JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 3, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i3.6588

ANALISIS KOMPARATIF BILSTM DAN BIGRU DENGAN WORD EMBEDDING GLOVE TERHADAP SENTIMEN PUBLIK TENTANG COVID-19 DI TWITTER

Vellanindhita Noorpramesari Yudoyono[1], Jimmy Maulana[2], Ahmad Riyo Alfath[3], Risma Melati[4], Mutiara Anastasya Sihaloho[5], Abdiansah[6]

[1,2]Universitas Sriwijaya; Jl. Palembang – Prabumulih KM.32 Kabupaten Ogan Ilir, Sumatera Selatan. 30662; Telp. (0711) 379249

Keywords:
Analisis Sentimen;
Twitter;
BiGRU; BiLSTM
GloVe Word Embedding.

Corespondent Email:
vellanindhita.tech@gmail.com

Abstrak. Pandemi COVID-19 telah memicu perdebatan publik di Twitter, yang dapat dijelaskan melalui analisis sentimen menggunakan pemrosesan bahasa alami. Karakter informal dan tak terstruktur cuitan Twitter menjadi tantangan tersendiri. Penelitian ini membandingkan kinerja arsitektur BiLSTM dan BiGRU dalam mengklasifikasikan sentimen publik terkait COVID-19. Model BiGRU dirancang dengan dua lapisan, disertai dengan implementasi GloVe embedding Twitter.27B.200d untuk representasi kata lebih baik dan dilengkapi dengan dilengkapi dropout, batch normalization, dan regularisasi L2, serta dioptimasi dengan AdamW, sedangkan BiLSTM menggunakan satu lapis standar. Hasil eksperimen menunjukkan BiGRU dua lapis mencapai validasi akurasi 86.78% dengan pelatihan yang lebih stabil dibandingkan BiLSTM 84.91% yang cenderung overfitting. Temuan ini mengindikasikan bahwa arsitektur double layer BiGRU lebih efektif memahami konteks dari cuitan yang tidak terstruktur Twitter, sehingga direkomendasikan untuk sistem analisis sentimen publik dan pengembangan pemrosesan bahasa alami di masa depan.

Abstract. *The COVID-19 pandemic has triggered public opinion on Twitter, which can be explained through sentiment analysis using natural language processing. The informal and unstructured nature of Twitter tweets poses a challenge. This study compares the performance of BiLSTM and BiGRU architectures in classifying public sentiment related to COVID-19. The BiGRU model is designed with two layers, accompanied by the implementation of GloVe embedding Twitter.27B.200d for better word representation and equipped with dropout, batch normalization, and L2 regularization, and optimized with AdamW, while BiLSTM uses a standard single layer. The experimental results show that the two-layer BiGRU achieves a validation accuracy of 86.78% with a more stable training graph compared to BiLSTM 84.91% which tends to overfit. These findings indicate that the double-layer BiGRU architecture is more effective in understanding the context of unstructured Twitter tweets, so it is recommended for public sentiment analysis systems and natural language processing developments in the future.*

1. PENDAHULUAN

Media sosial, khususnya Twitter, telah berkembang menjadi salah satu sumber utama untuk mengekspresikan opini publik secara waktu nyata, salah satunya dalam situasi ketika masyarakat tidak bisa bertatap muka secara langsung seperti pada pandemi COVID-19. Cuitan di Twitter mencerminkan respon emosional masyarakat terhadap isu-isu yang tengah berkembang, menjadikannya data yang kaya untuk dianalisis dalam konteks sosial dan kebijakan publik. Dalam konteks ini, analisis

sentimen berbasis teks merupakan pendekatan penting untuk memahami persepsi masyarakat dan mendukung pengambilan keputusan yang berbasis data.

Seiring dengan meningkatnya volume data tak terstruktur dari media sosial, pendekatan berbasis *deep learning* telah banyak digunakan dalam penyelesaian masalah analisis sentimen. Model seperti Long Short-Term Memory (LSTM) dan Gated Recurrent Unit (GRU) terbukti efektif dalam menangkap konteks sekuensial dari teks, terutama ketika dikembangkan dalam bentuk dua arah (bidirectional). Beberapa penelitian sebelumnya [1] telah menunjukkan performa tinggi dari model BiLSTM dan BiGRU dalam menyelesaikan masalah serupa, termasuk deteksi emosi dan prediksi opini publik.

Meskipun demikian, masih terdapat kesenjangan dalam pemahaman mengenai efektivitas dan efisiensi relatif dari kedua arsitektur tersebut dalam menangani teks media sosial yang bersifat informal dan tidak terstruktur. Banyak studi cenderung hanya mengevaluasi salah satu model atau tidak membahas stabilitas dan risiko *overfitting* yang muncul dalam implementasinya. Selain itu, masih terbatas studi yang secara eksplisit menguji pengaruh penggunaan *pre-trained word embedding* seperti GloVe untuk menyelesaikan masalah sentimen analisis pada cuitan bertema COVID-19.

Penelitian ini dilakukan untuk menjawab pertanyaan utama: model mana yang lebih unggul dalam klasifikasi sentimen publik mengenai COVID-19 dengan mempertimbangkan akurasi, stabilitas pelatihan, dan potensi *overfitting*. Studi ini juga menelusuri sejauh mana representasi kata dari GloVe Twitter.27B.200d berkontribusi terhadap peningkatan performa model dalam memahami konteks emosional dari teks Twitter. Dengan demikian, penelitian ini menawarkan kontribusi terhadap pengembangan model NLP yang lebih tepat guna untuk analisis sentimen pada data teks media sosial, sekaligus menghadirkan evaluasi komprehensif atas dua arsitektur populer dalam *deep learning*.

2. TINJAUAN PUSTAKA

2.1. Pemrosesan Bahasa Alami

Pemrosesan Bahasa Alami (Natural Language Processing/NLP) adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP memungkinkan komputer untuk membaca, memahami, dan menghasilkan bahasa manusia secara bermakna. Aplikasi NLP mencakup penerjemah otomatis, chatbot, sistem tanya-jawab, dan analisis sentimen. Teknologi ini memanfaatkan linguistik komputasional, statistik, serta teknik pembelajaran mesin [2].

2.2. Analisis Sentimen

Analisis sentimen adalah proses otomatis untuk mengidentifikasi opini dalam sebuah teks dan mengklasifikasikannya menjadi sentimen tertentu, misalnya: positif, negatif, atau netral.

Teknik ini penting untuk memahami persepsi publik di media sosial, layanan pelanggan, dan analisis pasar. Metodenya bisa berbasis leksikon atau pembelajaran mesin, di mana model dilatih dengan data berlabel untuk mengenali pola sentimen [3].

Selain itu, pendekatan serupa juga digunakan dalam tugas deteksi emosi berbasis suara dan teks, di mana model deep learning telah terbukti efektif dalam mengidentifikasi ekspresi emosional pengguna [8].

2.3. Pra-Pemrosesan Teks

Pra-pemrosesan teks merupakan tahap penting dalam analisis data teks yang bertujuan untuk membersihkan dan menyiapkan data agar dapat diproses lebih lanjut oleh model pembelajaran mesin. Tahapan ini mencakup berbagai proses seperti mengubah semua huruf menjadi huruf kecil (*lowercasing*), memecah teks menjadi satuan kata (tokenisasi), serta menghapus kata-kata umum yang tidak memberikan makna signifikan dalam analisis (*stopword removal*). Pembersihan teks bertujuan untuk menghilangkan komponen asing seperti tanda baca dan huruf khusus, selanjutnya *casefolding* dilakukan untuk mengubah setiap kata menjadi huruf kecil (lowercase), sehingga variasi kapitalisasi pada kata-kata yang sama dapat diseragamkan [4].

Selain itu, proses *stemming* atau *lemmatization* dilakukan untuk mengubah kata ke bentuk dasarnya. Teks juga dibersihkan dari simbol, angka, dan elemen yang tidak relevan atau bukan data yang diinginkan. Tahapan-tahapan ini bertujuan untuk mereduksi *noise*

dalam data sehingga model dapat lebih fokus pada informasi yang relevan dan meningkatkan performa prediksi secara keseluruhan.

2.4. Bidirectional Long Short-Term Memory **(BiLSTM)**

Long Short-Term Memory (LSTM) merupakan jaringan saraf yang dirancang untuk mengatasi *vanishing gradient* dan menangkap ketergantungan jangka panjang dalam data urutan. LSTM terdiri dari beberapa komponen utama yaitu *input gate*, *forget gate*, *memory cell*, dan *output gate*. BiLSTM merupakan pengembangan dari LSTM yang memproses data dalam dua arah, yaitu maju (*forward*) dan mundur (*backward*), sehingga dapat menangkap konteks dari kedua sisi urutan [5]

Proses pertama dalam LSTM adalah *forget gate* yang dirumuskan sebagai

$$f_t = \sigma_g(w_f\, x_t + U_f\, h_{t-1} + V_f\, c_{t-1} + b_f) \quad (1)$$

di mana *gate* ini menyaring informasi yang tidak relevan dari memori sebelumnya (c_{t-1}). Selanjutnya, *input gate* bekerja untuk menentukan informasi baru yang akan disimpan, dengan rumus.

$$i_t = \sigma_g(W_i\, x_t + U_i\, h_{t-1} + V_i\, c_{t-1} + b_i) \quad (2)$$

Kandidat memori baru dihitung menggunakan

$$c'_t = \sigma_c(W_c\, x_t + U_c h_t - 1 + V_c\, c_t - 1 + b_c) \quad (3)$$

yang akan dikombinasikan dengan memori sebelumnya dalam proses pembaruan memori

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c'_t \quad (4)$$

Setelah itu, *output gate* menentukan bagian mana dari memori yang akan digunakan sebagai output melalui

$$o_t = \sigma(W_o\, x_t + U_o h_{t-1} + V_o\, c_{t-1} + b_o) \quad (5)$$

dan hasil akhirnya berupa *hidden state* yang dihitung dengan

$$h_t = o_t \cdot \sigma(c_t)) \quad (6)$$

Bidirectional LSTM (BiLSTM) terdiri dari dua lapisan LSTM yang berjalan secara paralel, satu memproses urutan dari awal ke akhir (forward), dan satu lagi dari akhir ke awal (backward). Hasil dari kedua arah ini digabungkan untuk membentuk representasi konteks yang lebih kaya. BiLSTM terbukti lebih unggul dibandingkan LSTM satu arah karena mampu memahami informasi dari kedua sisi urutan. Keunggulan ini terlihat jelas pada berbagai tugas, seperti pemrosesan bahasa alami dan bahkan dalam domain non-teks seperti peramalan deret waktu keuangan, di mana BiLSTM tetap menunjukkan akurasi yang lebih tinggi [5][6].

2.5. Bidirectional Gated Recurrent Unit (BiGRU)

GRU merupakan varian dari RNN yang menyederhanakan struktur LSTM. GRU hanya menggunakan dua gerbang utama, yaitu *reset gate* dan *update gate*, serta *candidate hidden state*. Rumus-rumus utama dalam GRU meliputi *Reset Gate*, *Update Gate*, *Candidate Hidden State*, *Hidden State* [1].

Pertama *reset gate* mengatur seberapa banyak informasi lama dari *hidden state* sebelumnya yang perlu diabaikan.

$$R_t = \phi(X_t\, W_{xr} + H_{t-1} W_{hz} + b_r) \quad (7)$$

Selanjutnya, *update gate* (Z_t) mengontrol proporsi antara informasi lama dan informasi baru yang digunakan untuk memperbarui stat

$$Z_t = \phi(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \quad (8)$$

Informasi baru yang telah dikalkulasi disebut sebagai *candidate hidden state*, diperoleh dengan mempertimbangkan hasil dari *reset gate* terhadap *hidden state* sebelumnya.

$$\tilde{H}_t = \sigma(X_t W_{xh} + (R_t \odot H_{t-1}) W_{hh} + b_h) \quad (9)$$

Akhirnya, *hidden state* saat ini dihitung dengan menggabungkan *hidden state* sebelumnya dan kandidat berdasarkan bobot *update gate*

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (10)$$

BiGRU merupakan pengembangan dari GRU yang mengatasi keterbatasan pemrosesan satu arah. BiGRU menggunakan dua lapisan GRU yang berjalan paralel, satu memproses urutan data secara maju (*forward*) dan satu secara mundur (*backward*). Hasil dari kedua arah ini digabungkan untuk membentuk representasi fitur yang lebih utuh dan kontekstual terhadap data sekuensial [1].

Pendekatan ini membuat BiGRU unggul dalam memahami konteks secara menyeluruh, terutama pada tugas-tugas seperti pemrosesan bahasa alami dan prediksi berbasis waktu. Meski membutuhkan waktu pelatihan lebih lama dibanding GRU biasa, BiGRU mampu memberikan akurasi yang lebih tinggi dalam

berbagai aplikasi karena kekayaan informasi dari kedua arah pemrosesan [6][7].

2.6. Twitter sebagai Sumber Data

Twitter merupakan salah satu platform media sosial yang digunakan secara luas di seluruh dunia, termasuk di Indonesia. Menurut laporan dari We Are Social dan Hootsuite, pengguna aktif Twitter di Indonesia mencapai lebih dari 20 juta pengguna. Platform ini memungkinkan pengguna untuk mengekspresikan opini, berbagi informasi, dan berinteraksi dalam bentuk teks singkat yang dikenal sebagai cuitan. Karakteristik waktu nyata dan kemampuannya menyebarkan informasi dengan cepat menjadikan Twitter sebagai sumber data yang sangat potensial dalam penelitian analisis sentimen.

3. METODE PENELITIAN

Penelitian ini dilakukan menggunakan pendekatan *supervised deep learning* dengan membandingkan dua arsitektur jaringan saraf tiruan yaitu Bidirectional Long Short-Term Memory (BiLSTM) dan Bidirectional Gated Recurrent Unit (BiGRU). Dataset yang digunakan merupakan kumpulan tweet berbahasa Inggris mengenai pandemi COVID-19 yang diperoleh dari platform Kaggle, dan telah dilabeli secara manual ke dalam lima kelas sentimen: Positive, Negative, Extremely Positive, Extremely Negative, dan Neutral.

3.1. Pengumpulan Data

Data dikumpulkan dari dataset publik "COVID-19 NLP Text Classification" yang tersedia di Kaggle. Dataset ini terdiri dari kolom OriginalTweet sebagai input teks, dan kolom Sentiment sebagai label klasifikasi. Total data terdiri dari ribuan cuitan yang telah dilabeli dan mencerminkan sentimen publik terhadap pandemi COVID-19 pada awal tahun 2020.

3.2. Pra-Pemrosesan Data

Sebelum dilakukan pelatihan model, data melalui beberapa tahap pra-pemrosesan:

a. **Pembersihan Teks:** Menghapus URL, tag HTML, karakter non-ASCII, angka, tanda baca, mention, dan hashtag.
b. **Lowercasing:** Mengubah semua teks menjadi huruf kecil.
c. **Tokenisasi:** Memecah kalimat menjadi token atau kata.
d. **Stopwords Removal:** Menghapus kata-kata umum yang tidak memiliki makna signifikan.
e. **Lemmatization:** Mengubah kata ke bentuk dasarnya.
f. **Padding & Sequencing:** Teks dikonversi menjadi urutan angka menggunakan tokenizer dan disesuaikan panjangnya menggunakan teknik padding.

3.3. Transformasi Data

Setelah teks dibersihkan, dilakukan transformasi data menggunakan teknik word embedding GloVe (Global Vectors for Word Representation). Embedding GloVe versi Twitter.27B.200d digunakan agar model dapat memahami makna kata dalam konteks media sosial. Setiap kata direpresentasikan dalam bentuk vektor berdimensi 200. Studi sebelumnya menunjukkan bahwa integrasi antara GloVe dan model deep learning seperti LSTM mampu meningkatkan akurasi klasifikasi sentimen, terutama pada data Twitter bertema COVID-19 [9].

3.4. Pemodelan Data

Dua arsitektur model digunakan dalam penelitian ini:

a. **BiLSTM Double Layer**: Model ini terdiri dari dua lapisan LSTM dua arah yang dilanjutkan dengan dense layer dan output layer dengan fungsi aktivasi softmax.
b. **BiGRU Double Layer**: Serupa dengan BiLSTM, tetapi menggunakan dua lapisan GRU dua arah. Kedua model menggunakan dropout, batch normalization, dan regularisasi L2 untuk menghindari *overfitting*.

Proses pelatihan dilakukan menggunakan optimizer AdamW dan fungsi aktivasi ReLU, dengan early stopping dan model checkpoint untuk menyimpan model terbaik.

3.5. Evaluasi Model

Evaluasi dilakukan dengan mengamati nilai akurasi pada data validasi serta membandingkan grafik training loss dan validation loss selama proses pelatihan. Model

dinilai berdasarkan stabilitas kinerjanya dalam beberapa percobaan, serta kemampuannya untuk menghindari *overfitting*.

Penggunaan teknik seperti early stopping dan reduce learning rate on plateau membantu menghentikan pelatihan saat performa model stagnan dan meningkatkan efisiensi proses pembelajaran. Selain itu, visualisasi grafik loss digunakan untuk memantau konvergensi model dan mendeteksi potensi *overfitting*.

Akurasi akhir pada data uji digunakan sebagai indikator utama untuk membandingkan performa antara model BiLSTM dan BiGRU. Hasil menunjukkan bahwa kedua model memberikan performa yang baik, dengan BiGRU memiliki kecenderungan lebih stabil pada proses pelatihan.

4. HASIL DAN PEMBAHASAN

Penelitian ini membandingkan performa dua model deep learning, yaitu BiGRU dan BiLSTM, dalam mengklasifikasikan kondisi daun tanaman berdasarkan citra. Eksperimen dilakukan dengan menguji beberapa konfigurasi model, mengamati akurasi, loss, dan efisiensi pelatihan. Hasil menunjukkan bahwa kedua model mampu mencapai akurasi tinggi, dengan perbedaan signifikan dalam efisiensi dan stabilitas pelatihan.

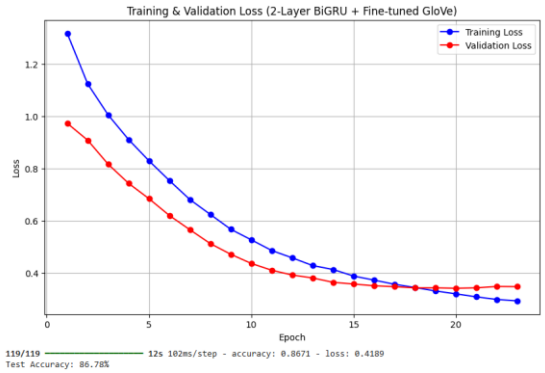
4.1 Hasil Eksperimen

Hasil pengujian menunjukkan bahwa model BiGRU menghasilkan akurasi pengujian tertinggi sebesar **86,78%** dengan *training loss* akhir sebesar 0,32 dan *validation loss* sebesar 0,41. Sementara itu, model BiLSTM mencatatkan akurasi pengujian sebesar **84,91%**, dengan *training loss* akhir 0,15 dan *validation loss* sebesar 0,49. Rangkuman hasil pengujian dari kedua model disajikan pada Tabel 1.

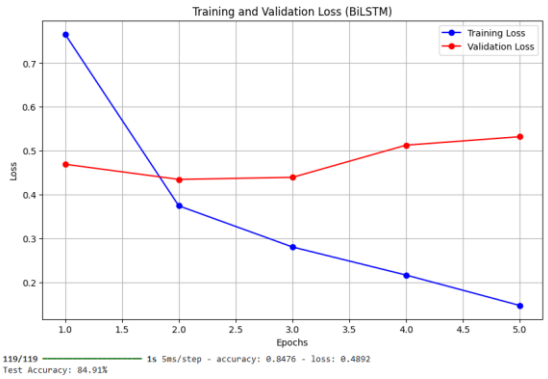
Tabel 1. Ringkasan Hasil Evaluasi Model

Model	Akurasi (%)	Akurasi Terakhir	Loss Terakhir	Overfitting
BiGRU	86,78	0,8671 (epoch 23)	0,4189 (epoch 23)	Tidak
BiLSTM	84,91	0,8476 (epoch 5)	0,4892 (epoch 5)	Ya

Visualisasi perbandingan nilai *loss* antara data pelatihan dan validasi dari masing-masing model ditunjukkan pada Gambar 1 dan Gambar 2.



Gambar 1. Grafik Training dan Validation Loss pada Model BiGRU



Gambar 2. Grafik Training dan Validation Loss pada Model BiLSTM.

Dari grafik tersebut terlihat bahwa model BiGRU menunjukkan tren penurunan yang konsisten pada kedua jenis *loss*, sedangkan pada model BiLSTM terjadi kenaikan *validation loss* setelah epoch ke-3, mengindikasikan overfitting.

Struktur arsitektur masing-masing model ditampilkan pada Gambar 3 dan Gambar 4.

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 50)	0
embedding (Embedding)	(None, 50, 200)	6,743,400
bidirectional (Bidirectional)	(None, 50, 256)	253,440
bidirectional_1 (Bidirectional)	(None, 256)	296,448
batch_normalization (BatchNormalization)	(None, 256)	1,024
dropout (Dropout)	(None, 256)	0
dense (Dense)	(None, 64)	16,448
dropout_1 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 3)	195

Total params: 7,310,955 (27.89 MB)
Trainable params: 7,310,443 (27.89 MB)
Non-trainable params: 512 (2.00 KB)

Gambar 3. Arsitektur Model BiGRU

Model: "functional_2"

Layer (type)	Output Shape	Param #
input_layer_2 (InputLayer)	(None, 50)	0
embedding_2 (Embedding)	(None, 50, 100)	3,371,600
bidirectional_2 (Bidirectional)	(None, 512)	731,136
dropout_2 (Dropout)	(None, 512)	0
dense_4 (Dense)	(None, 100)	51,300
dense_5 (Dense)	(None, 3)	303

Total params: 4,154,339 (15.85 MB)
Trainable params: 4,154,339 (15.85 MB)
Non-trainable params: 0 (0.00 B)

Gambar 4. Arsitektur Model BiLSTM

4.2 Pembahasan

Performa yang lebih unggul pada model BiGRU disebabkan oleh arsitektur dua lapis GRU, pemanfaatan pretrained embedding GloVe, serta penggunaan teknik regulasi seperti dropout, batch normalization, dan L2 regularization. Penggunaan optimizer AdamW dan callback seperti *ReduceLROnPlateau* serta *EarlyStopping* turut menjaga kestabilan selama pelatihan.

Sebaliknya, meskipun model BiLSTM menunjukkan akurasi awal yang baik, ia menunjukkan gejala overfitting lebih cepat, terutama karena peningkatan *validation loss* pada epoch selanjutnya. Hal ini menunjukkan bahwa BiLSTM kurang optimal dalam menangani generalisasi pada data yang belum terlihat sebelumnya.

Hasil ini sejalan dengan penelitian yang menemukan bahwa BiGRU memiliki kemampuan yang lebih baik dalam menangkap nuansa semantik dari data teks sosial. Hal ini menunjukkan bahwa BiGRU lebih sesuai digunakan pada tugas klasifikasi teks dengan konteks emosional atau opini publik, seperti selama pandemi COVID-19. [10][11]

4.3 Implikasi

Secara teoritis, hasil ini menegaskan bahwa BiGRU dapat menjadi alternatif yang lebih efisien dan akurat dibandingkan BiLSTM dalam tugas analisis sentimen. Dalam praktik, model BiGRU sangat relevan untuk digunakan pada sistem monitoring opini publik berbasis teks dalam skala besar, termasuk platform media sosial dan survei daring. Model ini juga cocok untuk implementasi pada perangkat dengan keterbatasan komputasi, karena efisiensi parameter yang lebih tinggi.

5. KESIMPULAN

1. Model BiGRU memberikan performa terbaik dalam tugas analisis sentimen publik terkait COVID-19 di Twitter, dengan akurasi tertinggi mencapai 86,78% dan stabilitas validasi yang konsisten.
2. Penggunaan komponen seperti GloVe pretrained embeddings, dropout, regularisasi, dan callbacks secara signifikan membantu dalam mengurangi overfitting serta meningkatkan kemampuan generalisasi model.
3. Meskipun BiLSTM juga menunjukkan akurasi yang kompetitif sebesar 84,91%, model ini memiliki kecenderungan overfitting pada epoch tertentu, sehingga memerlukan penyesuaian hyperparameter lanjutan.
4. Pendekatan kolaboratif dalam tim terbukti mempercepat proses tuning dan mengatasi hambatan teknis, yang berkontribusi positif terhadap efisiensi pengembangan model.
5. Kelebihan penelitian ini adalah berhasil menunjukkan bahwa BiGRU lebih stabil untuk data analisis sentimen publik, serta efektifnya teknik regulasi untuk meningkatkan akurasi.
6. Kekurangan utama terletak pada keterbatasan jumlah data yang digunakan serta kebutuhan waktu dan sumber daya yang tinggi dalam proses tuning model.
7. Pengembangan selanjutnya dapat difokuskan pada:
   a. Pengujian dengan dataset berukuran lebih besar dan lebih beragam secara topik maupun bahasa.

b. Eksperimen dengan arsitektur model hybrid seperti kombinasi BiGRU-CNN atau attention mechanism.
c. Pemanfaatan pendekatan AutoML atau Bayesian Optimization untuk tuning hyperparameter yang lebih efisien dan sistematis.

DAFTAR PUSTAKA

[1] Zhu, Q., Jiang, X., & Ye, R. (2021). Sentiment analysis of review text based on BiGRU-attention and hybrid CNN. *IEEE Access, 9*, 149077–149090. https://doi.org/10.1109/ACCESS.2021.3118537

[2] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, 3rd ed., draft ed., Jan. 12, 2025.

[3] B. Liu, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, May 2012.

[4] A. F. Rizki, W. Prihartono, and Fathurrohman, "Analisis sentimen ulasan Google Maps Rumah Sakit Khalishah di Cirebon dengan algoritma Naive Bayes," *J. Inform. dan Tek. Elektro Terapan (JITET)*, vol. 13, no. 2, p. 732, 2023. [Online]. Available: https://doi.org/10.23960/jitet.v13i2.6309

[5] Pranida, S. Z., & Kurniawardhani, A. (2022). Sentiment analysis of expedition customer satisfaction using BiGRU and BiLSTM. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM), 5*(1), 44–53. https://doi.org/10.24014/ijaidm.v5i1.17361

[6] S. Siami-Namini, N. Tavakoli, dan A. S. Namin, "A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM," 21 November 2019, *arXiv*: arXiv:1911.09512. doi: 10.48550/arXiv.1911.09512. https://arxiv.org/pdf/1911.09512

[7] J. K. Giustino dan Y. P. Santosa, "TOXIC COMMENT CLASSIFICATION COMPARISON BETWEEN LSTM, BILSTM, GRU, AND BIGRU," *Proxies J. Inform.*, vol. 7, no. 2, hlm. 115–127, Agu 2024. https://journal.unika.ac.id/index.php/proxies/article/view/12471

[8] S. Rahmadani, C. S. Rahayu, A. Salim, and K. N. Cahyo, "Deteksi Emosi Berdasarkan Wicara Menggunakan Deep Learning Model," JINTEKS (Jurnal Informatika Teknologi dan Sains), vol. 4, no. 3, pp. 220–224, 2022. https://jurnal.uts.ac.id/index.php/JINTEKS/article/view/1952

[9] C. K. Poetra, S. F. Pane, and R. N. S. Fatonah, "Meningkatkan Akurasi Long Short-Term Memory (LSTM) pada Analisis Sentimen Vaksin Covid-19 di Twitter dengan Glove," Jurnal Telematika, vol. 16, no. 2, pp. 85–90, 2023. https://journal.ithb.ac.id/telematika/article/view/400

[10] Zhang, J., Li, Z., Zhou, L., Yang, X., Jia, H., & Li, W. (2023). User sentiment analysis of COVID-19 via adversarial training based on the BERT-FGM-BiGRU model. *Systems, 11*(3), 129. https://doi.org/10.3390/systems11030129

[11] Xiaoyan, L., Raga, R. C., & Xuemei, S. (2022). GloVe-CNN-BiLSTM model for sentiment analysis on text reviews. *Journal of Sensors, 2022*, Article 7212366. https://doi.org/10.1155/2022/7212366