2025;
revised Sep 24, 2025, accepted Sep
25, 2025 Date of publication Sep
30, 2025. International Journal,
JITSI : Jurnal Ilmiah Teknologi
Sistem Informasi licensed under a
Creative Commons Attribution-
Share Alike 4.0 International
License



ABSTRAK

Studi ini memberikan tinjauan mendalam terhadap literatur mengenai penerapan deep learning dalam kategorisasi sentimen. Tujuan utamanya adalah untuk mengumpulkan data statistik terkait penelitian deep learning untuk analisis sentimen dan pengembangan model hibrida. Penelitian ini menemukan bahwa algoritma deep learning yang paling sering digunakan pada tahun 2022–2024 adalah BERT, LSTM, dan GRU, masing-masing dengan tingkat akurasi yang bervariasi. Secara khusus, GRU menunjukkan akurasi tertinggi (98%), diikuti oleh LSTM (93,58%) dan BERT (91,37%). Selain itu, sebanyak 31% dari publikasi yang ditinjau memodifikasi metode ini untuk membentuk model hibrida baru. Di antara model-model tersebut, kombinasi RoBERTa dan LSTM mencatatkan akurasi tertinggi sebesar 91,01%. Tinjauan sistematis ini mengeksplorasi perkembangan lanskap analisis sentimen dengan pendekatan deep learning, dengan fokus pada efektivitas model hibrida dalam meningkatkan akurasi klasifikasi.

Keywords / Kata Kunci — *Klasifikasi Sentime; neural network; prisma; nlp; literatur review*

CORRESPONDING AUTHOR

Bayu Yanuargi
Doctoral Program of Informatic, Faculty of Informatic Tech, Universitas Amikom
Email: bayu.yanuargi[at]students.amikom.ac.id

1. PENDAHULUAN

Pemrosesan bahasa alami (Natural Language Processing/NLP) semakin banyak digunakan dalam analisis sentimen di berbagai industri. Bidang ini, yang juga dikenal sebagai opinion mining, mempelajari opini, sentimen,

evaluasi, sikap, dan emosi yang terkandung dalam teks tertulis. Menurut [1], ini adalah topik kajian yang penting dan terus berkembang. Di era digital saat ini, NLP sering memanfaatkan data dari situs jejaring sosial seperti Twitter dan Facebook. Data ini kemudian diproses dan diubah ke dalam format standar agar lebih mudah dianalisis menggunakan teknologi NLP. Teknik NLP memungkinkan penerapan metode otomatis yang sangat efisien untuk mengekstraksi wawasan dari diskusi yang terjadi di media sosial. Penerapan teknik ini tidak hanya meningkatkan kemampuan untuk menafsirkan data dalam jumlah besar dengan cepat, tetapi juga memperluas batasan penggunaan NLP dalam memahami dinamika komunikasi manusia di era informasi [2].

Kategorisasi sentimen, sebagai komponen utama dalam Pemrosesan Bahasa Alami (NLP), sangat penting untuk membaca dan menafsirkan teks. Tujuan dasar dari analisis sentimen adalah untuk mengkategorikan teks berdasarkan polaritasnya. Polaritas mengacu pada sikap netral, negatif, atau positif yang terkandung dalam teks [3]. Analisis sentimen tidak hanya membantu dalam memahami nuansa emosional dari suatu teks, tetapi juga penting dalam berbagai aplikasi praktis seperti pemantauan media sosial, manajemen merek, dan layanan pelanggan. Namun, pekerjaan ini sering menghadapi banyak tantangan. Salah satu masalah utamanya adalah kesulitan dalam menafsirkan dan menentukan polaritas sentimen secara akurat. Proses ini menjadi sangat kompleks karena sifat bahasa alami yang seringkali ambigu dan subjektif [4].

Opinion mining, atau analisis sentimen, adalah praktik untuk menemukan dan mengklasifikasikan perasaan atau opini seseorang dengan mengevaluasi ulang suatu produk, layanan, atau peristiwa, apakah bersifat netral, negatif, atau positif [5]. Analisis sentimen memerlukan proses pra-pemrosesan atau analisis teks. Proses analisis teks mencakup persiapan dan pembersihan teks. Tergantung pada jenis aplikasi dan algoritma yang diinginkan, proses ini dapat melibatkan langkah-langkah dasar seperti menghapus tanda baca, URL, atau tagar. Algoritma penambahan teks membantu dalam memahami bagaimana teks disiapkan tanpa perlu dibaca oleh manusia, bagaimana mesin meninjau setiap karakter, dan bagaimana karakter-karakter tersebut diatur menggunakan semua kata yang tersedia dalam korpus, baik kata kunci utama maupun kata umum lainnya [6].

Klasifikasi sentimen adalah aspek penting dalam Pemrosesan Bahasa Alami (NLP) yang berfokus pada evaluasi dan interpretasi opini pengguna dalam bentuk teks. Proses ini umumnya dimulai dengan pengumpulan data ulasan dari internet, kemudian dilanjutkan dengan klasifikasi teks untuk memulai fase analisis menggunakan tiga metode utama dalam analisis sentimen: metode berbasis kamus, pembelajaran mesin, dan pendekatan hibrida yang menggabungkan kedua metode tersebut [7]. Media sosial, sebagai platform di mana individu bebas membagikan pandangan dan perasaan mereka mengenai berbagai topik, telah menjadi sumber data yang sangat kaya untuk analisis ini. Analisis ini sangat membantu dalam mengidentifikasi tren perilaku konsumen, sekaligus berperan dalam pengembangan produk dan peningkatan layanan yang ditawarkan. Teknik ini memungkinkan interpretasi yang lebih akurat dan mendalam terhadap data dalam jumlah besar yang dikumpulkan, serta memberikan wawasan yang lebih luas tentang sentimen konsumen [8].

Sejak kemunculannya pada tahun 1990-an, deep learning telah mengalami berbagai inovasi dan pembaruan signifikan, terutama dalam hal jaringan dan teknik analisis data, di mana evolusi ini mencakup pengembangan berbagai algoritma klasifikasi dengan pendekatan deep learning. Beberapa model yang paling menonjol adalah model mekanisme perhatian (attention mechanism), yang meningkatkan fokus model pada item data yang penting; jaringan saraf berulang (recurrent neural networks), yang efektif untuk data berurutan atau deret waktu; dan multi-layer perceptron, yang menjadi dasar bagi banyak jaringan saraf modern [9]. Di sisi lain, dengan pertumbuhan informasi yang sangat pesat di seluruh dunia dan meningkatnya popularitas internet, volume data yang tersedia juga meningkat secara signifikan, sehingga diperlukan pengorganisasian dan ekstraksi data yang efektif, cepat, akurat, dan menyeluruh agar dapat dimanfaatkan secara optimal. Dalam konteks ini, deep learning memberikan solusi yang sangat efektif terhadap permasalahan tersebut. Kekuatan teknik ini berasal dari kemampuannya untuk memproses data yang besar dan kompleks secara lebih intuitif dan otomatis [10].

Tujuan penelitian yang ingin kami capai dalam makalah ini adalah untuk memberikan tinjauan sistematis terhadap studi-studi terbaru yang menggunakan model jaringan saraf atau deep learning dalam mendukung aktivitas analisis sentimen. Tinjauan ini dilakukan dengan menelaah sejumlah makalah yang tersedia di basis data IEEE Xplore, dari tahun 2022 hingga 2024. Pemilihan makalah dibatasi pada jurnal yang tersedia secara akses terbuka, untuk memastikan bahwa informasi yang ditinjau dapat diakses secara luas oleh para peneliti dan praktisi. Kontribusi yang diharapkan dari tinjauan ini adalah memberikan pemahaman yang lebih mendalam tentang bagaimana deep learning telah diterapkan dalam analisis sentimen dan bagaimana hasilnya dapat memengaruhi perkembangan penelitian selanjutnya. Dengan memahami penggunaan dan kinerja model deep learning yang sudah ada, para peneliti dapat mengidentifikasi celah-celah yang ada serta potensi untuk penyempurnaan lebih lanjut. Hal ini sangat penting karena pemahaman yang mendalam terhadap kelebihan dan keterbatasan model saat ini memungkinkan pengambilan langkah-langkah inovatif untuk meningkatkan efektivitas analisis sentimen menggunakan teknologi deep learning.

Selain itu, tinjauan ini juga bertujuan untuk mendorong dialog dan kolaborasi di antara komunitas penelitian yang bergerak di bidang yang sama, sehingga dapat mempercepat penemuan solusi yang lebih efisien dan efektif dalam mengatasi berbagai tantangan dalam analisis sentimen. Dengan demikian, karya ini tidak hanya

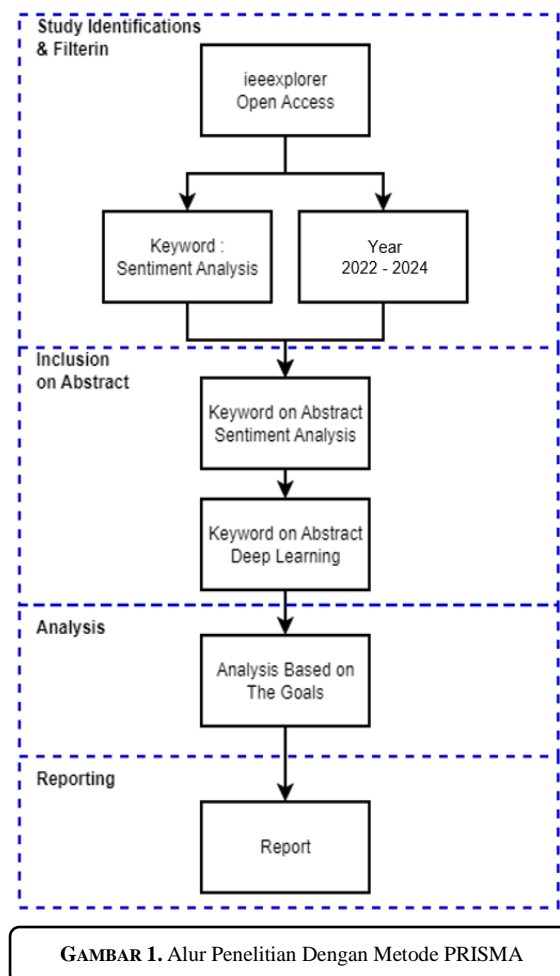
memberikan informasi, tetapi juga berperan sebagai wadah bagi penelitian di masa depan yang bertujuan meningkatkan kinerja dan penerapan pendekatan deep learning dalam analisis sentimen

2. METODOLOGI PENELITIAN

Metode PRISMA adalah kerangka kerja standar yang dirancang untuk mengatur dan melaporkan temuan dari tinjauan sistematis dan meta-analisis. PRISMA bertujuan untuk mendorong transparansi dan konsistensi dalam pelaporan penelitian, sehingga pembaca dapat memahami proses tinjauan sistematis serta menilai validitas dan reliabilitas hasilnya. Dengan mengikuti pedoman PRISMA, peneliti memberikan penjelasan yang jelas dan rinci tentang metodologi mereka, yang meningkatkan keterulangan penelitian serta memudahkan evaluasi secara kritis. Pendekatan yang terstandarisasi ini sangat penting untuk menjaga kualitas tinggi dalam tinjauan sistematis, sehingga memungkinkan sintesis temuan penelitian yang lebih dapat diandalkan dan menyeluruh. Makalah ini menerapkan metode PRISMA dalam proses tinjauan, agar hasilnya mudah dipahami oleh pembaca dan memiliki dasar ilmiah yang kuat. Berikut adalah langkah-langkah utama dalam metode PRISMA untuk melakukan tinjauan pustaka sistematis:

1. Mengidentifikasi semua studi yang relevan dengan melakukan pencarian menyeluruh di berbagai basis data dan sumber informasi.
2. Menyaring studi yang ditemukan melalui judul dan abstrak, lalu meninjau teks lengkap dari studi yang tampak relevan untuk memastikan bahwa studi tersebut memenuhi kriteria inklusi.
3. Mengumpulkan data dari studi yang memenuhi kriteria inklusi, termasuk informasi tentang metode, populasi, intervensi, hasil, dan temuan utama.
4. Menganalisis data yang telah dikumpulkan, yang dapat mencakup penggabungan data kuantitatif melalui meta-analisis atau sintesis naratif dari temuan.
5. Menyusun hasil tinjauan sesuai pedoman PRISMA, termasuk membuat diagram alur PRISMA yang menunjukkan bagaimana studi dipilih dan disertakan dalam tinjauan.

Tahapan yang disebutkan di atas umumnya digambarkan dalam diagram alur PRISMA pada Gambar 1 di bawah ini, yang merupakan representasi visual dari proses pemilihan dan penyaringan studi untuk tinjauan sistematis atau meta-analisis.



Mengacu pada diagram alur penelitian pada Gambar 1, proses identifikasi dan penyaringan dilakukan di situs web ieeexplore (<https://ieeexplore.ieee.org/>) dengan ketentuan sebagai berikut:

- Kata kunci pencarian adalah sentiment analysis
- Tahun pencarian adalah 2022 – 2024
- Status makalah adalah Open Access paper

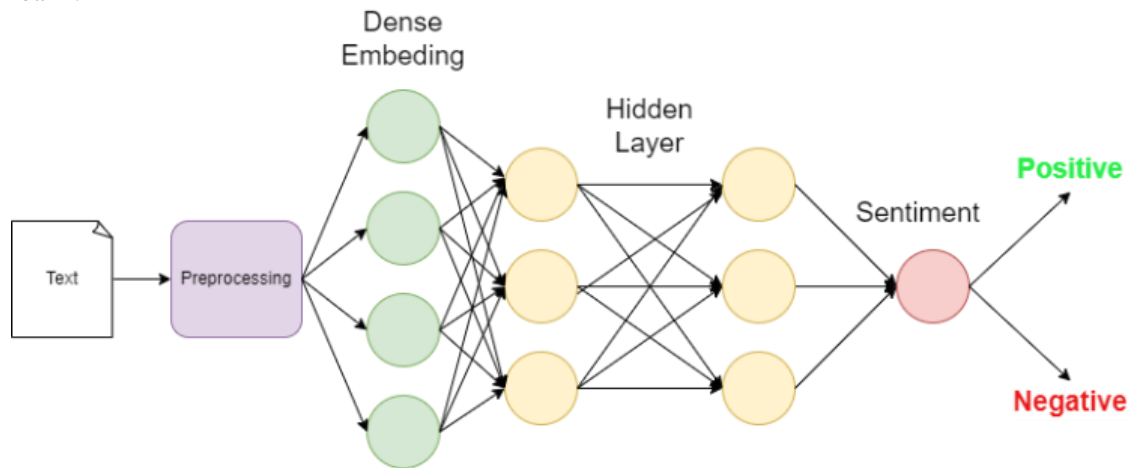
Dalam proses pencarian dan penyaringan, diperoleh sebanyak 476 makalah yang sesuai dengan tiga kriteria pencarian tersebut. Hasil pencarian sebanyak 476 makalah tersebut kemudian melalui dua tahap proses deteksi kriteria inklusi pada bagian abstrak sebagai berikut:

1. Kriteria inklusi awal melibatkan identifikasi abstrak yang memuat istilah "sentiment" dan "opinion mining." Langkah ini memastikan bahwa hanya studi yang relevan yang dipertimbangkan untuk analisis lebih lanjut. Setelah abstrak yang sesuai terdeteksi, makalah tersebut diseleksi untuk tahapan peninjauan selanjutnya. Metode ini memastikan fokus tetap pada studi yang berkaitan langsung dengan analisis sentimen dan opinion mining, sehingga proses tinjauan menjadi lebih efisien dan terarah dengan menyaring konten yang tidak relevan sejak awal. Dengan menerapkan kriteria ini, proses tinjauan menjadi lebih terstruktur dan memberikan dasar yang kuat untuk kajian sistematis yang komprehensif dan relevan.

- Kriteria inklusi kedua merupakan lanjutan dari tahap seleksi awal, dengan fokus pada pendeteksian keberadaan istilah "deep learning" dan "neural network" di dalam abstrak. Langkah ini memastikan bahwa hanya makalah yang secara eksplisit membahas konsep-konsep tersebut yang dipertimbangkan untuk tinjauan lebih lanjut. Dengan mendeteksi kata kunci tersebut pada abstrak, kriteria ini menyaring studi yang tidak relevan, sehingga mempersempit kumpulan literatur hanya pada studi yang paling relevan dengan tujuan penelitian. Pendekatan sistematis ini membantu menjaga relevansi dan kualitas dalam kajian sistematis, memastikan bahwa studi yang disertakan memberikan wawasan yang bermakna mengenai penerapan deep learning dan neural networks. Strategi ini sangat penting untuk menyusun kumpulan penelitian yang terfokus dan mendalam, serta memberikan kontribusi pada analisis dan pemahaman topik yang lebih kuat.

3. HASIL DAN PEMBAHASAN

Deep learning menggunakan pendekatan berlapis-lapis pada lapisan tersembunyi dalam jaringan saraf, yang membedakannya dari metode pembelajaran mesin tradisional. Berbeda dengan pendekatan konvensional, model jaringan saraf memiliki kemampuan untuk secara otomatis mempelajari dan mengekstraksi fitur dari data, sehingga secara signifikan meningkatkan akurasi dan kinerja. Kemampuan ini memungkinkan deep learning untuk menangani pola yang kompleks dan dataset yang sangat besar dengan lebih efektif, menjadikannya alat yang berguna untuk berbagai aplikasi dalam kecerdasan buatan dan analisis data, seperti yang ditunjukkan pada Gambar 2.



GAMBAR 2. Deep Learning Dalam Analisis Sentimen

Tahapan seleksi makalah berdasarkan metodologi menghasilkan sejumlah makalah seperti yang ditunjukkan pada Tabel 1 di bawah ini, di mana hasil akhir dari proses penyaringan untuk menentukan makalah yang akan digunakan dalam tinjauan adalah sebanyak 19 makalah yang telah lolos proses seleksi tahap ketiga. Tahapan seleksi makalah dapat dilihat pada Tabel 1 di bawah ini.

TABEL 1. Tahapan Inklusi dan Jumlah Paper

Tahapan	Filtering Process	Jumlah Paper
1	Identifikasi Kriteria	476
2	Kriteria Inklusi Pertama	93
3	Kriteria Inklusi Kedia	19

TABEL 2. Tahapan Inklusi dan Jumlah Paper

Basic Model	Jumlah Paper	References
LSTM	5	[11], [4], [5], [6], [7]
GRU	4	[8], [9], [12], [13]
CNN	3	[4], [14], [15]
BERT	6	[8], [14], [16], [6], [17], [18]
GCN	2	[14], [19]
RNN	1	[20]
CapsNETS	1	[21]

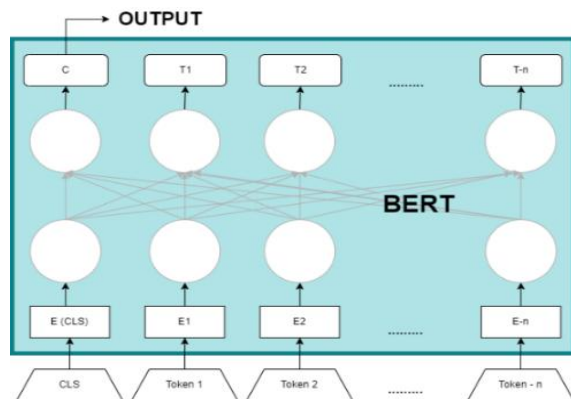
Proses tinjauan yang dilakukan dalam makalah ini didasarkan pada tiga pertanyaan penelitian sebagai berikut:

- RQ 1: Metode atau algoritma deep learning apa yang digunakan dalam penelitian Sentiment Analysis selama tahun 2022 – 2024?
- RQ 2: Metode deep learning mana yang memiliki akurasi terbaik?
- RQ 3: Bagaimana tren penggunaan algoritma hybrid dalam Sentiment Analysis?

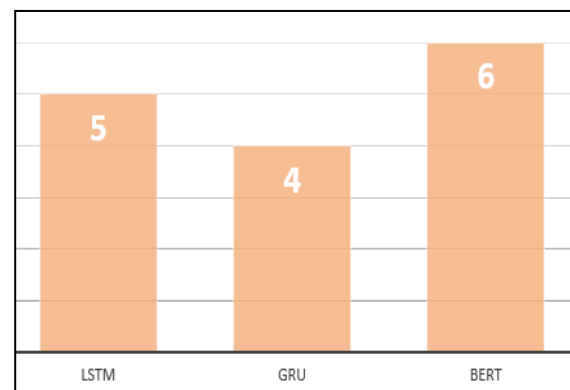
Berdasarkan ketiga pertanyaan penelitian yang telah ditetapkan, sembilan makalah telah dianalisis secara sederhana untuk memperoleh informasi terkait metode deep learning yang digunakan serta variasi akurasi yang diperoleh. Pada Tabel 2 di bawah ini, Anda dapat melihat jenis-jenis algoritma dasar yang digunakan dalam klasifikasi sentimen dalam makalah-makalah dari tahun 2022 hingga 2024.

Seperti yang disajikan dalam Tabel 3 di atas, mayoritas studi yang ditinjau menggunakan algoritma dasar BERT, yang menegaskan adopsinya yang luas dalam aplikasi penambangan teks dan opini. BERT telah mendapatkan popularitas yang signifikan karena kemampuannya yang luar biasa dalam menangani tugas-tugas bahasa yang kompleks, menjadikannya sebagai pilar utama dalam bidang pemrosesan bahasa alami (Natural Language Processing/NLP). Popularitas ini berasal dari performa kuatnya dalam memahami dan menganalisis data tekstual, menjadikannya solusi andalan bagi para peneliti dan praktisi yang mencari alat efektif untuk aplikasi berbasis teks. Tingginya frekuensi penggunaan BERT dalam berbagai makalah yang ditinjau menunjukkan fungsionalitas dan adaptabilitasnya yang tangguh dalam beragam tugas, sekaligus memperkuat reputasinya sebagai algoritma yang andal dan efisien dalam alur kerja NLP.

BERT, singkatan dari Bidirectional Encoder Representations from Transformers, adalah model berbasis transformer yang dirancang khusus untuk menangkap representasi dua arah yang mendalam dari data tersembunyi. Tidak seperti model tradisional, arsitektur BERT, seperti yang digambarkan dalam Gambar 3, memungkinkannya memproses informasi kontekstual dari dua arah secara bersamaan, sehingga memberikan pemahaman teks yang lebih mendalam dan bernuansa. Kemampuan ini membuat BERT sangat cocok untuk berbagai tugas pemrosesan bahasa alami, seperti klasifikasi teks, analisis sentimen, dan penjawaban pertanyaan. Dengan memanfaatkan sifat dua arahnya, BERT unggul dalam memahami konteks dan makna dalam kalimat, serta melampaui model-model sebelumnya dalam hal akurasi dan efisiensi. Fleksibilitas dan efektivitasnya telah mengukuhkan posisi BERT sebagai alat unggulan dalam kemajuan penelitian dan aplikasi NLP.



GAMBAR .3 Architecture Dari BERT Dalam NLP



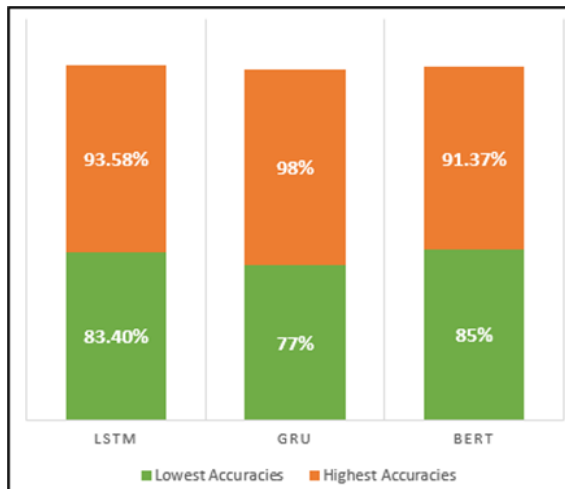
GAMBAR 4. Tiga Algoritma yang Lebih Sering Digunakan

Seperti yang digambarkan pada Gambar 3 di atas, BERT dikembangkan dengan menggabungkan dua konsep utama yang membuatnya sangat kuat untuk tugas pemrosesan bahasa alami (Natural Language Processing/NLP). Pertama, BERT menggunakan model transformer yang dalam, yang memungkinkannya untuk memproses kalimat panjang secara efektif melalui mekanisme perhatian (attention) yang canggih. Kedua, BERT beroperasi sebagai jaringan dua arah (bidirectional), yang memungkinkan model ini mempertimbangkan keseluruhan konteks dari suatu segmen teks ketika menganalisis setiap token. Kemampuan ganda ini memastikan pemahaman yang menyeluruh terhadap makna dan hubungan dalam teks, menjadikan BERT sebagai alat yang sangat efektif dalam menangkap pola linguistik yang halus [22]. Fitur-fitur ini telah menjadikan BERT sebagai fondasi dalam penelitian dan aplikasi NLP, terutama dalam penambangan teks dan opini.

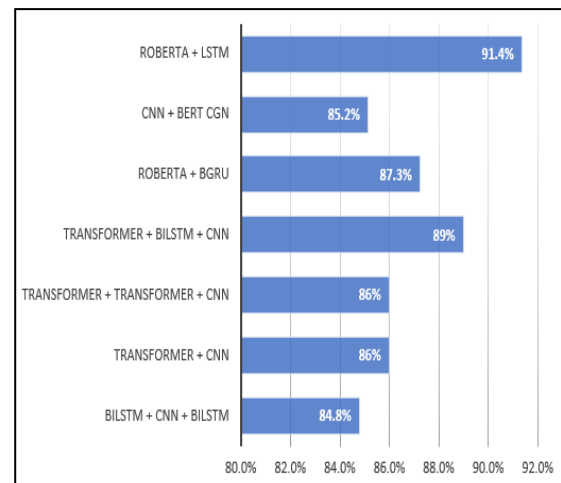
Proses peninjauan artikel seperti yang ditunjukkan pada Gambar 4 di atas, menunjukkan bahwa model GRU dasar mencapai akurasi tertinggi, yaitu hingga 98%. Menyusul di belakangnya adalah model LSTM dengan akurasi puncak sebesar 93,58%, sementara model BERT menunjukkan akurasi maksimum sebesar 91,37%. Hasil-hasil ini menyoroti kinerja kuat GRU dalam skenario tertentu, menjadikannya pesaing yang patut diperhatikan dalam tugas-tugas pemrosesan bahasa alami. Namun, meskipun mencapai kinerja terbaik, GRU menunjukkan variabilitas antar studi, dengan akurasi terendah tercatat sebesar 77%. Variabilitas ini menunjukkan bahwa meskipun GRU dapat unggul dalam kondisi optimal, konsistensinya dapat bergantung pada persyaratan tugas dan dataset tertentu.

Sebagai perbandingan, model LSTM menunjukkan kinerja yang lebih stabil, dengan akurasi minimum sebesar 83,4%, menyoroti keandalannya dalam memproses data berurutan dan memahami hubungan kontekstual. Demikian pula, model BERT yang dikenal karena pendekatannya yang dua arah dan pemahaman kontekstual yang kuat, mempertahankan akurasi minimum sebesar 85%, yang menunjukkan efektivitas yang konsisten dalam berbagai aplikasi. Seperti yang digambarkan pada Gambar 5, perbedaan dalam metrik kinerja ini memberikan

wawasan berharga mengenai kekuatan dan keterbatasan masing-masing model. Meskipun GRU unggul dalam akurasi puncak, kestabilan LSTM dan BERT menempatkan mereka sebagai pilihan yang andal untuk berbagai tugas yang lebih luas, menekankan pentingnya mempertimbangkan kasus penggunaan spesifik saat memilih algoritma.



GAMBAR 5. Akurasi dari tiga algoritma yang sering digunakan dan Jumlah Paper



GAMBAR 6. Akurasi dari Algoritma Hybrid dalam Paper Yang di Review digunakan dan Jumlah Paper

Mengingat dominasi algoritma ini dalam studi-studi yang ditinjau, tahap selanjutnya dari tinjauan sistematis untuk Pertanyaan Penelitian RQ2 dan RQ3 berfokus pada tiga algoritma teratas yang diidentifikasi dalam analisis ini. Seperti yang ditampilkan pada Gambar 4, ketiga algoritma tersebut adalah BERT, LSTM, dan GRU, yang masing-masing dikenal karena kekuatannya dalam memproses data tekstual dan memahami pola berurutan. Pemilihan algoritma ini mencerminkan penerapannya yang luas serta efektivitas yang telah terbukti dalam menangani berbagai tantangan NLP. Pendekatan sistematis ini akan memberikan wawasan yang lebih dalam mengenai kinerja dan kontribusi komparatif mereka, yang akan memperkaya pemahaman tentang teknik-teknik canggih dalam analisis teks.

Dalam beberapa penelitian, algoritma dasar dimodifikasi untuk mencapai hasil yang lebih optimal, mencerminkan perkembangan teknik pemrosesan bahasa alami yang terus berubah. Modifikasi yang diterapkan pada tiga algoritma utama—LSTM, GRU, dan BERT—melibatkan penggabungan arsitektur mereka ke dalam model hibrida serta penyesuaian alur algoritmanya. Peningkatan ini bertujuan untuk memanfaatkan keunggulan unik dari masing-masing algoritma sekaligus mengatasi keterbatasan individual mereka. Pendekatan hibrida memungkinkan pemahaman data yang lebih mendalam dengan mengintegrasikan komponen arsitektural yang berbeda, sehingga meningkatkan performa dalam berbagai tugas.

Seperti yang ditunjukkan pada Gambar 6, model-model hibrida menghasilkan rerata tingkat akurasi yang bervariasi, tergantung pada kombinasi spesifik yang digunakan. Misalnya, GRU, BERT, dan LSTM diintegrasikan dengan model lain seperti Graph Convolutional Networks (GCN) dan Convolutional Neural Networks (CNN). Kombinasi ini memanfaatkan kemampuan komplementer dari masing-masing algoritma, seperti pemahaman kontekstual dari BERT, kemampuan LSTM dalam menangkap ketergantungan berurutan, dan efisiensi komputasi GRU. Hasil penelitian menunjukkan bahwa model-model hibrida dapat melampaui performa model tunggal dengan lebih efektif mengatasi tantangan spesifik pada tiap tugas. Analisis ini menekankan pentingnya inovasi dalam arsitektur model untuk memaksimalkan potensi algoritma yang sudah ada sekaligus menjelajahi cakrawala baru dalam pemrosesan bahasa alami.

Menurut Gambar 6, model RoBERTa-LSTM mencapai akurasi tertinggi (rerata) di antara model-model yang dievaluasi, yakni sebesar 91,4%. Pendekatan hibrida ini menggabungkan kekuatan dari RoBERTa (BERT yang telah dioptimalkan secara robust) dan jaringan Long Short-Term Memory (LSTM), menjadikannya alat yang sangat kuat untuk analisis sentimen. Dengan memanfaatkan bobot pra-pelatihan RoBERTa, model ini secara efektif memetakan token ke dalam ruang embedding yang kaya dan bermakna, sehingga meningkatkan kemampuannya dalam memahami struktur bahasa yang kompleks.

Arsitekturnya kemudian mengintegrasikan LSTM, sebuah jaringan neural berulang yang dirancang untuk menangkap ketergantungan temporal dan fitur semantik penting dalam data berurutan. Kombinasi ini memungkinkan model RoBERTa-LSTM memproses dan menganalisis pola bahasa dengan presisi tinggi, karena embedding yang dihasilkan oleh RoBERTa disempurnakan melalui lapisan LSTM. Integrasi dari dua pendekatan ini menunjukkan peningkatan kinerja yang signifikan, sebagaimana dibuktikan oleh metrik akurasi model

4. KESIMPULAN

Makalah ini membahas dua pertanyaan penelitian utama: pertama, mengidentifikasi variasi dalam metode deep learning yang digunakan untuk klasifikasi sentimen; dan kedua, meninjau akurasi dari model-model tersebut serta potensi untuk dimodifikasi dan dikombinasikan. Melalui tinjauan sistematis terhadap penelitian yang dilakukan antara tahun 2022 hingga 2024, sebanyak 19 makalah yang berfokus pada model deep learning untuk analisis sentimen telah dianalisis. Di antara studi-studi tersebut, tiga algoritma muncul sebagai yang paling umum digunakan—BERT, LSTM, dan GRU. Temuan menunjukkan bahwa akurasi tertinggi dari model mencapai lebih dari 90%, sementara akurasi terendah tetap berada di atas 70%, yang menunjukkan ketangguhan pendekatan-pendekatan ini dalam tugas klasifikasi sentimen. Selain itu, 31% dari makalah yang ditinjau mengeksplorasi modifikasi model dengan menggabungkan algoritma untuk meningkatkan kinerja. Secara khusus, model hibrida RoBERTa + LSTM mencapai akurasi tertinggi sebesar 91,4%, menyoroti potensi pendekatan terintegrasi.

Metodologi yang digunakan dalam makalah ini sangat menekankan pada analisis kuantitatif untuk mengevaluasi pemanfaatan model deep learning di seluruh literatur yang ditinjau. Pendekatan ini mencakup analisis statistik terhadap variasi dan perkembangan model, memberikan perspektif yang jelas tentang evolusi teknik analisis sentimen. Meskipun akurasi model merupakan faktor penting, studi ini menyarankan agar tinjauan literatur sistematis di masa mendatang memperluas fokus untuk mencakup kinerja komputasional. Dengan demikian, peneliti dapat memberikan pemahaman yang lebih komprehensif mengenai pertukaran antara akurasi dan efisiensi komputasi, yang sangat penting untuk penerapan analisis sentimen di lingkungan dunia nyata.

Dengan menekankan aspek yang lebih luas dari kinerja model, makalah ini berkontribusi pada kemajuan bidang analisis sentimen. Makalah ini mendorong penelitian di masa depan untuk tidak hanya menyoroti akurasi model deep learning, tetapi juga biaya komputasi dan efisiensi keseluruhannya. Pendekatan semacam ini akan memungkinkan para peneliti dan praktisi untuk membuat keputusan yang lebih tepat dalam memilih atau mengembangkan model, yang pada akhirnya akan mendorong inovasi dalam klasifikasi sentimen dan penerapan deep learning.

REFERENSI

- [1] M. J. Page et al., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” *BMJ*, vol. 372, 2021, doi: 10.1136/bmj.n71.
- [2] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, “Sentiment analysis based on deep learning: A comparative study,” *Electronics (Switzerland)*, vol. 9, no. 3, Mar. 2020, doi: 10.3390/electronics9030483.
- [3] C.-G. Kim, Y.-J. Hwang, and C. Kamyod, “A Study of Profanity Effect in Sentiment Analysis on Natural Language Processing Using ANN,” *Journal of Web Engineering*, vol. 21, no. 3, pp. 751–766, 2022, doi: 10.13052/jwe1540-9589.2139.
- [4] M. Shukla and A. Kumar, “An Experimental Analysis of Deep Neural Network Based Classifiers for Sentiment Analysis Task,” *IEEE Access*, vol. 11, pp. 36929–36944, 2023, doi: 10.1109/ACCESS.2023.3266640.
- [5] M. Bilal, A. Khan, S. Jan, and S. Musa, “Context-Aware Deep Learning Model for Detection of Roman Urdu Hate Speech on Social Media Platform,” *IEEE Access*, vol. 10, pp. 121133–121151, 2022, doi: 10.1109/ACCESS.2022.3216375.
- [6] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, and K. M. Lim, “RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network,” *IEEE Access*, vol. 10, pp. 21517–21525, 2022, doi: 10.1109/ACCESS.2022.3152828.
- [7] Y. Li and N. Li, “Sentiment Analysis of Weibo Comments Based on Graph Neural Network,” *IEEE Access*, vol. 10, pp. 23497–23510, 2022, doi: 10.1109/ACCESS.2022.3154107.
- [8] K. Jahanbin and M. A. Z. Chahooki, “Aspect-Based Sentiment Analysis of Twitter Influencers to Predict the Trend of Cryptocurrencies Based on Hybrid Deep Transfer Learning Models,” *IEEE Access*, vol. 11, pp. 121656–121670, 2023, doi: 10.1109/ACCESS.2023.3327060.
- [9] A. Altaf et al., “Deep Learning Based Cross Domain Sentiment Classification for Urdu Language,” *IEEE Access*, vol. 10, pp. 102135–102147, 2022, doi: 10.1109/ACCESS.2022.3208164.
- [10] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, “Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model,” *IEEE Access*, vol. 10, pp. 39313–39324, 2022, doi: 10.1109/ACCESS.2022.3165621.

- [11] C. G. Kim, Y. J. Hwang, and C. Kamyod, "A Study of Profanity Effect in Sentiment Analysis on Natural Language Processing Using ANN," *Journal of Web Engineering*, vol. 21, no. 3, pp. 751–766, 2022, doi: 10.13052/jwe1540-9589.2139.
- [12] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model," *IEEE Access*, vol. 10, pp. 39313–39324, 2022, doi: 10.1109/ACCESS.2022.3165621.
- [13] J. Mao, Z. Qian, and T. Lucas, "Sentiment Analysis of Animated Online Education Texts Using Long Short-Term Memory Networks in the Context of the Internet of Things," *IEEE Access*, vol. 11, pp. 109121–109130, 2023, doi: 10.1109/ACCESS.2023.3321303.
- [14] H. T. Phan, N. T. Nguyen, and D. Hwang, "Aspect-Level Sentiment Analysis Using CNN Over BERT-GCN," *IEEE Access*, vol. 10, pp. 110402–110409, 2022, doi: 10.1109/ACCESS.2022.3214233.
- [15] M. Yekrangi and N. S. Nikolov, "Domain-Specific Sentiment Analysis: An Optimized Deep Learning Approach for the Financial Markets," *IEEE Access*, vol. 11, pp. 70248–70262, 2023, doi: 10.1109/ACCESS.2023.3293733.
- [16] J. K. Adarsh, V. T. Sreedevi, and D. Thangavelusamy, "Product Review System With BERT for Sentiment Analysis and Implementation of Administrative Privileges on Node-RED," *IEEE Access*, vol. 11, pp. 65968–65976, 2023, doi: 10.1109/ACCESS.2023.3275738.
- [17] R. Chandra and V. Kulkarni, "Semantic and Sentiment Analysis of Selected Bhagavad Gita Translations Using BERT-Based Language Framework," *IEEE Access*, vol. 10, pp. 21291–21315, 2022, doi: 10.1109/ACCESS.2022.3152266.
- [18] S. Mifrah and E. H. Benlahmar, "Sentence-Level Sentiment Classification A Comparative Study between Deep Learning Models," *Journal of ICT Standardization*, vol. 10, no. 2, pp. 339–352, 2022, doi: 10.13052/jicts2245-800X.10213.
- [19] J. Chen, H. Fan, and W. Wang, "Syntactic and Semantic Aware Graph Convolutional Network for Aspect-Based Sentiment Analysis," *IEEE Access*, vol. 12, pp. 22500–22509, 2024, doi: 10.1109/ACCESS.2024.3364353.
- [20] P. Durga and D. Godavarthi, "Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-RNN)," *IEEE Access*, vol. 11, pp. 108433–108447, 2023, doi: 10.1109/ACCESS.2023.3320738.
- [21] V. Diviya Prabha and R. Rathipriya, "Competitive Capsule Network Based Sentiment Analysis on Twitter COVID'19 Vaccines," *Journal of Web Engineering*, vol. 21, no. 5, pp. 1583–1602, 2022, doi: 10.13052/jwe1540-9589.2159.
- [22] S. Gundapu and R. Mamidi, "Transformer based Automatic COVID-19 Fake News Detection System," Jan. 2021, [Online]. Available: <http://arxiv.org/abs/2101.00180>