ZONAsi Jurnal Sistem Informasi
ZONAsi VOL. 7 NO. 1 JANUARI 2025
Page: 74 - 84
ISSN: 2656-7407 (Online)
2656-7393 (Print)

KLASIFIKASI SENTIMEN MENGGUNAKAN BIDIRECTIONAL LSTM DAN INDOBERT DENGAN DATASET TERBATAS

Ridho Illahi [1], Surya Agustian[2], Jasril[3], Febi Yanto[4]
[1,2,3,4]Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
Jl. HR. Soebrantas, Simpang Baru, Kota Pekanbaru, Riau, Telp. 0761 562223
e-mail: [1]12050117551@students.uin-suska.ac.id, [2] surya.agustian@uin-suska.ac.id, [3] jasril@uin-suska.ac.id, [4] febiyanto@uin-suska.ac.id

Abstrak

Untuk mendapatkan informasi melalui media sosial mengenai popularitas termasuk juga sentimen masyarakat terhadap para tokoh politik, masalah waktu menjadi krusial. Oleh karena itu, menganalisis hal tersebut menggunakan komputer menjadi pilihan. Namun menyediakan data yang berlabel cukup untuk pembelajaran mesin dalam kasus analisis sentimen dan tingkat kepopuleran dari media sosial, akan menyebabkan waktu yang dibutuhkan menjadi lebih lama. Penelitian ini mengembangkan model Bi-LSTM (Bidirectional Long Short-Term Memory) yang dipadukan dengan IndoBERT sebagai fitur representasi teks, untuk menganalisis sentimen dengan data training yang terbatas. Isu utama mengenai pengangkatan Kaesang sebagai ketua umum Partai Solidaritas Indonesia (PSI), hanya menyediakan 300 – 600 tweet berlabel positif, negatif dan netral untuk training. Data tweet setelah melalui tahap preprocessing dan vektorisasi BERT, dilatih pada metode klasifikasi Bi-LSTM. Langkah-langkah optimasi dilakukan sampai kepada hyperparameter tuning untuk menemukan model terbaik. Meski jumlah data train terbatas, model yang diusulkan berhasil mencapai F1-score sebesar 71% pada data validasi, dan 59% pada pengujian data test. Hasil penelitian mendemonstrasikan keefektifan metode Bi-LSTM dan fitur IndoBERT dalam mengukur opini masyarakat di media sosial. Framework metode klasifikasi yang diusulkan sangat potensial sebagai alat untuk menganalisis respons masyarakat terhadap isu politik terkini.

Kata kunci: *analisis sentimen, Bi-LSTM, IndoBERT, dataset terbatas*

Abstract

Time becomes a crucial factor in order to gather information from social media, regarding the popularity and sentiment towards political and public figures. Therefore, analyzing such matters using computers is the preferred approach. However, providing sufficient labeled data for machine learning in sentiment analysis and popularity measurement from social media can be time-consuming. This study developed a Bi-LSTM (Bidirectional Long Short-Term Memory) model combined with IndoBERT for text feature representation to analyze sentiment with limited training data. The main issue, regarding the appointment of Kaesang as the Chairman of the Indonesian Solidarity Party (PSI), only provided 300-600 labeled tweets (positive, negative, and neutral) for training. After preprocessing and BERT vectorization, the tweet data were trained using the Bi-LSTM classification method. Optimization steps, including hyperparameter tuning, were carried out to find the best model. Despite the limited training data, the proposed model achieved an F1-score of 71% on validation data and 59% on test data. The study demonstrates the effectiveness of the Bi-LSTM method and IndoBERT features in measuring public opinion on social media. The proposed classification framework shows great potential as a tool for analyzing public responses to current political issues.

Keywords: *sentiment analysis, Bidirectional LSTM, IndoBERT, limited dataset*

1. PENDAHULUAN

Media sosial telah menjangkau hampir separuh penduduk Indonesia, dengan 130 juta individu aktif menggunakan platform digital dari total populasi mencapai 256,4 juta jiwa. Secara persentase, sekitar 49 persen warga negara Indonesia terlibat secara aktif dalam berbagai platform media sosial [1]. Salah satunya media sosial yang cukup ramai digunakan yaitu Twitter, Twitter adalah platform jejaring sosial yang memungkinkan penggunanya berinteraksi melalui pesan singkat yang dikenal sebagai "Tweet." Melalui tweet ini, pengguna Twitter dapat berkomunikasi lebih dekat dengan sesama pengguna [2]. Hal ini menciptakan peluang besar untuk memahami pandangan masyarakat melalui analisis sentimen, yang merupakan proses mengidentifikasi opini positif, negatif, atau netral dari teks yang diunggah di media sosial.

Pada tahun 2024, menjelang pemilihan umum yang dijadwalkan pada bulan Februari, perhatian terhadap isu politik meningkat secara signifikan, terutama terkait pengangkatan Kaesang sebagai Ketua Umum PSI. Kaesang Pangarep, putra bungsu Presiden RI Joko Widodo, resmi dilantik sebagai Ketua Umum Partai Solidaritas Indonesia (PSI) pada acara Kopdarnas PSI yang berlangsung di Djakarta Theater, Jakarta Pusat, pada Senin, 25 September 2023 [3]. Pandangan dan respons masyarakat terhadap pengangkatan Kaesang sebagai Ketua Umum partai ini mengalami perkembangan pesat di berbagai media sosial, termasuk Twitter. Platform ini menjadi tempat utama bagi banyak orang untuk menyampaikan pendapat dan tanggapan mereka mengenai peran Kaesang sebagai Ketua Umum PSI [4]. Analisis sentimen, atau yang dikenal juga sebagai opinion mining, adalah proses otomatis untuk memahami, mengekstrak, dan mengolah data teks guna memperoleh informasi mengenai sentimen yang terkandung dalam sebuah kalimat opini. Pentingnya analisis sentimen dalam konteks politik seperti ini adalah untuk memahami opini publik secara lebih mendalam. Dengan menggunakan analisis sentimen, kita dapat melihat kecenderungan emosi dan persepsi masyarakat terhadap suatu isu, tokoh, atau peristiwa tertentu. Pada kasus pengangkatan Kaesang sebagai Ketua Umum PSI, analisis sentimen digunakan untuk memahami pendapat atau kecenderungan opini seseorang terhadap suatu isu atau objek tertentu [5], khususnya di Twitter, yang kemudian dapat dikategorikan ke dalam sentimen positif, negatif, atau netral.

Teknologi model bahasa berbasis Transformers, dengan fokus khusus pada BERT (*Bidirectional Encoder Representations From Transformer*), telah menghasilkan terobosan signifikan dalam bidang pemrosesan bahasa alami (NLP). Dalam periode waktu terkini, performa BERT menunjukkan kemampuan yang mengagumkan, terutama pada ranah analisis sentiment [6]. Kemampuan unik BERT dalam menangkap konteks kata secara dua arah memungkinkannya menyelami nuansa sentimen teks dengan akurasi tinggi. Model ini mampu memahami makna kata dari berbagai perspektif secara simultan. [7]. IndoBERT adalah model BERT yang telah dilatih khusus untuk bahasa Indonesia, menggunakan korpus bahasa Indonesia yang besar, terdiri dari empat miliar kata dalam proses pelatihan awalnya [8]. *Bidirectional* LSTM (*Bidirectional Long Short-Term Memory*) merupakan tumpukan dari LSTM (*Long Short-Term Memory*), di mana lapisan forward digunakan untuk menangkap konteks informasi sebelumnya, sementara lapisan backward digunakan untuk menangkap konteks informasi setelahnya [9]. bukan hanya satu pada urutan input: satu LSTM untuk urutan input asli, sementara yang lainnya untuk salinan urutan input yang telah dibalik. Dengan memberikan konteks tambahan bagi jaringan, pendekatan ini memungkinkan pembelajaran yang lebih cepat dan lebih menyeluruh dalam menyelesaikan permasalahan yang ada [10]. Menurut Hadab Khalid Obayes dkk[11], *Bidirectional* LSTM terbukti sangat bermanfaat dalam kasus yang membutuhkan konteks untuk input, seperti pada klasifikasi sentiment .

Analisis sentimen yang dilakukan oleh Lenggo Geni [12] pada tahun 2024, yang menggunakan model IndoBERT serta perbandingan dengan model berbasis machine learning dan lexicon (TextBlob, Multinomial *Naïve Bayes*, dan *Support Vector Machine*), hasil dari eksperimen ini dengan menggunakan model IndoBERT large-p1 memiliki kinerja terbaik dengan akurasi sebesar 83,5%. Model ini meningkatkan sistem baseline sebesar 48,5% dan 46,49% untuk TextBlob, 2,5% dan 14,49% untuk Multinomial *Naïve Bayes*, serta 3,5% dan 13,49% untuk Support Vector Machine dalam hal akurasi dan skor F-1. Dengan metode *Bidirectional* LSTM dibandingkan dengan metode LSTM dan hasil penelitian menunjukkan bahwa *Bidirectional* LSTM lebih unggul dibandingkan LSTM, dengan akurasi terbaik sebesar 91% dan training loss sebesar 28%, penelitian ini telah dilakukan oleh Dloifur Rohman Alghifari pada tahun 2022 [13]. Penelitian selanjutnya dilakukan oleh Gregorius Airlangga pada tahun 2024 dengan menggunakan metode RNN, LSTM, dan *Bidirectional* LSTM untuk

mendeteksi berita palsu dalam klasifikasi biner, hasil penelitian menunjukkan bahwa ketiga model mencapai akurasi tinggi sekitar 91%, dengan perbedaan kecil pada *precision* dan *recall* [14].

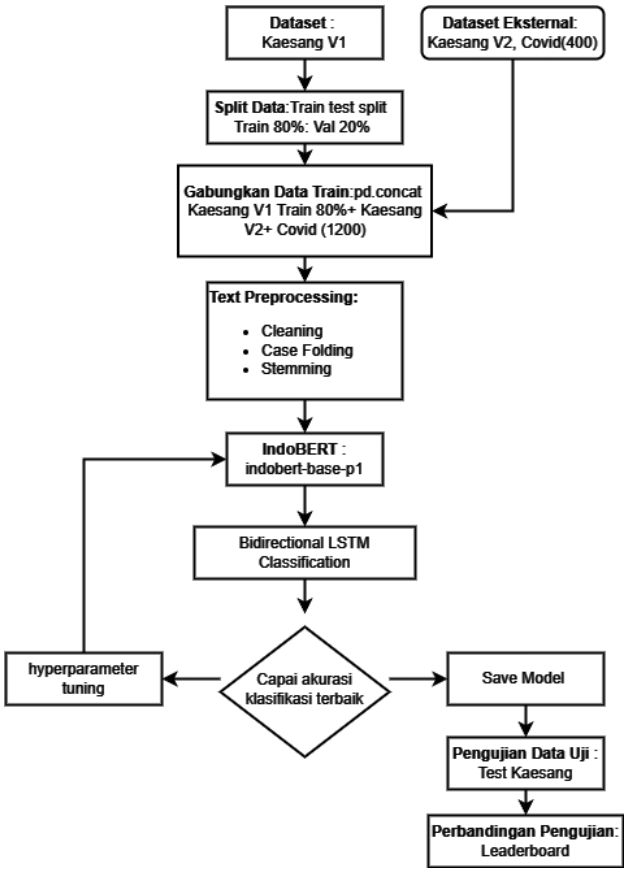
Penelitian klasifikasi sentimen tentang shared task telah dipaparkan dalam [15]. ini menitikberatkan pada permasalahan klasifikasi dengan keterbatasan data training, dimana hanya tersedia 300 tweet yang dinilai belum memadai untuk mendapatkan hasil yang maksimal. Fokus penelitian diarahkan pada pengkajian respon masyarakat terkait kontroversi naiknya Kaesang menjadi Ketua PSI di Twitter, dengan memanfaatkan pendekatan *Bidirectional* LSTM yang mengintegrasikan proses tokenisasi dan padding berbasis IndoBERT. Koleksi tweet yang terkumpul kemudian dievaluasi dan dikelompokkan menjadi tiga kategori sentimen: positif, negatif, atau netral berkaitan dengan kejadian tersebut. Pengolahan data dilaksanakan menggunakan keyword yang berkaitan dengan pengangkatan Kaesang, dengan penerapan metodologi *Bidirectional* LSTM dan IndoBERT dalam prosesnya.

2. METODE PENELITIAN

Dalam penelitian ini, digunakan pendekatan metodologis yang tersusun sistematis untuk menjamin validitas dan reliabilitas hasil penelitian. Rangkaian proses ini mencakup beberapa fase penting, dimulai dari perencanaan penelitian, penyiapan dataset, pemrosesan awal teks, implementasi *Bidirectional* LSTM bersamaan dengan proses vektorisasi menggunakan IndoBERT, hingga fase evaluasi model.

2.1. Tahapan Penelitian

Tahapan penelitian menguraikan serangkaian langkah sistematis yang akan ditempuh dalam menyelesaikan penelitian ini. Urutan tahapan tersebut dapat diamati melalui ilustrasi yang disajikan pada gambar 1 di bawah ini.



Gambar 1. Tahapan Peneltian

Sesuai dari gambar 1. Tahap Peneltian, Penelitian ini melibatkan berbagai tahapan, yang terdiri dari persiapan Dataset, Praproses Data, proses pengubahan teks menjadi vektor dengan IndoBERT,

implementasi metode klasifikasi *Bidirectional* LSTM, hyperparameter tuning, pengujian pada data uji, hingga perbandingan pengujian . Detail lengkap mengenai setiap tahapan akan dibahas pada bagian selanjutnya.

2.2. Dataset

Penelitian ini menggunakan dua kumpulan data yang diperoleh dari Twitter melalui metode *crawling*. Kumpulan data pertama berisi sentimen masyarakat terhadap pengangkatan Kaesang sebagai ketua PSI (Dataset Kaesang), sedangkan kumpulan kedua berisi sentimen terkait program vaksinasi Covid-19 (Dataset Covid). Dataset Kaesang menjadi perhatian utama penelitian karena dirancang khusus untuk mengkaji efektivitas pembelajaran mesin dengan data terbatas.

Dataset Kaesang diambil secara spesifik dalam periode 25 September hingga 3 Oktober 2023, menggunakan kata kunci 'Kaesang PSI. Proses pelabelan sentimen dilaksanakan melalui crowd sourcing dengan minimal empat anotator per tweet. Penentuan label (positif, netral, atau negatif) didasarkan pada suara terbanyak, dimana tweet tanpa label mayoritas dieliminasi dari dataset.

Untuk keperluan pelatihan, tersedia dua Dataset Kaesang V1, berisi 300 tweet terlabel. Data pengujian (Test Kaesang) terdiri dari 924 tweet dengan label gold standard yang tersimpan di server *leaderboard* [15]. Model terbaik dari penelitian akan diuji menggunakan data ini, dan hasil prediksinya akan diunggah ke sistem *leaderboard* untuk evaluasi.

Sebagai data eksternal, Dataset Kaesang V2 berjumlah 300 tweet terlabel dan Dataset Covid yang berasal dari penelitian sebelumnya[16], [17], mencakup 8.000 tweet berlabel dan akan dimanfaatkan sebagai data pelatihan tambahan untuk Dataset Kaesang, penelitian ini hanya menggunakan 400 tweet disetiap label positif, negative dan netral dengan jumlah tweet yang digunakan 1200 tweet pada Dataset Covid.

Tabel 1. Rincian Dataset

Dataset	Jumlah Tweet
Kaesang V1	300
Kaesang V2	300
Test Kaesang	924
Data Covid-19	1200

Berdasarkan informasi yang disediakan pada Tabel 1, penelitian ini hanya menggunakan satu dataset pelatihan, yaitu Kaesang V1 yang digunakan baik untuk pelatihan maupun validasi. Selanjutkan Dataset Kaesang V1 akan dibagi menjadi 80% data train 20% data validasi pembagian data menggunakan train test split dari library scikit-learn, dataset Train Kaesang V1 yang telah dibagi menjadi 80%, akan digabungak an untuk menambah jumlah data train dengan dataset KaesangV2 dan Data Covid menggunakan pd.concat dari pandas. Kemudian, dataset uji Kaesang digunakan sebagai data pengujian untuk mendapatkan skor evaluasi pada sistem peringkat (*leaderboard*).

2.3. *Text Preprocessing*

Dalam konteks pengolahan bahasa alami, menurut Ramadhansyah serangkaian tahapan dalam membersihkan dan mempersiapkan data teks merupakan bagian dari *text preprocessing* pada analisis sentimen , yang bertujuan agar data dapat diolah secara optimal oleh algoritma analisis sentimen [18]. Prosesnya meliputi pengubahan teks ke huruf kecil, pembersihan data, stemming, dan penanganan karakter khusus, sebelum dapat digunakan untuk klasifikasi, meskipun hanya menggunakan 300 tweet

untuk pelatihan, peneliti berusaha memaksimalkan hasil dengan menerapkan teknik *preprocessing* yang cermat.

a. Cleaning

Penenitian ini mengguanakan metode *Regex* dalam Cleaning data, *Regex* adalah rangkaian teks yang menggambarkan pola yang digunakan mesin regex untuk menemukan teks (atau posisi) dalam isi teks, biasanya untuk tujuan memvalidasi, menemukan, mengganti, atau membagi.

b. Case Folding

Case Folding adalah salah satu langkah dalam *preprocessing* teks yang bertujuan untuk menyederhanakan teks dengan mengubahnya menjadi format yang seragam, tanpa mengubah maknanya dalam melakukan *Case Folding* menggunakan *.lower()*.

c. Stemming

Stemming adalah proses memisahkan kata-kata yang memiliki awalan dan akhiran sehingga menjadi bentuk kata dasarnya, penellitain ini menggunakan stemming dari Sastrawi khusus Bahasa Indonesia .

Tabel 2.Hasil Text Preprocessing

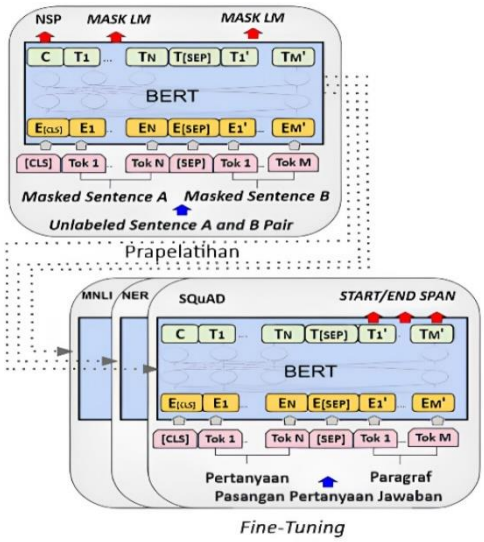
No	Proses	Sebelum proses	Setelah proses
1	Cleaing	Dukungan Capres 2024 PSI, Giring Ganesha: Tunggu Bro Kaesang! https://t.co/Z2zN6cqwbC https://t.co/rvelyGmuCE	Soal Dukungan Capres PSI Giring Ganesha Tunggu Bro Kaesang
2	Case Folding		soal dukungan capres psi giring ganesha tunggu bro kaesang
3	Stemming		soal dukung capres psi giring ganesha tunggu bro kaesang

Untuk hasil dari Text Preprocessing yang dilakukan dapat dilihat pada tabel atas ini, yang telah melewati 3 tahapan *Text Preprocessing* yaitu *Cleaning*, *Case Folding*, *Stemming*.

2.4. IndoBERT

Dalam penelitian ini, peneliti menggunakan model IndoBERT base p1 dari indobenchmark dengan fokus khusus pada proses Word Embedding dan tokenisasi. Implementasi dilakukan menggunakan tokenizer BertTokenizer dan model TFBertModel, keduanya dengan metode from_pretrained('indobenchmark/indobert-base-p1').

Proses tokenisasi IndoBERT mencakup fungsi tokenize_and_pad() yang berperan penting dalam encoding teks menjadi input_ids, pembuatan attention masks, serta padding dan truncation sesuai max_length yang ditentukan.



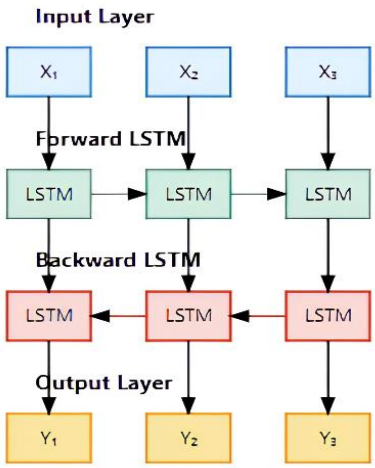
Gambar 2. Ilustrasi pre-training dan fine-tuning model BERT

Gambar 2 mengilustrasikan proses pre-training dan fine-tuning model BERT, yang melibatkan dua tahap utama: pre-training dengan menggunakan dua kalimat yang disamarkan (A dan B), dan fine-tuning untuk berbagai tugas NLP seperti MNLI, NER, SQuAD, dan SPAN.

Model IndoBERT base p1 memiliki karakteristik khusus, di antaranya dilatih pada korpus bahasa Indonesia, mampu menghasilkan representasi semantik yang akurat, dan mendukung pemodelan konteks bahasa Indonesia dengan presisi tinggi. Penggunaan model ini bertujuan mengoptimalkan ekstraksi fitur semantik dan kontekstual dari teks berbahasa Indonesia dalam proses analisis sentimen, dengan kemampuan memahami nuansa dan konteks bahasa lokal secara mendalam.

2.5. Klasifikasi dengan *Bidirectional* LSTM

Dalam penelitian ini, klasifikasi sentimen dilakukan menggunakan arsitektur *Bidirectional* LSTM yang dirancang secara komprehensif untuk mengekstraksi dan menganalisis fitur semantik dari teks berbahasa Indonesia. Pada arsitektur model, implementasi *Bidirectional* LSTM dilakukan setelah ekstraksi fitur dari lapisan embedding IndoBERT. Lapisan *Bidirectional* LSTM dikonfigurasi untuk mengembalikan urutan terakhir (return_sequences=False), yang memungkinkan model untuk menangkap representasi kontekstual dan temporal dari urutan token dalam teks.



Gambar 3. Arsitektur Bidirectional LSTM

Gambar 3 menggambarkan arsitektur *Bidirectional* LSTM yang digunakan dalam penelitian ini. Diagram tersebut menunjukkan proses kompleks di mana input (X1, X2, X3) diproses melalui dua lapisan LSTM paralel: lapisan maju (forward LSTM) yang memproses urutan dari kiri ke kanan, dan

lapisan mundur (backward LSTM) yang memproses urutan dari kanan ke kiri. Struktur ini memungkinkan model untuk menangkap konteks linguistik dari kedua arah, memberikan pemahaman yang lebih mendalam tentang nuansa makna dalam teks.

Dengan menggunakan parameter *Bidirectional*(), model dapat mengeksplorasi konteks kata dari dua arah - maju dan mundur - sehingga mampu menangkap dependensi kontekstual yang kompleks dalam data teks berbahasa Indonesia. Mekanisme ini memungkinkan jaringan untuk memproses informasi dari kedua arah input sequence, memberikan pemahaman yang lebih mendalam tentang konteks dan nuansa bahasa.

Setelah proses BiLSTM, dilakukan dropout untuk mencegah overfitting, yang selanjutnya disambung dengan lapisan dense bertingkat untuk klasifikasi akhir. Teknik dropout ini berperan penting dalam mencegah model terlalu bergantung pada fitur spesifik dalam data training, sehingga meningkatkan kemampuan generalisasi model. Arsitektur ini memungkinkan model untuk secara efektif mengekstraksi fitur semantik dan memahami struktur kompleks dalam tweet, yang pada akhirnya meningkatkan akurasi dalam klasifikasi sentimen.

2.6. Hyperparameter Tuning

Hyperparameter tuning adalah proses mengoptimalkan parameter model untuk meningkatkan performanya. Dalam penelitian ini, berfokus pada beberapa hyperparameter utama :

a. Learning Rate: Menentukan seberapa besar langkah yang diambil saat memperbarui bobot. Nilai yang diuji termasuk 0,00001, dan 0,000001.
b. *Batch Size*: Jumlah sampel yang diproses sebelum model diperbarui. Menguji ukuran batch 32, dan 64.
c. *Dropout Rate*: Proporsi neuron yang dihilangkan untuk mencegah overfitting. Nilai yang digunakan adalah 0.2, dan 0.3.
d. LSTM *Units*: Jumlah unit dalam layer LSTM. Mengeksplorasi konfigurasi dengan 64 dan 128.
e. Max Length: Panjang maksimum token yang digunakan saat tokenisasi, yang diatur pada 32 dan 64.

2.7. Pengujian pada data tes kaesang

Pada serangkaian data uji yang terdiri dari 924 tweet dan tidak pernah disentuh selama proses pelatihan, yang dimana data tes kaesang tidak memiliki label, model klasifikasi menggunakan *Bidirectional* LSTM dan IndoBert diuji coba. Data tes kaesang akan dilakukan prediksi untuk masing-masing kategori (positif, negatif, dan netral), prediksi yang dilakukan diambil dari 3 hasil pada presisi, *recall*, dan F1-*score* tertinggi dari output Hyperparameter Tuning.

3. Hasil dan Pembahasan

Penelitian menampilkan komparasi antara performa model pada kondisi awal dan setelah proses optimasi dilaksanakan. Tahap baseline menggambarkan kinerja awal model sebagai titik referensi, sedangkan tahap selanjutnya menunjukkan peningkatan performa yang diperoleh melalui penyempurnaan parameter secara sistematis.

3.1. Tahapan Awal (Baseline)

Metode baseline merupakan pendekatan fundamental dalam klasifikasi, dirancang untuk memberikan gambaran awal kinerja model sebelum proses optimasi. Metodologi ini mencakup penggunaan dataset campuran dari kaesang v1, dan data eksternal kaesang v2, dan data covid-19 (1200), dengan tahapan prapemrosesan meliputi cleaning teks dan case folding. Tokenisasi dan word embedding dilakukan menggunakan IndoBERT, dengan konfigurasi hyperparameter spesifik: Learning rate 0,000001,*batch size* 32, *dropout rate* 0,3, LSTM unit 64, dan max length 64.

Evaluasi klasifikasi pada data validasi menggunakan metrik macro-average menghasilkan F1-*score* 0.52, akurasi 0.60, presisi 0,55, dan *recall* 0,65. Hasil baseline ini berfungsi sebagai benchmark awal, memberikan wawasan mengenai kemampuan dasar model dalam mengolah dataset yang digunakan. Selanjutnya, performa ini akan menjadi titik referensi untuk mengukur potensi peningkatan melalui optimasi dan penyesuaian hyperparameter. Tahap akhir melibatkan pengujian model pada

dataset tes kaesang untuk mengevaluasi kemampuan klasifikasi pada data baru, dengan hasil performa *Bidirectional* LSTM tercatat pada tabel 3 entri nomor 6 dan tabel 5 urutan pertama.

3.2. Hasil Tahap Hyperparameter Tuning

Setelah tahap baseline ditentukan, penelitian ini melangkah ke tahap optimasi melalui proses hyperparameter tuning. Fokus utama adalah mengidentifikasi kombinasi parameter yang mampu meningkatkan performa model melampaui baseline awal. Eksplorasi ini menghasilkan dua tabel komparatif yang membandingkan performa dengan learning rate berbeda, yakni 0,00001 dan 0,000001, guna menemukan kombinasi yang optimal dalam proses klasifikasi.

Tabel 3. Hasil Kombinasi Hyperparameter Tuning 0,00001

No	Learning Rate	*Batch Size*	*Dropout Rate*	LSTM *Units*	Max Length (Tokenizer)	F1 *Score*	Akurasi	Presisi	*Recall*
1	0,00001	32	0.2	64	32	0.39	0.40	0.40	0.41
2	0,00001	32	0.2	64	64	0.67	0.66	0.67	0.67
3	0,00001	32	0.2	128	32	0.61	0.61	0.62	0.62
4	0,00001	32	0.2	128	64	0.62	0.63	0.65	0.63
5	**0,00001**	**32**	**0.3**	**64**	**32**	**0.70**	**0.71**	**0.71**	**0.71**
6	**0,00001**	**32**	**0.3**	**64**	**64**	**0.68**	**0.69**	**0.69**	**0.69**
7	**0,00001**	**32**	**0.3**	**128**	**32**	**0.71**	**0.71**	**0.71**	**0.71**
8	0,00001	32	0.3	128	64	0.69	0.70	0.69	0.69
9	0,00001	64	0.2	64	32	0.62	0.61	0.62	0.62
10	0,00001	64	0.2	64	64	0.65	0.64	0.64	0.64
11	0,00001	64	0.2	128	32	0.65	0.65	0.65	0.65
12	0,00001	64	0.2	128	64	0.65	0.66	0.65	0.65
13	0,00001	64	0.3	64	32	0.60	0.58	0.58	0.59
14	0,00001	64	0.3	64	64	0.62	0.64	0.64	0.65
15	0,00001	64	0.3	128	32	0.62	0.62	0.62	0.62
16	0,00001	64	0.3	128	64	0.64	0.66	0.65	0.65

Berdasarkan tabel 3, hasil dari tahap hyperparameter tuning menunjukkan bahwa model *Bidirectional* LSTM yang diintegrasikan dengan IndoBERT melalui serangkaian eksperimen memperoleh 3 performa terbaik pada tabel Kombinasi hyperparameter Tuning 0,00001, tepatnya pada entri nomor 5, 6, dan 7

Tabel 4. Hasil Kombinasi Hyperparameter Tuning 0,000001

No	Learning Rate	*Batch Size*	*Dropout Rate*	LSTM *Units*	Max Length (Tokenizer)	F1 *Score*	Akurasi	Presisi	*Recall*
1	0,000001	32	0.2	64	32	0.35	0.36	0.36	0.37
2	0,000001	32	0.2	64	64	0.62	0.61	0.62	0.62
3	0,000001	32	0.2	128	32	0.56	0.56	0.57	0.57
4	0,000001	32	0.2	128	64	0.57	0.58	0.60	0.58
5	0,000001	32	0.3	64	32	0.65	0.66	0.66	0.66
6	0,000001	32	0.3	64	64	0.63	0.64	0.64	0.64
7	0,000001	32	0.3	128	32	0.66	0.66	0.66	0.66
8	0,000001	32	0.3	128	64	0.64	0.65	0.64	0.64
9	0,000001	64	0.2	64	32	0.57	0.56	0.57	0.57

10	0,000001	64	0.2	64	64	0.60	0.59	0.59	0.59
11	0,000001	64	0.2	128	32	0.60	0.60	0.60	0.60
12	0,000001	64	0.2	128	64	0.60	0.61	0.60	0.60
13	0,000001	64	0.3	64	32	0.55	0.53	0.53	0.54
14	0,000001	64	0.3	64	64	0.57	0.59	0.59	0.60
15	0,000001	64	0.3	128	32	0.57	0.57	0.57	0.57
16	0,000001	64	0.3	128	64	0.59	0.61	0.60	0.60

Berdasarkan tabel 4 hasil dari tahap hyperparameter tuning memperoleh 3 performa terbaik pada tabel Kombinasi hyperparameter Tuning 0,000001, tepatnya pada entri nomor 5, 6, dan 7, tetapi kombinasi tersebut masih belum mencapai metrik tertinggi dibanding tabel 3, mencakup f1 *score*, akurasi, presisi, dan *recall* dalam keseluruhan rangkaian kombinasi. Maka yang akan digunakan untuk tahap selanjutnya yaitu pengujian model pada data test Kaesang diambil dari tabel 3 pada entri nomor 5, 6, dan 7.

3.3. Hasil Pengujian Model pada Data Test Kaesang

Setelah didapatkanya 3 performa terbaik hasil kombinasi dari hyperparameter tuning, maka akan dilakukan pengujian pada model yang telah dilatih, menggunakan data tes kaesang. Hasilnya bisa dilihat pada tabel 5 dibawah.

Table 5. Hasil Pengujian Model pada Data Tes Kaesang

RUN	Dataset	*Batch Size*	Drop out Rate	LSTM *Units*	Max Legth	F1 *Score*	Akurasi	Presisi	*Recall*
1	Kaesang V1+V2+Covid (1200)	32	0.3	64	64	0.52	0.60	0.55	0.65
2	Kaesang V1+V2+Covid (1200)	32	0.3	64	32	0.54	0.63	0.56	0.64
3	**Kaesang V1+V2+Covid (1200)**	**32**	**0.3**	**128**	**32**	**0.59**	**0.67**	**0.58**	**0.66**

Dari tabel di atas terlihat bahwa kombinasi *Batch Size* 32, *Dropout* 0.3 , LSTM *units* 128, *Max Legth* 32 memberikan hasil terbaik dengan f1-*Score* sebesar 0.59, akurasi 0.67, presisi 0.58, dan *recall* 0.66.

3.4. Perbandingan Pengujian pada Sistem *Leaderboard*

Di sistem *leaderboard* penelitian, tersedia hasil dari berbagai metode yang diuji dalam tugas klasifikasi sentimen. Setelah melakukan hyperparameter tuning dan mendapatkan hasil pengujian menggunakan data tes kaesang dari model yang telah dikembangkan, langkah selanjutnya melibatkan perbandingan performa.

Tujuan utama dari perbandingan ini adalah mengevaluasi kinerja model yang dikembangkan dalam penelitian, dengan mengukur seberapa efektif model tersebut dibandingkan metode lain, khususnya dalam hal akurasi dan F1-*score*. Berikut akan disajikan tabel perbandingan performa model yang digunakan dalam penelitian ini dengan metode-metode lain yang tercantum dalam *leaderboard*.

Table 6. Perbandingan Pengujian pada Sistem *Leaderboard*

Tim Peneliti	Metode	F1 *Score*	Akurasi	Presisi	*Recall*
BERT[19]	BERT Clasification	0.60	0.62	0.65	0.63

Penelitian ini	***Bidirectional* LSTM IndoBERT**	**0.59**	**0.67**	**0.58**	**0.66**
Safrizal [4]	SVM Fasttext	0.53	0.62	0.53	0.59
Organizer [15]	SVM TF - IDF	0.51	0.61	0.52	0.59
Muhammad Ravil [3]	NB PSO	0.50	0.59	0.52	0.58
Admin[15]	Baseline	0.40	0.45	0.49	0.48

Dari tabel pengujian yang dianalisis, tampak pencapaian yang signifikan dalam klasifikasi sentimen. Model BERT Classification menonjol dengan performa paling unggul, berhasil meraih f1-*score* tertinggi sebesar 0.60, yang mengindikasikan kemampuannya dalam memahami konteks dan memberikan klasifikasi sentimen dengan akurasi akurat.

Metode *Bidirectional* LSTM dan IndoBERT yang telah dioptimasi pun menunjukkan kinerja dengan f1-*score* 0.59 yang berhasil melampaui metode dari tim peneliti lainya. Hasil ini menggambarkan keberhasilan pendekatan yang digunakan dalam mengeksplorasi sentimen pada data yang dianalisis.

3.5. Pembahasan

Berdasarkan penelitian yang telah dilakukan dengan upaya inovatif dalam menganalisis sentimen publik dengan memanfaatkan kombinasi metode *Bidirectional* LSTM dan IndoBERT, dan melalui proses *hyperparameter tuning*, melakukan pengujian data tes dan perbandingan pengujian guna mengkaji tanggapan masyarakat seputar pengangkatan Kaesang Pangarep sebagai Ketua Umum PSI dan metode yang optimal dalam melakukan klasifikasi dengan keterbatasan data *training*. Dalam proses penelitian, kami berhasil mengolah dengan keberhasilan model dalam mengidentifikasi dan mengklasifikasikan sentimen mencapai akurasi 71% pada data latih dan 59% pada data uji.

4. KESIMPULAN

Penelitian ini mengimplementasikan metode *Bidirectional* LSTM dan model IndoBERT untuk analisis sentimen masyarakat terhadap pengangkatan Kaesang Pangarep sebagai Ketua Umum Partai Solidaritas Indonesia (PSI) di Twitter. Dataset yang digunakan terdiri dari 1.524 tweet yang dikumpulkan antara 25 September hingga 3 Oktober 2023, yang diberi label positif, negatif, dan netral. Dari total dataset, 300 tweet digunakan sebagai data pelatihan dan validasi yaitu Kaesang V1 dan 924 tweet sebagai data pengujian (Test Kaesang). Untuk memperluas data pelatihan, juga digunakan 300 tweet Kaesang V2 dan 1200 tweet terkait vaksin COVID-19 dari penelitian sebelumnya. Pembagian data pelatihan dilakukan dengan rasio 80% untuk pelatihan dan 20% untuk validasi.

Model ini memanfaatkan IndoBERT, yang diadaptasi khusus untuk bahasa Indonesia, dalam proses tokenisasi dan representasi teks. Bidirectional LSTM diterapkan untuk menangkap konteks sekuensial dari kedua arah, meningkatkan pemahaman model terhadap sentimen. Proses hyperparameter tuning dilakukan untuk mengoptimalkan performa model, dengan fokus pada parameter seperti learning rate, *batch size*, *dropout rate*, dan jumlah unit LSTM. Hasil eksperimen menunjukkan bahwa model mencapai akurasi tertinggi 0.71 dan F1 *score* 0.71, dan pengujian pada data tes kaesang mencapai akurasi 0.67 dan F1 *Score* 0.59, membuktikan efektivitas kombinasi metode dalam memahami sentimen pada keterbatasan data *training.*

Daftar Pustaka

[1] D. S. Puspitarini and R. Nuraeni, "Pemanfaatan Media Sosial Sebagai Media Promosi," *J. Common*, vol. 3, no. 1, pp. 71–80, 2019, doi: 10.34010/common.v3i1.1950.

[2] R. Rinaldo, A. P. Sari, and E. Fardiana, "Digital Opinion #Puanadalahharapan Di Media Sosial Twitter Menggunakan Social Network Analysis," *J. Ilm. Multidisiplin*, vol. 2, no. 01, pp. 19–29, 2023, doi: 10.56127/jukim.v2i01.407.

[3] M. Ravil, S. Agustian, M. Fikry, and F. Insani, "Peningkatan Performa Klasifikasi Sentimen

Tweet Kaesang Menggunakan Naïve Bayes dengan PSO pada Dataset Kecil," *KLIK Kaji. Ilm. …*, vol. 4, no. 6, pp. 2909–2917, 2024, doi: 10.30865/klik.v4i6.1939.
[4] Safrizal, S. Agustian, A. Nazir, and Yusra, "Klasifikasi Sentimen Terhadap Pengangkatan Kaesang Sebagai Ketua Umum Partai PSI Menggunakan Metode Support Vector Machine," *Technol. Sci.*, vol. 6, no. 1, pp. 216–225, 2024, doi: 10.47065/bits.v6i1.5340.
[5] Muhammad Abdillah, M. Fikry, Yusra, A. Nazir, and F. Insani, "Analisa sentimen terhadap kenaikan bbm di twitter (x) menggunakan naive bayes classifier," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 5, no. 1, pp. 65–74, 2024, doi: 10.37859/coscitech.v5i1.6954.
[6] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Naacl-Hlt 2019*, no. Mlm, pp. 4171–4186, 2018, [Online]. Available: https://aclanthology.org/N19-1423.pdf.
[7] M. F. Ramadhan, B. Siswoyo, and S. I. Cirebon, "Mengenal Model BERT dan Implementasinya untuk Analisis Sentimen Ulasan Game," *Sist. Inf. dan Teknol.*, vol. Vol. 8 No., pp. 395–398, 2024, [Online]. Available: http://seminar.iaii.or.id/index.php/SISFOTEK/article/view/520/444.
[8] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," 2020, [Online]. Available: http://arxiv.org/abs/2009.05387.
[9] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 41–46, 2022, doi: 10.29207/resti.v6i1.3676.
[10] A. Aziz Sharfuddin, M. Nafis Tihami, and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification," *2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018*, pp. 1–4, 2018, doi: 10.1109/ICBSLP.2018.8554396.
[11] H. K. Obayes, F. S. Al-Turaihi, and K. H. Alhussayni, "Sentiment classification of user's reviews on drugs based on global vectors for word representation and bidirectional long short-term memory recurrent neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 1, pp. 345–353, 2021, doi: 10.11591/ijeecs.v23.i1.pp345-353.
[12] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.
[13] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 89–99, 2022, doi: 10.34010/jamika.v12i2.7764.
[14] G. Airlangga, "Advancing fake news detection: a comparative study of RNN, LSTM, and Bidirectional LSTM Architectures," vol. 16, no. 1, pp. 13–23, 2024.
[15] S. Agustian *et al.*, "New Directions in Text Classification Research: Maximizing The Performance of Sentiment Classification from Limited Data Arah Baru Penelitian Klasifikasi Teks: Memaksimalkan Kinerja Klasifikasi Sentimen dari Data Terbatas," *Arxiv*, pp. 1–10, 2024, [Online]. Available: https://arxiv.org/abs/2407.05627.
[16] M. Ihsan, Benny Sukma Negara, and Surya Agustian, "LSTM (Long Short Term Memory) for Sentiment COVID-19 Vaccine Classification on Twitter," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 13, no. 1, pp. 79–89, 2022, doi: 10.31849/digitalzone.v13i1.9950.
[17] I. M. S. Putra, Putu Jhonarendra, and Ni Kadek Dwi Rusjayanthi, "Deteksi Kesamaan Teks Jawaban pada Sistem Test Essay Online dengan Pendekatan Neural Network," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 6, pp. 1070–1082, 2021, doi: 10.29207/resti.v5i6.3544.
[18] D. Ramadhansyah, A. Asrofiq, and Y. Yunefri, "Analisis Sentimen Ulasan penumpang maskapai penerbangan di Indonesia… ZONAsi," *J. Sist. Inf.*, vol. 6, no. 2, pp. 287–297, 2024.
[19] J. Pranata, S. Agustian, and E. Haerani, "Penggunaan Model Bahasa indoBERT pada Metode Random Forest Untuk Klasifikasi Sentimen Dengan Dataset Terbatas," vol. 6, no. 3, pp. 1668–1676, 2024, doi: 10.47065/bits.v6i3.6335.