Analisis Sentimen Capres dan Cawapres pada Media Sosial X untuk Prediksi Pemenang Pemilu 2024 Menggunakan Algoritma BERT-BiLSTM

Tugas Akhir

Diajukan untuk memenuhi salah satu syarat memperoleh gelar sarjana

dari Program Studi S1 Informatika

Fakultas Informatika

Universitas Telkom

1301200325

Shalsabila Yasmin Ridwan

Program Studi Sarjana S1 Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2025

LEMBAR PENGESAHAN

ANALISIS SENTIMEN CAPRES DAN CAWAPRES PADA MEDIA SOSIAL X UNTUK PREDIKSI PEMENANG PEMILU 2024 MENGGUNAKAN ALGORITMA BERT-BiLSTM

Sentiment Analysis for Presidential and Vice Presidential Candidates' Winner Prediction in the 2024 Election on Social Media X Using BERT-BiLSTM Algorithm

NIM : 1301200325

Shalsabila Yasmin Ridwan

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana S1 Informatika

Fakultas Informatika

Universitas Telkom

Bandung, 16 Juli 2025

Menyetujui

Pembimbing I,

Dr. YULIANT SIBARONI, S.Si., M.T.

NIP : 00750036

Ketua Program Studi

Sarjana S1 Informatika,

MAHMUD DWI SULISTIYO, S.T., M.T., Ph.D.

NIP : 13880017

LEMBAR PERNYATAAN

Dengan ini saya, Shalsabila Yasmin Ridwan menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul "Analisis Sentimen Capres dan Cawapres pada Media Sosial X untuk Prediksi Pemenang Pemilu 2024 Menggunakan Algoritma BERT-BiLSTM" beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung risiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya.

Bandung, 16 Juli 2025

Yang Menyatakan

Shalsabila Yasmin Ridwan

Analisis Sentimen Capres dan Cawapres pada Media Sosial X untuk Prediksi Pemenang Pemilu 2024 Menggunakan Algoritma BERT-BiLSTM

Shalsabila Yasmin Ridwan[1] , Dr. Yuliant Sibaroni, S.Si., M.T.[2]

[1,2]Fakultas Informatika, Universitas Telkom Bandung
[1]syrltelkom@student.telkomuniversity.ac.id, [2]yuliant@telkomuniversity.ac.id

Abstrak

Pemilihan umum (pemilu) di Indonesia selalu menjadi topik menarik bagi seluruh pengguna media sosial termasuk media sosial X, atau yang sebelumnya dikenal dengan Twitter. Melalui postingan-postingan para pengguna, kita dapat melihat gambaran opini masyarakat: positif, negatif, atau netral–dengan melakukan analisis sentimen. Akan tetapi, dibutuhkan waktu dan usaha yang banyak untuk mengklasifikasi sentimen para pengguna dikarenakan banyaknya jumlah postingan yang ditemukan. Untuk melihat opini masyarakat secara efektif, diperlukan metode yang juga efektif dan efisien. Untuk memecahkan masalah ini, penelitian ini mengajukan metode klasifikasi teks menggunakan *Bidirectional Encoder Representations from Transformers* (BERT) untuk pemrosesan bahasa alami dalam rangka memudahkan komputer dalam memahami bahasa selayaknya manusia dengan algoritma *Deep Learning* BiLSTM. Dataset yang digunakan dalam penelitian ini adalah kumpulan postingan dari Twitter atau 'X' dari hasil *data crawling* menggunakan keyword nama dari masing-masing calon Presiden dan Wakil Presiden Pemilu 2024. Evaluasi algoritma dilakukan dengan menggunakan Stratified 10-Fold Cross Validation sehingga rata-rata dari hasil akurasi terbaik yang didapatkan dari analisis sentimen adalah 95,8% dengan nilai Precision 93,2%, Recall 92,4% dan F-1 92,2%

Kata Kunci : pemilu, X, analisis sentimen, BERT, BiLSTM, BERT-BiLSTM

Abstract

The general election (Pemilu) in Indonesia has always been an intriguing topic for all social media users, including on the platform X, previously known as Twitter. Through users' posts, we can gain an overview of public opinion—whether positive, negative, or neutral—by conducting sentiment analysis. However, classifying users' sentiments require a significant amount of time and effort due to the large volume of posts available. To effectively gauge public opinion, an effective and efficient method is needed. To solve this problem, this study proposes a text classification method using *Bidirectional Encoder Representations from Transformers* (BERT) for natural language processing to help the computer to understand language as humans do, in combination with the BiLSTM deep learning algorithm. The dataset used in this research is a collection of posts from Twitter or 'X' from the result of data crawling using the keyword the name of each Presidential and Vice Presidential candidates in the 2024 Election. The algorithm's evaluation was done with a Stratified 10-Fold Cross Validation so that the average of the best accuracy obtained from the algorithm is 95,8% with an average Precision score of 93,2%, Recall score of 92,4%, and an F-1 score of 92,2%.

Keywords : election, X, sentiment analysis, BERT, BiLSTM, BERT-BiLSTM

1. Pendahuluan

Latar Belakang

Indonesia merupakan negara yang menganut sistem demokrasi di mana kedaulatan tertinggi ada pada kekuasaan rakyat dengan menjunjung supremasi hukum sebagai negara hukum, sebagaimana ditegaskan dalam UUD 1945 [1]. Salah satu sarana demokrasi untuk mewujudkan sistem pemerintahan negara yang berkedaulatan rakyat sebagaimana diamanatkan oleh UUD 1945 adalah pemilihan umum. Proses demokrasi juga terwujud melalui prosedur pemilu untuk memilih wakil rakyat dan pejabat publik lainnya [2]. Pada tahun 2024 ini, Indonesia telah melakukan pemilu dalam memilih Presiden dan Wakil Presiden yang ke-8. Kemudian, pada tanggal 20 Maret 2024, Komisi Pemilihan Umum (KPU) telah menetapkan bahwa pemenang dari pemilu tahun 2024 adalah pasangan capres dan cawapres nomor urut 2, yaitu Prabowo Subianto dan Gibran Rakabuming [3]. Hasil pengumuman ini telah membuat ramai media sosial.

Proses untuk mengetahui jumlah sentimen positif, negatif, atau netral dalam topik tertentu di media sosial dipelajari dalam bidang analisis sentimen yang merupakan bidang studi yang menganalisis opini,

sentimen, evaluasi, penilaian, sikap, dan emosi terhadap sebuah entitas seperti produk, layanan, organisasi, orang, peristiwa, dan atribut-atributnya [4].

Sebuah penelitian sebelumnya telah melakukan analisis sentimen dengan menggunakan dataset postingan X mengenai pemilu tahun 2019 dan membandingkan algoritma *Machine Learning* dan *Deep Learning*. Untuk algoritma *Machine Learning*, penelitian tersebut mendapatkan hasil akurasi 84,04% untuk SVM, 82,74% untuk *Logistic Regression*, dan 82,08% untuk Multinomial Naïve Bayes di mana masing-masing metode menggunakan fitur TF-IDF. Untuk algoritma *Deep Learning*, penelitian tersebut mendapatkan hasil akurasi 84,20% untuk LSTM, 84,05% untuk CNN, 84,30% untuk CNN+LSTM, 84,50% untuk GRU+LSTM, dan hasil tertinggi didapat dengan menggunakan metode *Bidirectional* LSTM dengan akurasi sebesar 84,60% [5]. Sebuah penelitian lain yang menggunakan dataset yang sama juga telah melakukan analisis sentimen dengan menggabungkan metode *lexicon-based* dan SVM dan mendapatkan hasil akurasi sebesar 92,5% [6]. Selain kedua penelitian di atas, terdapat juga sebuah penelitian analisis sentimen YouTube reviews menggunakan IndoBERT dan mendapat akurasi sebesar 74% [7]. Ketiga penelitian di atas melakukan analisis sentimen untuk klasifikasi teks menjadi positif, negatif, atau netral dan belum dikaitkan untuk prediksi mengenai suatu masalah.

Mengenai analisis sentimen untuk prediksi, pada 2020, terdapat sebuah penelitian melakukan analisis sentimen untuk prediksi harga saham dengan membandingkan penggunaan BERT dan Word2Vec dengan menggunakan 3 dataset. Akurasi yang didapat adalah 95,20%, 98,46% dan 95,60% untuk masing-masing dataset menggunakan BERT dan 73,06%, 76,54%, dan 73,63% menggunakan Word2Vec [8]. Analisis sentimen juga dapat digunakan untuk memprediksi kepribadian Myers-Briggs pengguna X menggunakan *Machine Learning Classifiers* dan BERT berdasarkan tipe Myers-Briggs Personality Type Indicator (MBTI). Dengan membagi tipe kepribadian Myers-Briggs menjadi empat model, yaitu IE (Introverts/Extroverts), NS (Intuitives/Sensors), PJ (Perceivers/Judgers), dan FT (Feelers/Thinkers), didapatkan bahwa pada keempat model tersebut, SVM mencapai hasil tertinggi dengan akurasi 84,28%, 88,20%, 83,45%, dan 78,56% untuk masing-masing model [9]. Kemudian, pada 2022, sebuah penelitian melakukan analisis sentimen untuk memprediksi kejahatan menggunakan pendekatan hybrid. Penelitian tersebut menggunakan pendekatan analisis sentimen untuk mendeteksi ancaman keamanan secara *real-time* melalui platform media sosial X dengan menggabungkan metode BERT dalam NLP dan metode *lexicon-based* untuk analisis sentimen dan prediksi. Penelitian tersebut mencapai hasil akurasi sebesar 94,91% dalam memprediksi kejahatan berdasarkan data X [10].

Pada penelitian kali ini, peneliti bermaksud untuk melakukan analisis sentimen untuk memprediksi pemenang pemilu 2024 berdasarkan postingan di X. Selain itu, untuk meningkatkan akurasi metode lebih jauh lagi, peneliti menggabungkan algoritma BiLSTM dengan BERT, yaitu sebuah arsitektur neural network berbasis transformer yang sebelumnya telah dilatih menggunakan data yang banyak dan telah terbukti sangat efektif dalam melaksanakan berbagai tugas NLP termasuk analisis sentimen [11] dengan menggunakan BERT untuk *Word Embedding* sebelum diproses algoritma BiLSTM.

Topik dan Batasannya

Berdasarkan latar belakang di atas, penelitian ini mengangkat rumusan masalah sebagai berikut:

a. Bagaimana akurasi dari algoritma BERT-BiLSTM untuk analisis sentimen?

b. Bagaimana kebenaran hasil prediksi calon pemenang yang sebelumnya diprediksi menang berdasarkan klasifikasi analisis sentimen?

Agar penelitian lebih terfokus, Batasan masalah yang diambil adalah sebagai berikut:

a. Menggunakan dataset dari X

b. Hasil klasifikasi teks adalah positif, negatif, atau netral

c. Hasil prediksi adalah pasangan capres dan cawapres yang diprediksi akan menang berdasarkan analisis sentimen yang dilakukan

d. Topik yang digunakan adalah seputar pemilu 2024 seperti : Anies Baswedan, Muhaimin Iskandar, Prabowo Subianto, Gibran Rakabuming, Ganjar Pranowo, dan Mahfud MD.

e. Algoritma BERT digunakan sebagai *word embedding*.

Tujuan

Tujuan yang akan dicapai oleh penelitian ini berdasarkan rumusan masalah di atas adalah :

a. Analisis algoritma BiLSTM untuk proses klasifikasi postingan X menggunakan pendekatan BERT sebagai *Word Embedding*

b. Mengukur performansi algoritma dalam klasifikasi postingan X

c. Memprediksi pasangan capres dan cawapres mana yang sebelumnya diprediksi akan memenangkan pemilu 2024 berdasarkan model algoritma yang dihasilkan.

2. Studi Terkait

Dalam beberapa tahun terakhir, terdapat banyak penelitian berkaitan dengan klasifikasi teks untuk analisis sentimen dan penelitian-penelitian yang berkaitan dengan pemanfaatan analisis sentimen untuk memprediksi suatu masalah.

Dalam analisis sentimen yang dilakukan oleh Hidayatullah, dkk. [5], analisis sentimen dilakukan menggunakan algoritma *machine learning* SVM, *Logistic Regression*, dan Multinomial Naïve Bayes dengan menambahkan fitur TF-IDF pada masing-masing algoritma dan algoritma *Deep Learning* seperti LSTM, CNN, CNN+LSTM, GRU+LSTM, dan *Bidirectional* LSTM menggunakan dataset Twitter dengan topik pemilu 2019. Analisis Sentimen pada dataset yang sama juga dilakukan pada 2019 oleh Seno, Danar W., dan Arief Wibowo [6] dengan menggunakan algoritma SVM dan metode *lexicon-based*. Hasil menunjukkan bahwa SVM berbasis metode *lexicon-based* mencapai hasil akurasi tertinggi. Elmitwalli, Sherif dan Mehegan, John [12] juga melakukan analisis sentimen menggunakan dataset IMDB *reviews* dan Sentiment140 dengan algoritma Vader, BiLSTM, BERT, dan GPT-3 dengan hasil menunjukkan GPT-3 unggul. Bello, dkk [13] juga melakukan analisis sentimen dan bereksperimen dengan menggabungkan masing-masing algoritma CNN, LSTM, dan BiLSTM dengan BERT. Hasil eksperimen menunjukkan bahwa kombinasi algoritma-algoritma tersebut dengan BERT berhasil meningkatkan nilai akurasi, presisi, *recall*, dan *F-1 score*.

Selain Bello, dkk [13], Talaat, Amira [14] juga melakukan analisis sentimen dengan menggabungkan BERT dengan *Bidirectional Long Short Term Memory* (BiLSTM) dan *Bidirectional Gated Recurrent Unit* (BiGRU). Hasil akurasi terbaik yang didapatkan adalah 89,52%. Kemudian, Oswari, dkk [7] juga melakukan analisis sentimen terhadap YouTube *Reviews* menggunakan IndoBERT *Fine-Tuning*. IndoBert adalah bagian dari model BERT yang sudah dispesifikasikan untuk bahasa Indonesia. Berdasarkan hasil yang diberikan, model yang dikembangkan dapat menggunakan IndoBERT. Adapun untuk akurasi yang didapat adalah 74%. Ketujuh penelitian di atas melakukan analisis sentimen untuk mengklasifikasi teks. Akan tetapi, penelitian-penelitian tersebut belum mengaitkan analisis sentimen untuk melakukan prediksi terhadap suatu masalah.

Tabel 1. Hasil Penelitian Analisis Sentimen untuk Klasifikasi Teks

Author	Dataset	Metodologi	Akurasi (%)
Hidayatullah, dkk. [5]	Postingan X mengenai Pemilu 2019	SVM + TF-IDF	84,04
		Logistic Regression + TF-IDF	83,15
		Multinomial Naïve Bayes + TF-IDF	82,13
		LSTM	84,20
		CNN	84,05
		CNN + LSTM	84,30
		GRU + LSTM	84,50
		Bidirectional LSTM	84,60
Seno, Danar W., dan Arief Wibowo [6]	Postingan X mengenai Pemilu 2019	SVM berbasis Lexicon-based	92,5
Elmitwalli, Sherif dan Mehegan, John [12]	IMDB reviews	VADER	58,31
		Bi-LSTM	84,12
		BERT	86,27
		GPT-3	88,12
	Sentiment 140	VADER	58,31
		Bi-LSTM	75,23

		BERT	81,41
		GPT-3	88,12
Bello, dkk [13]	6 dataset mengenai postingan dari X yang dikumpulkan dari Kaggle dan digabung menggunakan Python	CNN	89
		RNN	90
		Bi-LSTM	90
		BERT-CNN	93
		BERT-RNN	93
		BERT-BiLSTM	93
Talaat, Amira [14]	Tweets about Apple computer	DistilBERT + GLG90	88,04
		RoBERTa + GLG90	90,18
Oswari, dkk [7]	Youtube Reviews about Lesgian, Gay, Bisexual, and Transgender (LGBT)	IndoBERT + Lexicon-based labeling	74

Mengenai analisis sentimen untuk prediksi, pada 2020, Sebastian, W dan Isa, S [8] melakukan analisis sentimen untuk prediksi harga saham dengan membandingkan penggunaan BERT dan Word2Vec dengan menggunakan 3 dataset. Hasil penelitian tersebut membuktikan bahwa akurasi yang dicapai dengan menggunakan BERT pada ketiga dataset jauh lebih bagus melebihi akurasi yang dicapai Word2Vec dengan akurasi tertinggi dari ketiga dataset adalah 98,46%. Kemudian, pada 2021, Kaushal, dkk [9] melakukan analisis sentimen dan memprediksi kepribadian Myers-Briggs pengguna X menggunakan *Machine Learning Classifiers* dan BERT. Dengan membagi kepribadian Myers-Briggs menjadi empat model, yaitu IE (*Introverts/Extroverts*), NS (*Intuitives/Sensors*), PJ (*Perceivers/Judgers*), dan FT (*Feelers/Thinkers*), didapatkan bahwa pada keempat model tersebut, SVM mencapai hasil tertinggi dengan akurasi 84,28%, 88,20%, 83,45%, dan 78,56% untuk masing-masing model. Lalu, pada 2022, Boukabous, M dan Azizi M [10] melakukan sebuah penelitian analisis sentimen untuk memprediksi kejahatan dengan menggunakan pendekatan *hybrid*, yaitu menggabungkan metode BERT dan *lexicon-based* dan mencapai hasil akurasi sebesar 94,91%.

Tabel 2. Hasil Penelitian Analisis Sentimen untuk Prediksi

Author	Keterangan Dataset			Metodologi	Akurasi (%)
Sebastian, W dan Isa, S [8]	Analisis Sentimen untuk Prediksi Harga Saham	Saham Apple		BERT	95,20
				Word2Vec	73,06
		Saham Micosoft		BERT	98,46
				Word2Vec	76,54
		Saham Nike		Bert	95,60
				Word2Vec	73,63
Kaushal, dkk [9]	Prediksi Tipe Kepribadian Myers-Briggs dan Analisis Sentimen Pengguna X	Postingan X dan tipe MBTI	IE (Introverts,Extroverts)	KNN	77,18
				Logistic Regression	82,25
				Multinomial Naïve Bayes	79,98
				Random Forest	77,09
				Stochastic Gradient Descent	81,46
				SVM	84,28
			NS (Intuitives, Sensors)	KNN	84,37
				Logistic Regression	88,01
				Multinomial Naïve Bayes	85,52
				Random Forest	86,12
				Stochastic Gradient Descent	87,41

				SVM	88,20
			FT (Feelers, Thinkers)	KNN	63,90
				Logistic Regression	79,35
				Multinomial Naïve Bayes	81,05
				Random Forest	76,35
				Stochastic Gradient Descent	78,19
				SVM	83,45
			PJ (Perceivers, Judgers)	KNN	63,85
				Logistic Regression	73,67
				Multinomial Naïve Bayes	73,40
				Random Forest	65,47
				Stochastic Gradient Descent	72,57
				SVM	78,56
Boukabous, M dan Azizi M [10]	Prediksi Kejahatan Menggunakan Pendekatan Analisis Sentimen	Dataset dari X hasil crawling		Hybrid (lexicon-based + BERT)	94,91

Berdasarkan penelitian-penelitian sebelumnya, dapat disimpulkan bahwa aplikasi BERT kepada BiLSTM dapat meningkatkan akurasi algoritma tersebut sehingga kali ini, peneliti ingin mencoba metode BERT-BiLSTM dalam melakukan analisis sentimen untuk prediksi menggunakan dataset dari X.

3. Sistem yang Dibangun

Alur kerja pada penelitian ini diawali dengan pengumpulan dataset berdasarkan aturan-aturan yang telah dibuat. Dataset tersebut kemudian masuk pada tahap *preprocessing*. Hasil dari tahapan ini digunakan untuk pembuatan model algoritma BiLSTM setelah diubah ke dalam representasi vektor numerik menggunakan metode BERT.

3.1. Dataset

Dataset yang digunakan dalam penelitian ini menggunakan dataset yang diambil dari *Twitter* atau X menggunakan teknik *Crawling*. X merupakan salah satu aplikasi media sosial yang paling bayak digunakan. Data yang diambil dari dataset ini adalah berupa *tweet* yang berisi opini atau pendapat publik mengenai kandidat pasangan Calon Presiden dan Wakil Presiden Indonesia dengan menggunakan bahasa Indonesia. Dataset yang digunakan sebanyak 6000 data yang terdiri dari 2000 data untuk masing-masing pasangan calon kandidat dengan atribut data, label, dan tag. Penulis juga menggunakan 18000 data tambahan yang terdiri dari 6000 data yang belum dilabeli untuk masing-masing pasangan calon kandidat untuk prediksi sentimen.

3.2. Labelling

Dataset yang telah dikumpulkan telah diberi label secara manual oleh peneliti. Labelling pada data dibagi menjadi dua skenario, yaitu : dataset menggunakan label ‘sarkas’ dan dataset tidak menggunakan label ‘sarkas’.

3.3. Preprocessing

Preprocessing merupakan tahap yang dilakukan sebelum masuk ke tahap *processing data* dengan algoritma BERT-BiLSTM. Data-data yang telah dikumpulkan akan diolah untuk memastikan tidak ada masalah dalam data yang akan digunakan dan menjadikan data dalam bentuk yang dapat diolah algoritma. Berikut beberapa tahap *preprocessing data* dalam penelitian ini :

Tahap pertama adalah *data cleansing* atau pembersihan data. Dalam dataset dengan jumlah besar, besar kemungkinan adanya data duplikat dan banyak simbol atau karakter yang tidak diperlukan seperti *mentions* (contoh: @AniesBaswedan, @Prabowo, @Gibran, dll), URL (contoh: https://X.com, www.google.com, www.facebook.com, dll) dan karakter-karakter spesial lainnya (contoh: @,#,$,%, !, ^, &, *, dst). *Data cleansing* dilakukan untuk membersihkan data dari berbagai elemen yang dapat

mengurangi akurasi algoritma BERT-BiLSTM, menghindari overfitting, dan meningkatkan efisiensi dengan menghapus elemen-elemen yang tidak diperlukan saat melakukan analisis sentimen.

Tahap kedua adalah *case folding*. *Case folding* digunakan untuk mengubah huruf kecil menjadi besar atau huruf besar menjadi kecil (*uppercase/lowercase*) dengan tujuan menyamakan bentuk kata agar tidak dianggap berbeda hanya karena perbedaan huruf besar dan huruf kecil.

Tahap ketiga adalah *Tokenization.* *Tokenization* merupakan proses memecahkan kalimat menjadi beberapa kata yang dipisahkan menggunakan karakter agar bisa diproses oleh model algoritma.

Tahap keempat adalah *stopwords removal*. *Stopwords* merupakan kata yang sering muncul dalam teks tetapi tidak mengandung informasi atau makna yang berguna untuk analisis seperti yang, di, ke, untuk, pada, adalah, ini, itu, dan lain-lain. *Stopwords* umumnya merupakan kata-kata yang sangat umum yang tidak berhubungan langsung dengan informasi penting dalam teks seperti kata penghubung.

Tabel 3. Daftar Kata *Stopwords* dalam Kamus Sastrawi

Daftar Kata *Stopwords* dalam Kamus Sastrawi
'saya', 'bagi', 'seterusnya', 'walau', 'sedangkan', 'yaitu', 'para', 'selagi', 'ketika', 'juga', 'kenapa', 'ok', 'bahwa', 'dll', 'tanpa', 'dengan', 'kami', 'itu', 'selain', 'dst', 'agak', 'ia', 'mereka', 'kita', 'kemana', 'anu', 'setelah', 'lain', 'masih', 'ini', 'begitu', 'agar', 'dapat', 'apakah', 'demikian', 'seharusnya', 'boleh', 'pasti', 'ya', 'saat', 'hanya', 'saja', 'mengapa', 'seraya', 'tentang', 'sebab', 'namun', 'anda', 'terhadap', 'sampai', 'bisa', 'kecuali', 'sebelum', 'kepada', 'itulah', 'dia', 'nanti', 'tidak', 'harus', 'atau', 'untuk', 'amat', 'oh', 'tapi', 'sesudah', 'adalah', 'bagaimanapun', 'daripada', 'melainkan', 'oleh', 'sekitar', 'antara', 'karena', 'seolah', 'mari', 'dan', 'serta', 'yakni', 'ada', 'guna', 'ke', 'dsb', 'hal', 'supaya', 'pula', 'apalagi', 'seperti', 'telah', 'sebagai', 'setiap', 'sudah', 'kembali', 'dahulu', 'pun', 'nggak', 'pada', 'toh', 'yang', 'kah', 'ingin', 'di', 'sebetulnya', 'lagi', 'dalam', 'demi', 'jika', 'dari', 'secara', 'akan', 'menurut', 'setidaknya', 'dua', 'dimana', 'tolong', 'sesuatu', 'sambil', 'sementara', 'maka', 'dulunya', 'tetapi', 'belum', 'sehingga', 'tentu'

Tabel 4. Daftar Kata *Stopwords* yang ditambahkan ke dalam Kamus

Daftar Kata *Stopwords* Tambahan
'yg', 'dg', 'rt', 'dgn', 'ny', 'klo', 'klw', 'amp', 'biar', 'bikin', 'bilang', 'krn', 'nya', 'nih', 'sih', 'si', 'tau', 'tdk', 'tuh', 'utk', 'ya', 'jd', 'jgn', 'sdh', 'aja', 'n', 't', 'hehe', 'pen', 'u', 'nan', 'loh', 'rt', '&', 'https', 'http', 'wk', 'wkwk', 'wkwkwk', 'wkwkwkwk', 'haha', 'hehe', 'eh', 'hmm', 'om', 'deh', 'ah', 'pada', 'kita', 'bisa', 'dong', 'lah', 'sama', 'orang', 'mereka', 'katanya', 'coy', 'bos', 'gw', 'lu', 'elo', 'gua', 'gue', 'stlh', 'sth', 'elu', 'tp', 'tuh', 'tuhnya', 'sdgkan', 'yi', 'ktk', 'jg', 'knp', 'mrk', 'kmn', 'msh', 'dpt', 'blh', 'sehrsnya', 'sj', 'ttg', 'thd', 'bs', 'kec', 'sblm', 'kpd', 'unt', 'hrs', 'adl', 'bgmnpun', 'spt', 'tlh', 'sbg', 'sdh', 'pd', 'lg', 'dlm', 'jk', 'scr', 'mnrt', 'dmn', 'tlg', 'ttpi', 'blm', 'shg', 'kyk', 'kayak', 'bahwa', 'dr', 'wong', 'nich', 'nih', 'hadiuuh', 'hadeuh', 'p', 'd'

Tahap kelima adalah *Stemming*. *Stemming* merupakan proses untuk menghilangkan imbuhan-imbuhan yang ada dalam sebuah kata untuk diubah menjadi bentuk dasarnya. Tujuan dari proses ini adalah meningkatkan efisiensi dan akurasi model dengan menyederhanakan kata.

Tabel 5. Contoh *Stemming*

Sebelum	**Sesudah**
Merasakan	Rasa
Menyukai	Suka
Berbicara	Bicara

Memamerkan Melakukan	Pamer Lakukan

Tahap ketujuh adalah membagi dataset menjadi dua. Dataset yang telah selesai diolah akan dibagi menjadi dua bagian: Data latih (*data train*) dan data uji (*data test*). Data latih digunakan untuk melatih model algoritma BERT-BiLSTM sedangkan data uji digunakan untuk menguji model yang telah dilatih dengan data latih dan dinilai akurasinya. Pembagian dataset menggunakan metode Stratified 10-K Fold Split.

3.4. Word Embedding menggunakan BERT

Setelah melakukan tahap *preprocessing*, dataset diubah ke dalam bentuk representasi vektor numerik yang digunakan untuk menangkap makna dan hubungan semantic antar kata. Tujuan utama dari *word embedding* adalah untuk mengubah kata-kata yang telah diproses dalam dataset yang berbentuk *string* (teks) menjadi bentuk yang dapat diproses oleh model algoritma BiLSTM.

Jenis model dari algoritma BERT sendiri sudah banyak. Dalam penelitian kali ini, dikarenakan dataset yang digunakan adalah komentar dalam bahasa Indonesia, peneliti menggunakan model IndoBERT untuk hasil yang lebih optimal.

3.5. Klasifikasi Data Sentimen Menggunakan BiLSTM

Setelah data diproses menjadi bentuk yang dapat diolah oleh model, peneliti menggunakan dataset yang telah diubah ke dalam bentuk representasi vektor untuk melatih dan menguji model algoritma BiLSTM. Data tersebut dibagi menggunakan metode *Stratified* 10-*fold Cross-Validation* dimana dataset dibagi menjadi sepuluh subset. Sembilan subset digunakan untuk melatih model dan satu subset yang tersisa digunakan untuk menguji model dengan tujuan mengevaluasinya. Proses ini dilakukan sepuluh kali untuk melatih model dan meningkatkan akurasinya.

3.6. Prediksi Data Sentimen Menggunakan BiLSTM

Setelah model algoritma BiLSTM dilatih dan diuji menggunakan dataset yang sudah dilabeli, model tersebut kemudian digunakan untuk melakukan prediksi sentimen menggunakan dataset yang tidak dilabeli. Dataset tersebut juga melalui tahap *preprocessing* yang sama seperti dataset sebelumnya.

3.7. Metrik Evaluasi

Pada penelitian ini, metrik evaluasi yang digunakan adalah *10-fold cross-validation, accuracy, precision, recall,* dan *F-1 Score. 10-fold cross-validation* merupakan metode dimana dataset dibagi menjadi 10 subset dan proses *training* dan *testing* akan dilakukan 10 kali dimana satu subset dari dataset digunakan untuk data uji dan sisa subset digunakan sebagai data latih. *Accuracy* dalam analisis sentimen untuk prediksi menyatakan sejauh mana model atau sistem dapat mengklasifikasikan analisis sentimen atau memprediksi pemenang pemilu sesuai dengan dataset dari X seperti yang terlihat pada persamaan berikut:

$$Q = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision atau presisi merupakan metrik yang digunakan untuk mengukur seberapa banyak dari tweets yang benar-benar postif, negatif, atau netral seperti yang terlihat pada persamaan berikut:

$$Precision = \frac{TP}{TP + FP}$$

Recall merupakan metrik yang digunakan untuk mengukur seberapa banyak dari tweet yang berhasil diidentifikasi sebagai positif, negatif, atau netral oleh sistem seperti yang terlihat pada persamaan berikut:

$$Recall = \frac{TP}{TP + FN}$$

Sedangkan *F-1 Score* merupakan metrik kombinasi dari *precision* dan *recall* untuk mengukur performa atau memberikan gambaran performa secara keseluruhan seperti yang terlihat pada persamaan berikut:

$$F - 1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. Evaluasi

4.1. Hasil Pengujian

Pada penelitian kali ini, analisis sentimen dilakukan berdasarkan dua skenario, yaitu ketika menggunakan label sarkas dan ketika tidak menggunakan label sarkas dalam dataset. Label sarkas dalam dataset ini dikategorikan sebagai kalimat negatif yang bisa disalah artikan sebagai kalimat positif. Berdasarkan dua skenario ini, hasil pengujian yang diperoleh algoritma (%) adalah sebagai berikut :

Tabel 6. Hasil Pengujian Menggunakan Dataset dengan Label 'Sarkas'

Fold ke- / Pasangan	1	2	3	4	5	6	7	8	9	10	Rata-rata
Paslon 1	67,3	85,4	97,5	99	99	97,5	100	99,5	98	96	93,9
Paslon 2	63,9	79	96,1	97,9	100	100	100	89,2	98,3	100	92,4
Paslon 3	78,6	87,9	96,7	98,6	99,5	100	97,7	97,7	99,5	100	95,6

Tabel 7. Hasil Pengujian Menggunakan Dataset yang Tidak Menggunakan Label 'Sarkas'

Fold ke- / Pasangan	1	2	3	4	5	6	7	8	9	10	Rata-rata
Paslon 1	62,3	80,4	96,5	98,5	99,5	100	100	99	100	100	93,6
Paslon 2	63,5	79,8	92,3	99,6	99,1	99,6	98,3	99,6	99,6	100	93,1
Paslon 3	79,1	88,4	96,7	99,1	100	96,3	100	99,5	99,1	99,5	95,8

Berdasarkan model-model yang telah dilatih dan diuji sebelumnya, model-model tersebut digunakan untuk melakukan analisis sentimen untuk prediksi menggunakan dataset komentar baru yang belum dilabeli. Hasil tertinggi diperoleh Paslon 1 (Anies-Muhaimin) dengan nilai sentimen positif sebesar 71,76%, sentimen netral sebesar 8,99%, dan sentimen negatif/sarkas sebesar 19,25%

Tabel 8. Hasil Prediksi Analisis Sentimen Ketiga Paslon

Sentimen / Pasangan	Positif (%)	Netral (%)	Negatif/Sarkas (%)
Paslon 1	71,76	8,99	19,25
Paslon 2	28,82	25,63	45,54
Paslon 3	65,53	20,32	14,16

4.2. Analisis Hasil Pengujian

Hasil pengujian tertinggi yang didapat dari penelitian kali ini adalah nilai akurasi sebesar 95,8% sedangkan nilai tertinggi yang dicapai oleh Sebastian, W dan Isa, S [8] yang menggunakan BERT untuk analisis sentimen untuk prediksi harga saham adalah 98,46%.

Terdapat beberapa hal yang dapat memengaruhi akurasi algoritm BERT-BiLSTM. Diantaranya :

a. Jumlah dataset. Dataset yang digunakan untuk melatih dan menguji model algoritma berjumlah 6000 data dimana 2000 data komentar dialokasikan untuk masing-masing paslon. Dataset yang lebih banyak mungkin akan memengaruhi hasil akurasi karena bahan latih yang lebih banyak.
b. Pelabelan dataset. Pelabelan data dilakukan secara manual. Bisa jadi terdapat label kalimat yang dilabeli positif/negatif/sarkas/netral oleh peneliti keliru dikarenakan pelabelan dilakukan sendiri secara manual.
c. Penggunaan bahasa *slang*, singkatan-singkatan, dan bahasa daerah yang digunakan oleh pengguna bisa jadi tidak terdeteksi dikarenakan bukan bahasa Indonesia baku.
d. Penggunaan kalimat sarkas yang dapat disalahartikan menjadi kalimat positif.

e. Jumlah unit neuron pada BiLSTM dapat memengaruhi kapasitas model untuk belajar. Tapi jika terlalu tinggi, dapat meningkatkan risiko overfitting jika data tidak cukup banyak. Jika terlalu rendah, kemampuan model dalam menangkap pola yang kompleks akan berkurang.

5. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil kesimpulan sebagai berikut :

a. Untuk melakukan analisis sentimen untuk prediksi, peneliti melakukan *preprocessing* dataset. Setelah itu, data diubah ke dalam representasi vektor numerik menggunakan model algoritma IndoBERT. Evaluasi dan pelatihan model algoritma BiLSTM dilakukan menggunakan metode *Stratified* 10-*fold Cross Validation* dan menghitung akurasi dan hasil evaluasi performansi tiap *fold*-nya.

b. Berdasarkan hasil pengujian yang dilakukan tingkat akurasi tertinggi algoritma BERT-BiLSTM dalam melakukan analisis sentimen untuk prediksi pemenang pemilu 2024 adalah sebesar 95,8% dengan nilai 95,8% dengan nilai *Precision* 93,2%, *Recall* 92,4% dan *F-1* 92,2%.

c. Menurut hasil analisis sentimen untuk prediksi, pemenang dari pemilu 2024 adalah pasangan Calon Presiden dan Calon Wakil Presiden nomor urut 1, yaitu Anies Baswedan dan Muhaimin Iskandar dengan sentimen positif sebesar 71,76%, sentimen netral sebesar 25,63% dan sentimen negatif sebesar 19,25%.

Adapun saran yang dapat diberikan pada penelitian ini adalah sebagai berikut :

a. Menambah jumlah dataset yang digunakan untuk melatih dan menguji model.
b. Pelabelan dilakukan secara berkelompok untuk menghindari kekeliruan sentimen atau mencoba menggunakan metode auto labelling.
c. Mengubah bahasa-bahasa *slang* atau bahasa daerah ke dalam Bahasa Indonesia baku agar dapat dikenali model.
d. Mencoba jumlah unit neuron pada BiLSTM yang lebih tinggi untuk meningkatkan kapasitas model untuk belajar terlebih jika menggunakan dataset dengan jumlah yang lebih besar.

Daftar Pustaka

[1] D. Nuruddin, H. A. Muhasim, M. Hi, and C. A. Press, *HUKUM TATA NEGARA INDONESIA*. [Online]. Available: www.cvalfapress.my.id

[2] A. E. Subiyanto, “Pemilihan Umum Serentak yang Berintegritas sebagai Pembaruan Demokrasi Indonesia,” *Jurnal Konstitusi*, vol. 17, no. 2, p. 355, Aug. 2020, doi: 10.31078/jk1726.

[3] F. C. ’Farisa, “Hasil Lengkap Pemilu 2024: Pilpres dan Pileg,” https://nasional.kompas.com/read/2024/03/21/11334381/hasil-lengkap-pemilu-2024-pilpres-dan-pileg#google_vignette. Accessed: May 04, 2024. [Online]. Available: https://nasional.kompas.com/read/2024/03/21/11334381/hasil-lengkap-pemilu-2024-pilpres-dan-pileg#google_vignette

[4] B. Liu, “Sentiment Analysis and Opinion Mining,” Morgan & Claypool Publishers, 2012.

[5] A. F. Hidayatullah, S. Cahyaningtyas, and A. M. Hakim, “Sentiment Analysis on Twitter using Neural Network: Indonesian Presidential Election 2019 Dataset,” *IOP Conf Ser Mater Sci Eng*, vol. 1077, no. 1, p. 012001, Feb. 2021, doi: 10.1088/1757-899x/1077/1/012001.

[6] D. W. Seno and A. Wibowo, “Analisis Sentimen Data Twitter Tentang Pasangan Capres-Cawapres Pemilu 2019 Dengan Metode Lexicon Based Dan Support Vector Machine,” *Jurnal Ilmiah FIFO*, vol. 11, no. 2, p. 144, Nov. 2019, doi: 10.22441/fifo.2019.v11i2.004.

[7] T. Oswari, T. Yusnitasari, and S. Wijay, “Sentiment Analysis of Indonesian YouTube Reviews About Lesbian, Gay, Bisexual, and Transgender (LGBT) using IndoBERT Fine Tuning,” vol. 15, no. 1, pp. 2088–1541, 2024, doi: 10.24843/LKJITI.2024.v15.i01.p03.

[8] W. Sebastian, “Stock Price Prediction Using BERT and Word2Vec Sentiment Analysis,” *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 9, pp. 5430–5433, Sep. 2020, doi: 10.30534/ijeter/2020/85892020.

[9] P. Kaushal, N. B. B P, P. M S, K. S, and A. K. Koundinya, “Myers-briggs Personality Prediction and Sentiment Analysis of Twitter using Machine Learning Classifiers and BERT,” *International Journal of Information Technology and Computer Science*, vol. 13, no. 6, pp. 48–60, Dec. 2021, doi: 10.5815/ijitcs.2021.06.04.

[10] M. Boukabous and M. Azizi, “Crime prediction using a hybrid sentiment analysis approach based on the bidirectional encoder representations from transformers,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 2, pp. 1131–1139, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp1131-1139.

[11] S. Mann, J. Arora, M. Bhatia, R. Sharma, and R. Taragi, “Twitter Sentiment Analysis Using Enhanced BERT,” in *Lecture Notes in Electrical Engineering*, Springer Science and Business Media Deutschland GmbH, 2023, pp. 263–271. doi: 10.1007/978-981-19-6581-4_21.

[12] S. Elmitwalli and J. Mehegan, “Sentiment analysis of COP9-related tweets: a comparative study of pre-trained models and traditional techniques,” *Front Big Data*, vol. 7, 2024, doi: 10.3389/fdata.2024.1357926.

[13] A. Bello, S. C. Ng, and M. F. Leung, “A BERT Framework to Sentiment Analysis of Tweets,” *Sensors*, vol. 23, no. 1, Jan. 2023, doi: 10.3390/s23010506.

[14] A. S. Talaat, “Sentiment analysis classification system using hybrid BERT models,” *J Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00781-w.