

Perbandingan Kinerja *Machine Learning* dan *Deep Learning* untuk Analisis Sentimen Fufufafa

Darusman^{1,*}, Windu Gata¹

¹ Program Studi Magister Ilmu Komputer; Universitas Nusa Mandiri Jakarta; Jl.Raya Jatiwaringin Cipinang Melayu, Kec.Makasar, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, No.Telpn (021)28534471; e-mail: 14230029@nusamandiri.ac.id; windu@nusamandiri.ac.id;

* Korespondensi: e-mail: 14230029@nusamandiri.ac.id

Diterima: 14 Januari 2025; Review: 03 Mei 2025; Disetujui: 16 Juni 2025

Cara sitasi: Darusman D, Gata W. 2025. Perbandingan Kinerja *Machine Learning* dan *Deep Learning* untuk Analisis Sentimen Fufufafa. Information System for Educators and Professionals. Vol 10(1): 1-12.

Abstrak: Analisis sentimen pada data media sosial, khususnya Twitter, menghadapi tantangan dalam mengklasifikasikan opini yang dinamis dan kompleks ke dalam kategori positif, negatif, dan netral. Penelitian ini membandingkan kinerja beberapa model machine learning dan deep learning dalam menganalisis sentimen terhadap fenomena fufufafa di Twitter. Model yang diuji meliputi Naive Bayes, Random Forest, Support Vector Machine, Logistic Regression, Decision Tree, dan Long Short-Term Memory (LSTM). Tujuan utama penelitian ini adalah mengevaluasi efektivitas masing-masing model dalam mengklasifikasikan sentimen secara akurat. Hasil eksperimen menunjukkan bahwa Decision Tree mencapai akurasi tertinggi sebesar 96%, diikuti oleh LSTM dengan 95%. Model Random Forest, SVM, dan Logistic Regression mencatat akurasi sebesar 94%, sedangkan Naive Bayes memiliki akurasi terendah yaitu 81%, terutama karena keterbatasannya dalam menangani sentimen yang kompleks. Model LSTM unggul dalam menangkap konteks temporal dan hubungan antar kata sehingga memberikan prediksi lebih akurat meski membutuhkan sumber daya komputasi lebih besar. Temuan ini menegaskan bahwa deep learning, khususnya LSTM, lebih efektif dalam menganalisis sentimen pada data sosial media yang dinamis dibandingkan metode machine learning konvensional. Penelitian ini memberikan dasar untuk pengembangan metode analisis sentimen yang lebih efisien dan akurat di masa depan.

Kata kunci: Analisis Sentimen, Media Sosial Twitter, *Machine Learning*, *Deep Learning*

Abstract: Sentiment analysis on social media data, particularly Twitter, faces challenges in classifying dynamic and complex opinions into positive, negative, and neutral categories. This study compares the performance of several machine learning and deep learning models in analyzing sentiments related to the fufufafa phenomenon on Twitter. The models tested include Naive Bayes, Random Forest, Support Vector Machine, Logistic Regression, Decision Tree, and Long Short-Term Memory (LSTM). The primary objective is to evaluate the effectiveness of each model in accurately classifying sentiments. Experimental results show that the Decision Tree achieved the highest accuracy of 96%, followed by LSTM with 95%. Random Forest, SVM, and Logistic Regression recorded accuracies of 94%, while Naive Bayes had the lowest accuracy at 81%, mainly due to its limitations in handling complex sentiments. The LSTM model excels in capturing temporal context and word relationships, providing more accurate predictions despite requiring greater computational resources. These findings confirm that deep learning, particularly LSTM, is more effective in analyzing sentiments on dynamic social media data compared to conventional machine learning methods. This study provides a foundation for developing more efficient and accurate sentiment analysis methodologies in the future.

Keywords: Sentiment Analysis, Twitter Social Media, *Machine Learning*, *Deep Learning*

1. Pendahuluan

Dalam era digital saat ini, media sosial telah menjadi platform penting bagi individu untuk berbagi informasi, opini, dan perasaan terhadap berbagai isu yang sedang berkembang. Twitter, sebagai salah satu platform media sosial terbesar di dunia, menjadi ruang interaksi yang penuh dinamika bagi penggunanya [1][2][3]. Melalui fitur *tweet*, pengguna dapat dengan cepat menyampaikan pandangan atau reaksi terhadap suatu fenomena, baik itu terkait politik, sosial, ekonomi, maupun hiburan [4][5]. Setiap *tweet* berpotensi mencerminkan sentimen positif, negatif, atau netral dari pengguna terhadap topik tertentu [6][7]. Oleh karena itu, analisis sentimen terhadap data yang ada di media sosial, khususnya *twitter*, menjadi sangat penting untuk memahami pandangan dan perasaan publik terhadap berbagai isu yang sedang populer.

Fenomena yang tengah ramai diperbincangkan di *twitter*, seperti fufufafa, menjadi salah satu topik yang menarik untuk dianalisis. Istilah ini, meskipun mungkin tidak familiar di kalangan sebagian orang, ternyata mampu menciptakan respons yang beragam di kalangan pengguna *twitter* [8][9]. Dengan semakin maraknya pembicaraan tentang fufufafa, menjadi relevan untuk melakukan analisis sentimen guna mengetahui bagaimana persepsi masyarakat terhadap fenomena tersebut, apakah cenderung positif, negatif, atau bahkan netral [10][11]. Melalui analisis sentimen ini, penelitian bertujuan menggali lebih dalam pengaruh dan dampak fenomena tersebut di dunia maya, sekaligus mengetahui bagaimana persepsi dan reaksi pengguna Twitter terhadap fenomena yang relatif baru dan unik ini.

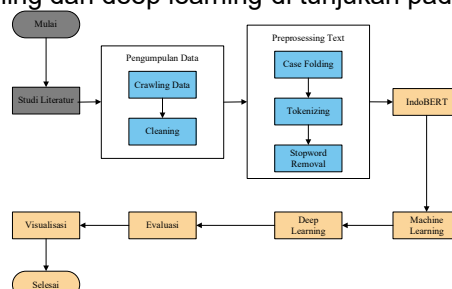
Selain itu, model berbasis *transformer* seperti *BERT (Bidirectional Encoder Representations from Transformers)* dan adaptasinya untuk bahasa Indonesia, yaitu *IndoBERT*, telah memberikan terobosan dalam analisis sentiment [12][13][14] [15]. *IndoBERT* dirancang khusus untuk menangani nuansa bahasa Indonesia, menjadikannya sangat relevan untuk analisis sentimen dalam konteks lokal. Penelitian menunjukkan bahwa *IndoBERT* memiliki kemampuan untuk memahami konteks bahasa yang kompleks, sehingga menghasilkan akurasi yang lebih tinggi dibandingkan pendekatan lainnya [16][17] [18].

Pendekatan berbasis machine learning dan deep learning telah berkembang menjadi metode utama dalam analisis sentimen berbasis *teks. transformer*, seperti *indobert*, dirancang khusus untuk menangkap nuansa kompleks bahasa indonesia, menawarkan kemampuan analisis yang unggul [19][20][21]. Berbagai penelitian menunjukkan bahwa model berbasis *transformer* ini mampu menghasilkan akurasi yang lebih tinggi dalam menganalisis konteks semantik dibandingkan metode konvensional. Dengan demikian, penerapan model seperti *indobert* pada analisis sentimen terhadap fenomena fufufafa di *twitter* dapat memberikan wawasan yang lebih mendalam tentang opini publik sekaligus mendorong pengembangan metodologi yang lebih efektif dan akurat [22][23][24].

Dalam penelitian ini, pendekatan yang digunakan melibatkan perbandingan metode machine learning dan deep learning untuk menganalisis sentimen terhadap fenomena fufufafa di *twitter*. Pendekatan ini bertujuan tidak hanya untuk memberikan wawasan mengenai persepsi publik terhadap fenomena tersebut, tetapi juga untuk mengeksplorasi keunggulan relatif antara metode klasik dan modern dalam analisis sentimen berbasis *teks*. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan teknik analisis sentimen yang lebih efisien dan akurat di masa depan.

2. Metode Penelitian

Metode penelitian yang akan di buat tentang analisis sentimen fufufafa di *twitter* perbandingan machine learning dan deep learning di tunjukan pada Gambar 1.



Sumber : Hasil Penelitian (2025)

Gambar 1 Metode Penelitian

Studi Literatur

Pada tahap awal penelitian ini, dilakukan kajian literatur terhadap jurnal-jurnal yang relevan mengenai analisis sentiment. Berdasarkan kajian ini, penelitian ini mengadaptasi metodologi yang sesuai, termasuk penerapan model *transformer* seperti *IndoBERT*, untuk menganalisis sentimen terhadap fenomena fufufafa yang tengah populer di *twitter*, serta membandingkannya dengan metode machine learning dan deep learning untuk memperoleh wawasan yang lebih mendalam.

Pengumpulan Data

Data penelitian ini diperoleh menggunakan pustaka Python *scraping_twitter* dengan bantuan Twitter API berdasarkan kata kunci *fufufafa* [25]. Pengumpulan data dilakukan dari 15 Juni hingga 15 November 2024, mencakup teks tweet, waktu publikasi, jumlah retweet, dan like [26]. Data yang diperoleh kemudian disimpan dalam format Excel untuk memudahkan proses pra-pemrosesan, yang meliputi penghapusan duplikasi, mention, hashtag, URL, retweet prefix, dan karakter non-alfanumerik [27]. Proses ini bertujuan memastikan bahwa data yang digunakan bersih dan relevan untuk analisis sentimen.

Preprocessing Text

Pada tahap *pra-pemrosesan data*, peneliti melakukan beberapa langkah penting, termasuk *case folding* untuk mengubah semua huruf menjadi kecil guna memastikan konsistensi *format teks* [28], *tokenizing* untuk memecah *teks* menjadi unit kata (*token*) yang memudahkan analisis, serta [29], *stopword removal* untuk menghapus kata-kata tidak relevan seperti kata hubung atau ekspresi umum. Langkah-langkah ini memastikan bahwa data yang digunakan fokus pada informasi yang relevan untuk analisis sentimen[30].

IndoBERT

Penelitian ini memanfaatkan model *transformer IndoBERT* sebagai pendekatan utama untuk analisis sentimen. *IndoBERT* digunakan untuk menangkap representasi semantik *teks* di *twitter* dan mengklasifikasikannya ke dalam kategori positif, negatif, atau netral [31]. Prosesnya meliputi *pra-pemrosesan data*, pembobotan *teks* dengan *tokenizer IndoBERT*, dan pelatihan model. Hasil analisis *IndoBERT* dibandingkan dengan metode machine learning dan deep learning lainnya untuk mengevaluasi keunggulannya dalam menganalisis fenomena fufufafa di media sosial.

Machine Learning

Penelitian ini juga menerapkan beberapa algoritma machine learning, termasuk *Naive Bayes*, *Random Forest*, *Support Vector Machine (SVM)*, *Logistic Regression*, dan *Decision Tree*. Setiap algoritma diuji pada data sentimen yang telah diproses sebelumnya [32]. Evaluasi kinerja dilakukan menggunakan *metrik akurasi*, *precision*, *recall*, dan *F1-score* untuk menentukan efektivitas masing-masing metode. Hasilnya dibandingkan dengan model berbasis *transformer* dan deep learning untuk memberikan wawasan yang lebih komprehensif dalam analisis sentimen fenomena fufufafa di *twitter*.

Deep Learning

Penelitian ini menggunakan algoritma *Long Short-Term Memory (LSTM)* untuk analisis sentimen pada data twitter terkait fenomena fufufafa [33]. Model LSTM dirancang dengan lapisan *embedding* untuk representasi vektor kata, lapisan LSTM untuk menangkap hubungan temporal, dan lapisan *dense* untuk klasifikasi sentimen. Model dilatih menggunakan fungsi *loss binary crossentropy* dengan *optimizer Adam*, dan kinerjanya dievaluasi menggunakan metrik *akurasi*, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa LSTM efektif dalam menangkap *konteks teks* dan memberikan hasil analisis sentimen yang akurat.

Evaluasi

Evaluasi penelitian ini menggunakan *confusion matrix* untuk mengukur performa model dalam klasifikasi data. *Confusion matrix* membantu menilai akurasi klasifikasi dengan membandingkan hasil prediksi model terhadap label yang sebenarnya [34]. Dari *matriks* ini, dihitung *metrik evaluasi* seperti *accuracy*, *precision*, *recall*, dan *F1-score*, yang memberikan

gambaran komprehensif tentang kemampuan model dalam menganalisis sentimen secara akurat dan efektif [35].

Visualisasi

Pada tahap *visualisasi*, *wordcloud* digunakan untuk menampilkan kata-kata yang paling sering muncul dalam cuitan berdasarkan sentimen positif, negatif, dan netral. Wordcloud memvisualisasikan frekuensi kata-kata dengan ukuran *font* yang proporsional, sehingga kata-kata yang sering muncul akan terlihat lebih besar [36]. Visualisasi ini membantu peneliti mengidentifikasi kata kunci utama dalam setiap kategori sentimen, mempermudah pemahaman tentang opini yang dominan dalam analisis sentimen.

3. Hasil dan Pembahasan

Pada penelitian ini, algoritma machine learning seperti Naive Bayes, Random Forest, Support Vector Machine (SVM), Logistic Regression, dan Decision Tree, serta algoritma deep learning Long Short-Term Memory (LSTM) diterapkan untuk analisis sentimen terhadap fenomena fufufafa di Twitter. Proses manual dari algoritma tersebut dilakukan secara konseptual, namun memiliki keterbatasan dalam mengolah data besar dan kompleks. Oleh karena itu, pengolahan dan pelatihan model menggunakan tools Python seperti scikit-learn dan TensorFlow yang memungkinkan optimasi parameter, pengolahan fitur TF-IDF, serta evaluasi model secara efisien dan akurat. Hasil pengujian menunjukkan Decision Tree memperoleh akurasi tertinggi sebesar 96%, diikuti oleh LSTM dengan 95%, sementara Random Forest, SVM, dan Logistic Regression mencapai 94%, dan Naive Bayes memiliki akurasi terendah 81% karena kesulitan menangani kompleksitas sentimen.

Perbandingan hasil manual dan tools menegaskan pentingnya pemanfaatan tools modern untuk analisis sentimen pada dataset besar seperti Twitter, yang mampu mempercepat proses pelatihan dan evaluasi model secara tepat. Model LSTM unggul dalam menangkap konteks temporal dan hubungan antar kata, sehingga menghasilkan prediksi yang lebih akurat meskipun membutuhkan sumber daya komputasi lebih besar. Secara keseluruhan, meskipun model machine learning klasik masih efektif, deep learning, khususnya LSTM, memberikan keunggulan dalam memahami nuansa linguistik kompleks yang sering ditemui di media sosial, sehingga menjadi pilihan yang lebih baik untuk analisis sentimen data dinamis dan tidak terstruktur.

Pengumpulan Data

Pengumpulan data dilakukan melalui dua tahap utama *crawling* dan *cleaning*. Tahap pertama, *crawling*, dilakukan dengan menarik data dari *twitter* menggunakan kata kunci fufufafa melalui pustaka *scraping_twitter* yang terintegrasi dengan *twitter API*. Data dikumpulkan antara 15 Juni hingga 15 November 2024, menghasilkan 1.000 *tweet* yang relevan. Tahap kedua, *cleaning*, bertujuan untuk membersihkan data dari elemen tidak relevan seperti *mention*, *hashtag*, *URL*, karakter *non-alfanumerik*, dan duplikasi. Proses ini memastikan data berkualitas tinggi dan siap digunakan untuk analisis sentimen lebih lanjut.

Preprocessing

Pada tahap *preprocessing data* hasil *crawling* yang tidak terstruktur diolah menjadi bentuk yang lebih terorganisir. Proses ini dilakukan dengan menggunakan pustaka python di *google colab*. Langkah pertama adalah *case folding*, yaitu mengubah semua huruf menjadi kecil untuk memastikan konsistensi. Kemudian, dilakukan *tokenizing* untuk memecah *teks* menjadi kata-kata, dan *stopword removal* untuk menghapus kata-kata umum yang tidak penting. Setelah melalui tahap ini, data menjadi bersih dan siap untuk analisis lebih lanjut.

Case Folding

Proses *case folding* dilakukan untuk mengubah semua kata dalam *teks* menjadi huruf kecil, sehingga memastikan konsistensi dalam analisis tanpa terpengaruh perbedaan kapitalisasi. Setelah tahap *case folding*, data yang telah diproses dapat dilihat pada Gambar 2.

	full_text	username	created_at
0	boneka jokowi prabroro fufufafa satpol pp maka...	sastri200531	Thu Nov 14 23:38:37 +0000 2024
1	atau pemain game league of legend ada chara ny...	kulimain	Thu Nov 14 23:38:32 +0000 2024
2	gak akan milih cagub yang didukung fufufafa	yuaran29544	Thu Nov 14 23:38:06 +0000 2024
3	lagi nanyain ijazah mulyono dan fufufafa	Sudrajat04j	Thu Nov 14 23:37:37 +0000 2024
4	fufufafa sama dengan mulyono pencitraan saja	andry_h	Thu Nov 14 23:37:27 +0000 2024

Sumber : Hasil Penelitian (2025)

Gambar 2 Hasil Case Folding

Pada Gambar 2 hasil *case folding* untuk mengubah semua huruf dalam teks menjadi huruf kecil agar konsisten.

Tokenizing

Tokenizing adalah proses untuk memisahkan setiap kata dalam dokumen agar dapat dianalisis secara terpisah. Setelah proses *tokenizing*, data yang telah diproses disajikan pada Gambar 3.

	full_text
0	[boneka, jokowi, prabroro, fufufafa, satpol, p...
1	[pemain, game, league, of, legend, chara, nyi,...
2	[gak, milih, cagub, didukung, fufufafa]
3	[nanyain, ijazah, mulyono, fufufafa]
4	[fufufafa, sama, mulyono, pencitraan]

Sumber : Hasil Penelitian (2025)

Gambar 3 Hasil Tokenizing

Pada Gambar 3 hasil tokenizing berupa daftar kata dari setiap tweet untuk memecah teks menjadi kata-kata atau token.

Stopword Removal

Stopword *removal* adalah proses untuk menghapus kata-kata yang sering muncul namun tidak memberikan informasi signifikan, seperti kata sambung atau kata umum lainnya yang tidak relevan dengan analisis sentimen. Hasil dari *proses stopwords removal* disajikan pada Gambar 4.

	full_text	username	created_at
0	boneka jokowi prabroro fufufafa satpol pp maka...	sastri200531	Thu Nov 14 23:38:37 +0000 2024
1	pemain game league of legend chara nyi blorong...	kulimain	Thu Nov 14 23:38:32 +0000 2024
2	gak milih cagub didukung fufufafa	yuaran29544	Thu Nov 14 23:38:06 +0000 2024
3	nanyain ijazah mulyono fufufafa	Sudrajat04j	Thu Nov 14 23:37:37 +0000 2024
4	fufufafa sama mulyono pencitraan	andry_h	Thu Nov 14 23:37:27 +0000 2024

Sumber : Hasil Penelitian (2025)

Gambar 4 Hasil Stopword Removal

Gambar 4 hasil setelah stopwords dihilangkan untuk menghapus kata umum yang tidak penting dari teks.

IndoBERT

Hasil analisis menunjukkan bahwa *IndoBERT* sesuai dengan hasil yang diperoleh dalam mengklasifikasikan *tweet* sebagai sentimen positif, negatif, atau netral. Perbandingan dengan model lain, seperti LSTM, menunjukkan bahwa *IndoBERT* lebih unggul dalam hal akurasi dan kemampuan menangkap konteks yang lebih kompleks. Mayoritas *tweet* terkait fufufafa cenderung bersifat positif, mengindikasikan bahwa fenomena ini lebih diterima dengan respon humoris dan dukungan dari pengguna *twitter*.

	full_text	Sentiment	Label
0	boneka jokowi prabroro fufufafa satpol pp maka...	Negative	0
1	main game league of legend chara nyi blorong n...	Neutral	1
2	gak milih cagub dukung fufufafa	Positive	2
3	nanyain ijazah mulyono fufufafa	Negative	0
4	fufufafa sama mulyono citra	Neutral	1

Sumber : Hasil Penelitian (2025)

Gambar 5 Hasil Labeling Menggunakan IndoBERT

Pada Gambar 5 mengklasifikasikan tweet ke sentimen negatif, netral, dan positif.

Machine Learning

Penelitian ini mengevaluasi efektivitas lima algoritma machine learning *Naive Bayes*, *Random Forest*, *Support Vector Machine (SVM)*, *Logistic Regression*, dan *Decision Tree* untuk analisis sentimen terhadap fenomena fufufafa di Twitter. Dataset *tweet* diproses melalui pembersihan teks dan ekstraksi fitur menggunakan *TF-IDF*. Hasil eksperimen menunjukkan bahwa *Decision Tree* memiliki akurasi tertinggi sebesar 96%, diikuti oleh *Random Forest*, *SVM*, *Logistic Regression* dan yang masing-masing mencapai akurasi 94%. *Naive Bayes* menunjukkan akurasi terendah, yaitu 81%. Meskipun *Naive Bayes* cepat dalam pelatihan, hasilnya kurang efektif dibandingkan algoritma lain yang lebih kompleks. Secara keseluruhan, *Decision Tree* terbukti paling unggul dalam menganalisis sentimen, meskipun *Random Forest*, *SVM* dan *Logistic Regression* memberikan hasil yang sangat baik pada dataset twitter yang kompleks.

Tabel 1 Hasil Model Naive Bayes

	Precision	Recall	F1-Score	Support
Negative	0,00	0.00	0.00	2
Neutral	0.95	0.84	0.89	94
Positive	0.43	0.60	0.50	5
Accuracy			0.81	101
Macro Avg	0.46	0.48	0.46	101
Weighted Avg	0.91	0.81	0.86	101

Sumber : Hasil Penelitian (2025)

Tabel 1 menunjukkan kinerja algoritma *Naive Bayes* dalam analisis sentimen data *twitter*. Model ini memiliki performa terbaik pada sentimen netral dengan *F1-score* 0,89, namun untuk mengklasifikasikan sentimen negatif dengan nilai 0,00 pada semua metrik. Secara keseluruhan, *akurasi* model mencapai 0,81, menunjukkan keterbatasan *Naive Bayes* dalam menangani sentimen kompleks, terutama untuk kategori negatif.

Tabel 2 Hasil Model Random Forest

	Precision	Recall	F1-Score	Support
Negative	0.00	0.00	0.00	2
Neutral	0.94	1.00	0.97	94
Positive	1.00	0.20	0.33	5
Accuracy			0.94	101
Macro Avg	0.65	0.40	0.43	101
Weighted Avg	0.92	0.94	0.92	101

Sumber : Hasil Penelitian (2025)

Tabel 2 menunjukkan kinerja algoritma *Random Forest* dalam analisis sentimen data *twitter*. Model ini memiliki performa terbaik pada kategori netral dengan *F1-score* 0,97. Namun, model untuk mengklasifikasikan sentimen negatif dengan *F1-score* 0,00 dan memiliki kinerja terbatas pada kategori positif dengan *F1-score* 0,33. Secara keseluruhan, akurasi model mencapai 0,94, menunjukkan efektivitasnya dalam mengklasifikasikan sentimen netral, namun kurang optimal untuk kategori lainnya.

Tabel 3 Hasil Model Support Vector Machine

	Precision	Recall	F1-Score	Support
Negative	0.00	0.00	0.00	2
Neutral	0.94	1.00	0.97	94
Positive	1.00	0.20	0.33	5
Accuracy			0.94	101
Macro Avg	0.65	0.40	0.43	101
Weighted Avg	0.92	0.94	0.92	101

Sumber : Hasil Penelitian (2025)

Tabel 3 menunjukkan kinerja *Support Vector Machine* (SVM) dalam analisis sentimen data *twitter*. Model ini memiliki performa terbaik pada kategori netral dengan *F1-score* 0,97, namun untuk mengklasifikasikan sentimen negatif dengan *F1-score* 0,00. Untuk kategori positif, *F1-score* tercatat 0,33. Secara keseluruhan, akurasi model mencapai 0,94, menunjukkan efektivitas yang tinggi pada kategori netral tetapi kurang optimal pada kategori lainnya.

Tabel 4 Hasil Model Logistic Regression

	Precision	Recall	F1-Score	Support
Negative	0.00	0.00	0.00	2
Neutral	0.95	0.99	0.97	94
Positive	0.67	0.40	0.50	5
Accuracy			0.94	101
Macro Avg	0.54	0.46	0.49	101
Weighted Avg	0.92	0.94	0.93	101

Sumber : Hasil Penelitian (2025)

Tabel 4 menunjukkan kinerja *Logistic Regression* dalam analisis sentimen data *twitter*. Model ini memiliki performa terbaik pada kategori netral dengan *F1-score* 0,97, tetapi untuk mengklasifikasikan sentimen negatif dengan *F1-score* 0,00. Untuk kategori positif, *F1-score* mencapai 0,50. Secara keseluruhan, akurasi model adalah 0,94, menunjukkan efektivitas tinggi pada sentimen netral namun keterbatasan pada kategori negatif dan positif.

Tabel 5 Hasil Model Decision Tree

	Precision	Recall	F1-Score	Support
Negative	0.00	0.00	0.00	2
Neutral	0.96	1.00	0.98	94
Positive	1.00	0.60	0.75	5
Accuracy			0.96	101
Macro Avg	0.65	0.53	0.58	101
Weighted Avg	0.94	0.96	0.95	101

Sumber : Hasil Penelitian (2025)

Tabel 5 menunjukkan kinerja *Decision Tree* dalam analisis sentimen data *twitter*. Model ini mencapai performa terbaik pada kategori netral dengan *F1-score* 0,98 dan kategori positif dengan *F1-score* 0,75, tetapi untuk mengklasifikasikan sentimen negatif dengan *F1-score* 0,00. Secara keseluruhan, akurasi model mencapai 0,96, menunjukkan bahwa *Decision Tree* sangat efektif dalam menganalisis sentimen netral namun memiliki keterbatasan pada kategori negatif.

Deep Learning

Penelitian ini mengevaluasi penerapan algoritma deep learning *Long Short-Term Memory* (LSTM) untuk analisis sentimen terhadap fenomena fufufafa di *twitter*. Model LSTM diterapkan untuk mengklasifikasikan sentimen *tweet* ke dalam kategori positif, negatif, dan netral setelah melalui proses prapemrosesan data. Hasil eksperimen menunjukkan bahwa LSTM mencapai akurasi 0.95%, dengan kinerja yang baik pada *metrik precision, recall, dan F1-score*. Model ini terbukti efektif dalam menangani konteks temporal dan memahami hubungan antar kata dalam *tweet*, termasuk sentimen yang lebih kompleks seperti ironi atau sindiran.

Tabel 6 Hasil Model *Long Short-Term Memory* (LSTM)

	Precision	Recall	F1-Score	Support
Negative	0.00	0.00	0.00	2
Neutral	0.95	1.00	0.97	94
Positive	1.00	0.40	0.57	5
Accuracy			0.95	101
Macro Avg	0.65	0.47	0.52	101
Weighted Avg	0.93	0.95	0.93	101

Sumber : Hasil Penelitian (2025)

Tabel 6 menunjukkan kinerja model LSTM dalam analisis sentimen data *twitter*. Model ini menunjukkan hasil terbaik pada kategori netral dengan *F1-score* 0,97, namun tidak dapat mengklasifikasikan sentimen negatif dengan *F1-score* 0,00. Pada kategori positif, model memperoleh *F1-score* sebesar 0,57. Secara keseluruhan, model LSTM mencapai akurasi 0,95, menunjukkan efektivitas tinggi pada kategori netral, namun masih memiliki keterbatasan dalam menangani kategori negatif dan positif.

Evaluasi

Hasil evaluasi ini kemudian disajikan dalam bentuk tabel perbandingan, seperti yang terlihat pada tabel 7, yang berisi ringkasan hasil *confusion matrix* untuk setiap skenario pengujian. Tabel ini memungkinkan analisis yang lebih mendalam mengenai kekuatan dan kelemahan model dalam mengklasifikasikan data. Dengan melihat distribusi nilai peneliti dapat mengidentifikasi pola kesalahan yang terjadi serta area yang memerlukan perbaikan, sehingga model dapat dioptimalkan untuk menghasilkan prediksi yang lebih akurat di masa depan.

Tabel 7 Confusion Matrix Machine Learning dan Deep Learning

Model	Accuracy	Precision	Recall	F-Measure
Naive Bayes	0.81	0.90	0.81	0.85
Random Forest	0.94	0.92	0.94	0.91
Support Vector Machine	0.94	0.92	0.94	0.91
Logistic Regression	0.94	0.91	0.94	0.92
Decision Tree	0.96	0.94	0.96	0.94
LSTM	0.95	0.93	0.95	0.93

Sumber : Hasil Penelitian (2025)

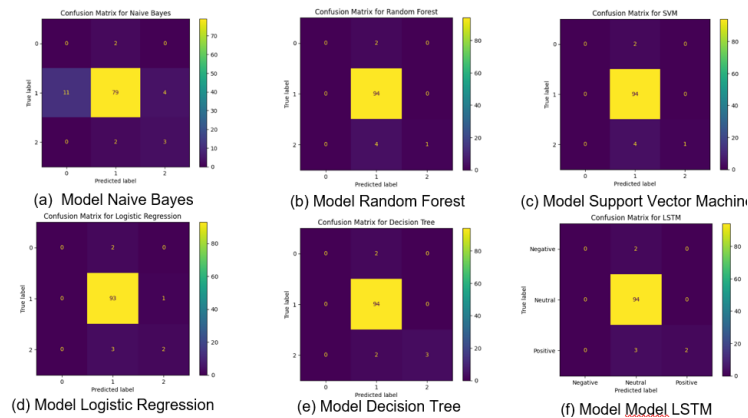
Tabel 7 memperlihatkan perbandingan kinerja beberapa model machine learning dan deep learning dalam analisis sentimen data *twitter*. Decision Tree mencatat akurasi tertinggi sebesar 0,96 dengan *F-Measure* 0,94, diikuti oleh LSTM dengan akurasi 0,95 dan *F-Measure* 0,93. Model *Random Forest*, *Support Vector Machine*, dan *Logistic Regression* memiliki akurasi yang sama sebesar 0,94, dengan *F-Measure* berkisar antara 0,91 hingga 0,92. *Naive Bayes* mencatat akurasi terendah sebesar 0,81 dan *F-Measure* 0,85, menunjukkan performa yang lebih rendah dibandingkan model lainnya. Tabel ini menggambarkan bahwa *Decision Tree* dan LSTM unggul dalam klasifikasi sentimen.

Tabel 8 Hasil Akurasi Machine Learning dan Deep Learning

Model	Accuracy
Naive Bayes	0.81
Random Forest	0.94
Support Vector Machine	0.94
Logistic Regression	0.94
Decision Tree	0.96
LSTM	0.95

Sumber : Hasil Penelitian (2025)

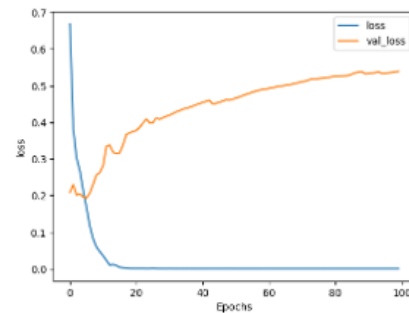
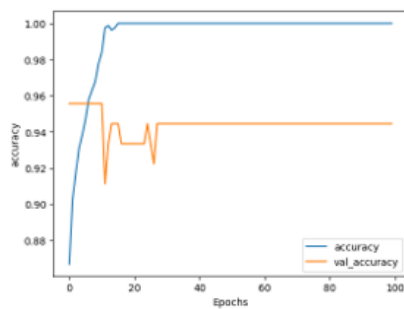
Tabel 8 menunjukkan akurasi berbagai model machine learning dan deep learning dalam analisis sentimen. Decision Tree memiliki akurasi tertinggi sebesar 0,96, diikuti oleh LSTM dengan 0,95. *Random Forest*, *Support Vector Machine*, dan *Logistic Regression* mencatatkan akurasi yang sama sebesar 0,94, sedangkan *Naive Bayes* memiliki akurasi terendah sebesar 0,81. Hasil ini menyoroti keunggulan *Decision Tree* dan LSTM dalam analisis sentimen.



Sumber : Hasil Penelitian (2025)

Gambar 6 Hasil Confusion Matrix Model Machine Learning dan Deep Learning

Gambar 6 ini menunjukkan matriks konfusi enam model (*Naive Bayes*, *Random Forest*, *SVM*, *Logistic Regression*, *Decision Tree*, dan *LSTM*). *Decision Tree* dan *LSTM* memiliki kinerja terbaik, sedangkan *Naive Bayes* menunjukkan performa rendah, terutama pada kategori negatif. Visualisasi ini mempermudah evaluasi akurasi klasifikasi setiap model.



Sumber : Hasil Penelitian (2025)

Gambar 7 Training dan Validation Accuracy Deep Learning LSTM

Gambar 7 menunjukkan akurasi pelatihan dan validasi model LSTM. Akurasi pelatihan stabil mendekati 1, dan validasi konsisten, menunjukkan performa model yang baik tanpa *overfitting*.

Visualisasi

Setelah data dibersihkan dan diklasifikasikan, hasil sentimen divisualisasikan menggunakan *wordcloud*, yang menunjukkan kata-kata paling sering muncul berdasarkan frekuensinya. Kata yang lebih besar mencerminkan kemunculan lebih tinggi, memudahkan identifikasi kata kunci dominan dalam data *teks*. Visualisasi ini dihasilkan menggunakan *script python*, seperti yang ditampilkan pada Gambar 8.



Sumber : Hasil Penelitian (2025)

Gambar 8 *Wordcloud*

4. Kesimpulan

Penelitian ini membandingkan kinerja berbagai model machine learning dan deep learning dalam analisis sentimen terhadap fenomena fufufafa di Twitter. Hasil menunjukkan bahwa Decision Tree memiliki akurasi tertinggi sebesar 96%, diikuti oleh LSTM dengan 95%, sementara Random Forest, SVM, dan Logistic Regression mencapai 94%, dan Naive Bayes mencatatkan akurasi sebesar 81% karena keterbatasan dalam menangani sentimen kompleks. LSTM terbukti unggul dalam menangkap konteks temporal dan hubungan antar kata, meskipun membutuhkan lebih banyak sumber daya komputasi. Hasil ini mengonfirmasi bahwa model deep learning, khususnya LSTM, lebih efektif dalam menganalisis sentimen yang dinamis dan beragam dibandingkan model machine learning. Penelitian ini memberikan dasar bagi pengembangan model yang lebih efisien dan penerapan analisis sentimen dalam berbagai konteks, seperti media sosial lainnya, untuk memahami opini publik secara lebih mendalam dan akurat. Ke depan, penelitian ini dapat dikembangkan dengan mengintegrasikan model transformer seperti IndoBERT untuk meningkatkan akurasi dan kemampuan menangkap nuansa bahasa yang lebih kompleks, serta memperluas penerapan ke platform media sosial dan bidang analisis opini publik lainnya guna mendukung pengambilan keputusan berbasis data secara lebih efektif.

Referensi

- [1] J. P. · A. P. G. · D. M. · S. M. and S. R. · A. G. · A. D. · S. M. · J. Si, "A survey and comparative study on negative sentiment analysis in social media data," *Multimed. Tools Appl.*, 2024.
- [2] A. G. Himani Maheshwari, Dharmendra Yadav , Umesh Chandra, "Twitter Sentiment Analysis in the Crisis Between Russia and Ukraine Using the Bert and LSTM Model," *Int. Conf. Comput. Commun. Intell. Syst.*, 2023.
- [3] J. V. T. · K. L. · A. Agrawal, "Twitter sentiment analysis using ensemble of multi-channel model based on machine learning and deep learning techniques," *Springer*, 2024.
- [4] A. A. Rizal Fitrianto, Arda Surya Editya, Mochammad Alamin, Anggay Pramana, "Classification of Indonesian Sarcasm Tweets on X Platform Using Deep Learning," *2024 7th Int. Conf. Informatics Comput. Sci.*, 2024.
- [5] and U. N. Chih-Hsueh Lin, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J. Big Data*, 2023.
- [6] A. U. I. K. Ahmet Göçen, Mahat Maalim Ibrahim, "Public attitudes toward higher education using sentiment analysis and topic modeling," *Discov. Artificial Intell.*, 2024.
- [7] S. H. A. Silviya, · S. Julia Faith, · R. Seetha, and · M. Hemalatha, "Performance Evaluation of Natural Language Processing Algorithms for Sentiment Analysis," *Springer Nat. Comput. Sci.*, 2024.
- [8] Y. M. Andry Chowanda, "Indonesian Sentiment Analysis Model from Social Media by Stacking BERT and BI-LSTM," *2022 3rd Int. Conf. Artif. Intell. DATA Sci.*, 2022.
- [9] A. Muhamad Fahmi, Faturahman Yudanto, Naurah Nazhifah, Yunita Sari, "Deep Learning Approach for Aspect-Based Sentiment Analysis on Indonesian Hospitals

- Reviews," *IEEE*, 2024.
- [10] S. S. Muhamad Rizky Yanuar, "Aspect Extraction for Tourist Spot Review in Indonesian Language using BERT," *IEEE*, 2024.
 - [11] S. R. Feliks Victor Parningotan Samosir, "Sentiment Analysis of TikTok Comments on Indonesian Presidential Elections Using IndoBERT," *IEEE*, 2024.
 - [12] H. R. Hendri Ahmadian, Taufik Fuadi Abidin, "Transformer-based Indonesian Language Model for Emotion Classification and Sentiment Analysis," *Int. Conf. Inf. Technol. Computing*, 2023.
 - [13] J. Computer and V. P. C. Christiella Setiawan, Shilvia Meidhi, Ivan Parmonangan, "Sentiment Analysis of Skincare Product Reviews in Indonesian Language using IndoBERT and LSTM," *IEEE 9th Inf. Technol. Int. Semin.*, 2023.
 - [14] B. K. P. Shreyas Agrawal, Sumanto Dutta, "Sentiment Analysis of Short Informal Text by Tuning BERT - Bi-LSTM Model," *IEEE*, 2021.
 - [15] B. S. Kundeti Naga Prasanthi, Rallabandi Eswari Madhavi, Degala Naga Sai Sabarinadh, "A Novel Approach for Sentiment Analysis on social media using BERT & ROBERTA TransformerBased Models," *IEEE 8th Int. Conf. Conver. Technol.*, 2023.
 - [16] J. L. Yifan Shen, "Comparison of Text Sentiment Analysis based on Bert and Word2vec," *IEEE 3rd Int. Conf. Front. Technol. Inf. Comput.*, 2021.
 - [17] P. H. Kui Zhang, Fuji Ren, Min Hu, "Sentiment Analysis of Chinese Product Reviews Based on BERT Word Vector and Hierarchical Bidirectional LSTM," *IEEE Int. Conf. Comput. Sci. Artif. Intell. Electron. Eng.*, 2021.
 - [18] X. L. Jianqiong Xiao, "Aspect-level sentiment analysis based on BERT fusion multi-attention," *2022 14th Int. Conf. Intell. Human-Machine Syst. Cybern.*, 2022.
 - [19] X. Z. Zesheng Shi, Tianhao Cao, "Incorporating BERT with Naive Bayes into Neutral Sentiment Analysis," *IEEE 3rd Int. Conf. Data Sci. Comput. Appl.*, 2023.
 - [20] B. V. K. Yalanati Ayyappa, S. A. Sudhabatthula Padma, Priya Student, and S. G. Tamma Priya, Vardhan Reddy, "Forecasting Equity Prices using LSTM and BERT with Sentiment Analysis," *Proc. Int. Conf. Inven. Comput. Technol. (ICICT 2023)*, 2023.
 - [21] B. Jlif, · Chaima Abidi, and · Claude Duvallet, "Beyond the use of a novel Ensemble based Random Forest-BERT Model (Ens-RF-BERT) for the Sentiment Analysis of the hashtag COVID19 tweets," *Soc. Netw. Anal. Min.*, 2024.
 - [22] M. S. I. Fahmida Khanam, Anik Chakraborty, Md Habib, "Bangla Sentiment Analysis On Highly Imbalanced Data Using Hybrid CNN-LSTM & Bangla BERT," *2024 3rd Int. Conf. Adv. Electr. Electron. Eng. (ICAEEE 2024)*, 2024.
 - [23] R. M. R Prasanna Kumar, Bharathi Mohan G, Elakkiya R, Charan Kumar M, "Automated Sentiment Classification of Amazon Product Reviews using LSTM and Bidirectional LSTM," *Int. Conf. Evol. Algorithms Soft Comput. Tech.*, 2023.
 - [24] B. R. Roy, M. H. O. Rashid, M. M. Haque, and A. Sarker, "A Comparative Study of Sentiment Analysis on Public Opinion Data Using BERT and LSTM Architecture Based on a Transformers Approach," *26th Int. Conf. Comput. Inf. Technol.*, 2023.
 - [25] F. P. Y. Evaristus Didik Madyatmadja, Brigita Christabel Surya Winata, Evelyn Pradhan and C. and D. J. M. S. Feladiva Annrastia Adrian, Raihan Mahardhika, "Sentiment Analysis on User Reviews of Snapchat in Indonesia," *J. Comput. Sci.*, p. 158.167, 2024, doi: 10.3844/jcssp.2025.158.167.
 - [26] B. S. Ainapure, R. N. Pise, P. Reddy, B. Appasani, A. Srinivasulu, and M. S. K. and N. Bizon, "Sentiment Analysis of COVID-19 Tweets Using Deep Learning and Lexicon-Based Approaches," <https://www.mdpi.com/journal/sustainability>, 2023.
 - [27] S. Z. and S. W. Jiawei Xue, Yanhong Li, Zixuan Li, Yue Cui, "A Cross-Domain Generative Data Augmentation Framework for Aspect-Based Sentiment Analysis," <https://www.mdpi.com/journal/electronics>, 2023.
 - [28] Z. M. and H. Z. Hongchan Li, Yu Ma, "Weibo Text Sentiment Analysis Based on BERT and Deep Learning," <https://www.mdpi.com/journal/applsci>, 2021.
 - [29] H. W. Zhu He and and X. Zhang, "Multi-Task Learning Model Based on BERT and Knowledge Graph for Aspect-Based Sentiment Analysis," <https://www.mdpi.com/journal/electronics>, 2023.
 - [30] B. Omarov and Z. Zhumanov, "Bidirectional Long-Short-Term Memory with Attention Mechanism for Emotion Analysis in Textual Content," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, 2023.

- [31] E. D. and A. Chowanda, "Indonesian personal financial entity extraction using indoBERT-BiGRU-CRF model," *Dave Chowanda J. Big Data*, pp. 2–24, 2024.
- [32] L. A. and Y. Zhang, "A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning," <https://www.mdpi.com/journal/computers>, 2024.
- [33] N. C. Dang, M. N. Moreno-García, and and F. D. la Prieta, "Sentiment Analysis Based on Deep Learning: A Comparative Study," <http://www.mdpi.com/journal/electronics>, 2020.
- [34] W. and R. K. Ahmad Jazuli, "Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback," <https://www.mdpi.com/journal/applsci>, pp. 1–28, 2024.
- [35] A. N. Ma'aly and A. D. F. and H. F. , Dita Pramesti, "Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm," <https://www.mdpi.com/journal/information>, pp. 1–33, 2024.
- [36] W. O. V. Darryl Rayhan Wijaya, Gusti Made Arya Sasmita, "Sentiment Analysis of Indonesian Citizens on Electric Vehicle Using FastText and BERT Method," *J. Inf. Syst. Informatics*, vol. 6, no. 3, pp. 1360–1372, 2024.