JIIFKOM (Jurnal Ilmiah Informatika & Komputer) STTR Cepu
E-Mail: jiifkom@sttrcepu.ac.id

Analisis Sentimen Komentar YouTube Berbahasa Indonesia Menggunakan Algoritma *Machine Learning* dan *Deep Learning*

Fahreza Alhadi[a*], Adhika Pramita Widyassari[b]
fahrezaalhadi47@gmail.com[a*]

Sekolah Tinggi Teknologi Ronggolawe Cepu

Kata kunci:
Analisis sentimen, Machine learning, Deep learning

Tentang naskah:
-diterima, 15 Januari 2026
-diterbitkan, 26 Januari 2026

Keyword:
Sentiment analysis, Machine Learning, Deep learning

Intisari

Penelitian ini membahas analisis sentimen komentar YouTube berbahasa Indonesia dengan membandingkan kinerja algoritma machine learning dan deep learning. Data penelitian berupa 364 komentar YouTube yang dikumpulkan menggunakan YouTube Data API dari sebuah video debat publik terkait kasus meme Bahlil Lahadalia. Komentar diklasifikasikan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral, dengan distribusi kelas yang tidak seimbang. Tahapan preprocessing teks meliputi pembersihan teks, normalisasi, tokenisasi, stopword removal, dan stemming menggunakan Sastrawi. Representasi fitur TF-IDF unigram dan bigram digunakan untuk model machine learning, sedangkan LSTM menggunakan pendekatan embedding berbasis tokenisasi. Algoritma yang diuji meliputi Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor, Decision Tree, Multi-Layer Perceptron (MLP), dan LSTM. Evaluasi performa dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa SVM kernel linear memiliki performa terbaik, diikuti oleh MLP dan Naive Bayes, sedangkan LSTM menunjukkan performa terendah akibat keterbatasan ukuran dan distribusi dataset. Temuan ini menunjukkan bahwa algoritma machine learning tradisional masih efektif untuk analisis sentimen berbahasa Indonesia pada data terbatas.

Abstract

YouTube provides a comment section that reflects public sentiment toward an issue. The large volume of comments makes manual analysis inefficient, requiring automated sentiment analysis. This study compares the performance of machine learning and deep learning algorithms in classifying Indonesian YouTube comments. The dataset consists of 364 comments collected using the YouTube Data API and classified into positive, negative, and neutral sentiments. The data were processed through text preprocessing and represented using TF-IDF for machine learning models, while the deep learning model employed a Long Short-Term Memory (LSTM) architecture. The evaluated algorithms include Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor, Decision Tree, Multi-Layer Perceptron (MLP), and LSTM. Model performance was evaluated using accuracy, precision, recall, and F1-score. The results show that SVM achieved the best performance, followed by MLP and Naive Bayes, while LSTM performed the worst due to limited data size and class distribution. This study indicates that traditional machine learning remains effective for sentiment analysis on small Indonesian-language datasets.

1. Pendahuluan

Media sosial telah menjadi salah satu sarana utama bagi masyarakat dalam menyampaikan opini, pandangan, serta respons terhadap berbagai isu dan konten digital. YouTube sebagai platform berbagi video dengan tingkat interaksi pengguna yang tinggi menyediakan kolom komentar yang aktif, dimana pengguna bebas mengekspresikan pendapatnya. Komentar-komentar tersebut merepresentasikan sentimen publik terhadap suatu konten dan berpotensi memberikan informasi berharga mengenai persepsi serta respons masyarakat secara luas, termasuk dalam konteks isu politik dan debat publik (Fahrezi et al., 2025). Namun, tingginya volume komentar menyebabkan analisis secara manual menjadi tidak efisien, sehingga diperlukan pendekatan otomatis berbasis komputasi untuk memetakan kecenderungan opini masyarakat (Mufriz, 2024).

Analisis sentimen merupakan salah satu bidang dalam *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini atau sikap pengguna terhadap suatu objek berdasarkan data teks. Dalam konteks Bahasa Indonesia, analisis sentimen menghadapi tantangan tersendiri, antara lain penggunaan bahasa tidak baku, campuran bahasa asing, singkatan, serta ragam bahasa informal atau bahasa gaul. Tantangan tersebut banyak dijumpai pada komentar YouTube, sehingga tahapan pra-pemrosesan (*preprocessing*) menjadi krusial untuk memastikan data teks yang bersih sebelum diproses oleh model klasifikasi (Misrun et al., 2024).

Berbagai pendekatan telah diterapkan dalam analisis sentimen, baik menggunakan algoritma *machine learning* maupun *deep learning*. Algoritma *machine learning* tradisional seperti *Naive Bayes*, *Support Vector Machine* (SVM), dan Multi-Layer Perceptron (MLP) diketahui memiliki kinerja yang relatif stabil pada dataset berukuran kecil hingga menengah. Studi oleh (Wijaya et al., 2025) menunjukkan bahwa SVM mampu memberikan akurasi yang baik dalam klasifikasi komentar YouTube dengan bantuan leksikon, sementara (Maulana & Fahmi, 2024) menemukan bahwa SVM sering kali mengungguli Naive Bayes dalam hal akurasi pada data opini publik. Di sisi lain, pendekatan *deep learning* seperti *Long Short-Term Memory* (LSTM) memiliki kemampuan dalam mempelajari pola bahasa yang lebih kompleks dan kontekstual, yang sangat relevan untuk menganalisis teks debat politik yang dinamis (Fahrezi et al., 2025). Meskipun demikian, model seperti LSTM umumnya memerlukan jumlah data yang lebih besar dan sumber daya komputasi yang lebih tinggi untuk mencapai performa optimal. Perbedaan karakteristik tersebut menimbulkan kebutuhan untuk membandingkan kinerja kedua pendekatan dalam konteks data berbahasa Indonesia.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melakukan komparasi kinerja beberapa algoritma *machine learning* dan *deep learning* dalam mengklasifikasikan sentimen komentar YouTube berbahasa Indonesia. Data penelitian dikumpulkan melalui proses *crawling* menggunakan YouTube Data API, kemudian melalui tahapan *preprocessing* teks sebelum digunakan dalam proses pelatihan dan pengujian model klasifikasi.

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menghasilkan dataset komentar YouTube berbahasa Indonesia yang dapat digunakan untuk analisis sentimen.
2. Menganalisis pengaruh tahapan *preprocessing* teks terhadap kualitas data dan kinerja model klasifikasi.
3. Mengevaluasi kinerja beberapa algoritma *machine learning* dan *deep learning* dalam tugas klasifikasi sentimen.
4. Menentukan algoritma dengan performa terbaik berdasarkan metrik evaluasi yang digunakan.

Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai efektivitas berbagai algoritma klasifikasi dalam analisis sentimen berbahasa Indonesia, khususnya pada dataset dengan ukuran terbatas.

2. Kerangka Teori

2.1 *Analisis Sentimen*

Analisis sentimen merupakan salah satu cabang dalam *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini, sikap, atau penilaian pengguna terhadap suatu objek berdasarkan data teks. Dalam penelitian analisis sentimen, teks umumnya diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral untuk memahami kecenderungan opini publik secara otomatis dan efisien (Utari & Wibowo, 2025).

2.2 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) merupakan bidang ilmu yang mempelajari interaksi antara komputer dan bahasa manusia. NLP memungkinkan sistem komputer untuk memahami, mengolah, dan menganalisis bahasa alami dalam bentuk teks maupun ujaran, serta berperan penting dalam keseluruhan proses pengolahan teks — mulai dari *preprocessing* hingga representasi fitur, sehingga data teks dapat diproses secara efektif oleh algoritma klasifikasi (Nursyanti et al., 2024).

2.3 *Preprocessing Teks*

Preprocessing teks merupakan tahapan awal dalam pengolahan data teks yang bertujuan untuk membersihkan dan

menyiapkan data sebelum dilakukan pemodelan. Tahapan umum meliputi *case folding, cleaning, tokenization, stopword removal,* dan *stemming*. Proses ini meningkatkan kualitas data sehingga berdampak positif terhadap kinerja model klasifikasi — sesuai dengan praktik yang dilaporkan dalam studi perbandingan algoritma NLP (Utari & Wibowo, 2025).

Pelabelan sentimen pada data komentar dilakukan secara otomatis menggunakan aturan *sentiment score* berbasis leksikon, di mana nilai skor digunakan untuk menentukan kelas sentimen (positif, negatif, netral) sebagai ground truth — pendekatan yang banyak dikombinasikan dengan pembelajaran mesin dalam literatur NLP Bahasa Indonesia (Nursyanti et al., 2024).

2.4 Representasi Fitur Teks

Representasi fitur teks bertujuan untuk mengubah data teks menjadi bentuk numerik agar dapat diproses oleh algoritma *machine learning*. Salah satu metode representasi fitur yang umum digunakan dalam analisis sentimen berbasis teks adalah Term Frequency–Inverse Document Frequency (TF-IDF), yaitu metode pembobotan kata yang mempertimbangkan frekuensi kemunculan suatu kata dalam sebuah dokumen serta tingkat kepentingannya terhadap keseluruhan korpus.

Dalam penelitian ini, pembobotan TF-IDF dihitung secara otomatis menggunakan TfidfVectorizer dari library scikit-learn untuk menghasilkan representasi vektor numerik berbasis unigram dan bigram. Representasi ini digunakan sebagai masukan bagi model klasifikasi sentimen berbasis *machine learning*, yaitu *Naive Bayes, Decision Tree, K-Nearest Neighbor, dan Support Vector Machine*. Penggunaan TF-IDF memungkinkan model untuk mengekstraksi informasi penting dari teks komentar YouTube berdasarkan bobot kata yang relevan terhadap kelas sentimen, sejalan dengan temuan (Sutriawan; Mutmainnah, Siti; Lorosae, Teguh Ansyor; Ramadhan, 2025) yang menyatakan bahwa representasi TF-IDF dan TF-IDF berbasis n-gram efektif dalam menangkap informasi tekstual yang relevan untuk meningkatkan kinerja algoritma klasifikasi teks.

2.5 Machine Learning untuk Analisis Sentimen

Pendekatan *machine learning* sering digunakan dalam analisis sentimen karena kemampuannya dalam klasifikasi teks secara efisien dengan kebutuhan komputasi yang relatif rendah. Beberapa algoritma umum termasuk *Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN),* dan *Decision Tree*. Dalam literatur nasional, SVM dan *Naive Bayes* banyak diuji kinerjanya untuk komentar bahasa Indonesia — dengan SVM cenderung unggul dalam memisahkan kelas sentimen (Pandie et al., 2025; Utari & Wibowo, 2025)

2.6 *Deep Learning* untuk Analisis Sentimen

Deep learning merupakan pendekatan berbasis jaringan saraf dalam yang mampu mempelajari pola dan representasi kompleks dari data. Long Short-Term Memory (LSTM) merupakan salah satu arsitektur dalam Recurrent Neural Network (RNN) yang dirancang untuk memproses data berurutan serta mempertahankan informasi jangka panjang. Dalam analisis sentimen, LSTM sering digunakan karena kemampuannya dalam menangkap ketergantungan antar kata dalam urutan teks, sehingga dapat menghasilkan pemodelan konteks yang lebih baik dibandingkan pendekatan *machine learning* tradisional.

Meskipun demikian, penerapan LSTM pada dataset berukuran kecil dan tidak seimbang perlu dilakukan secara hati-hati karena berpotensi menyebabkan overfitting atau hasil evaluasi yang kurang stabil, sejalan dengan temuan (Naquitasia, Risca; Fudholi, Dhomas Hatta; Iswari, 2022) yang menunjukkan bahwa performa model *deep learning* dapat dipengaruhi oleh keterbatasan jumlah data dan tahapan prapemrosesan, sehingga berpotensi menimbulkan kesalahan prediksi pada pengujian tertentu.

3. Metodologi

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasional. Eksperimen dilakukan dengan membandingkan kinerja beberapa algoritma klasifikasi sentimen berbasis *machine learning* dan *deep learning* terhadap data komentar YouTube berbahasa Indonesia. Variabel bebas dalam penelitian ini adalah jenis algoritma klasifikasi yang digunakan, sedangkan variabel terikatnya adalah

performa model klasifikasi sentimen yang diukur menggunakan metrik evaluasi klasifikasi.

3.2 Sumber Data

Data penelitian diperoleh dari komentar pada satu video YouTube mengenai kasus penyebaran meme Bahlil Lahadalia yang diproses secara hukum. Video yang dianalisis berjudul "Sengit! Debat Feri Amsari–Waketum AMPG Soal Kasus Meme Bahlil Lahadalia: Bawa Delik Agama? (https://youtu.be/Mi_Tzg8gTpU?si=mrA8yqX4LrOctyya)", yang diunggah pada 22 Oktober 2025. Pengambilan data dilakukan menggunakan YouTube Data API melalui Google API Client untuk mengekstraksi komentar secara otomatis. Seluruh komentar yang dikumpulkan merupakan komentar berbahasa Indonesia yang merepresentasikan tanggapan pengguna terhadap konten video tersebut.

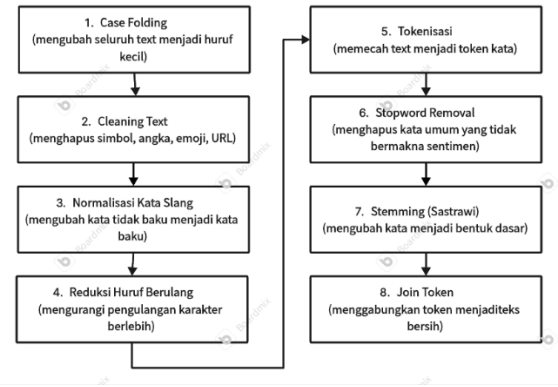
Total data yang berhasil dikumpulkan berjumlah 364 komentar, yang terdiri dari 78 komentar positif, 55 komentar negatif, dan 231 komentar netral. Data komentar ini selanjutnya digunakan sebagai sumber utama dalam proses analisis sentimen untuk mengidentifikasi kecenderungan opini publik terhadap isu pemolisian kasus meme Bahlil Lahadalia yang dibahas dalam video tersebut..

3.3 Proses Pengumpulan Data

Pengumpulan data dilakukan secara otomatis melalui proses crawling menggunakan YouTube Data API. Data yang diperoleh berupa teks mentah (raw text) komentar pengguna. Sebelum tahap *preprocessing*, data diperiksa untuk memastikan tidak terdapat komentar kosong maupun komentar duplikat, sehingga data yang digunakan layak untuk diproses lebih lanjut dalam tahap analisis sentimen.

3.4 Preprocessing Teks

Tahapan *preprocessing* teks dilakukan untuk meningkatkan kualitas data sebelum dilakukan ekstraksi fitur dan pelatihan model. Tahapan *preprocessing* yang diterapkan meliputi case folding, pembersihan teks dengan menghapus simbol, angka, emoji, dan URL, normalisasi kata untuk menyamakan bentuk kata tidak baku, reduksi huruf berulang, tokenisasi untuk memecah teks menjadi unit kata, penghapusan stopword, serta stemming menggunakan algoritma Sastrawi untuk mengubah kata ke bentuk dasarnya. Tahapan *preprocessing* bertujuan untuk membersihkan dan menormalkan teks sebelum dilakukan tahap representasi fitur. Alur *preprocessing* teks secara keseluruhan ditunjukkan pada Gambar 3.1:



Gambar 3.1 Diagram Alur Preprocessing Teks

3.5 Algoritma Klasifikasi

Penelitian ini menggunakan enam algoritma klasifikasi yang terdiri dari pendekatan *machine learning* dan *deep learning*. Algoritma *machine learning* yang digunakan meliputi *Naive Bayes, Support Vector Machine* (SVM) dengan kernel linear, *K-Nearest Neighbor* (KNN), dan *Decision Tree*. Selain itu, digunakan Multi-Layer Perceptron (MLP) sebagai representasi jaringan saraf berbasis *machine learning*, serta *Long Short-Term Memory* (LSTM) sebagai pendekatan *deep learning*. Pemilihan algoritma tersebut bertujuan untuk membandingkan kinerja model klasik dan model berbasis neural network dalam tugas klasifikasi sentimen.

3.6 Pembagian Data

Dataset dibagi menjadi data latih dan data uji dengan proporsi 80% untuk data latih dan 20% untuk data uji. Pembagian data dilakukan secara acak dengan mempertahankan proporsi kelas sentimen pada masing-masing subset data. Pembagian ini diterapkan secara konsisten pada seluruh model untuk memastikan keadilan dalam perbandingan performa.

3.7 Perangkat dan Lingkungan Kerja

Eksperimen dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab. Library yang digunakan meliputi scikit-learn untuk implementasi algoritma *machine learning* serta Keras untuk pemodelan LSTM. Lingkungan komputasi berbasis cloud ini digunakan karena kemudahan akses serta

dukungan terhadap pemrosesan data dan pelatihan model klasifikasi teks.

3.8 Evaluasi Model

Evaluasi performa model dilakukan menggunakan beberapa metrik evaluasi, yaitu accuracy, precision, recall, f1-score, serta confusion matrix. Penggunaan lebih dari satu metrik evaluasi bertujuan untuk memberikan gambaran kinerja model secara komprehensif, khususnya dalam kondisi distribusi kelas yang tidak seimbang. Metrik precision, recall, dan f1-score digunakan untuk menilai kemampuan model dalam mengklasifikasikan masing-masing kelas sentimen secara lebih representatif dibandingkan hanya menggunakan akurasi.

4. Hasil dan Pembahasan

4.1 Hasil Preprocessing

Tahap *preprocessing* menghasilkan data teks yang telah dibersihkan dan dinormalisasi sehingga siap digunakan pada tahap pelabelan sentimen dan pemodelan klasifikasi. Proses ini bertujuan untuk mengurangi noise pada komentar YouTube yang umumnya mengandung variasi bahasa tidak baku, pengulangan huruf, simbol, serta struktur kalimat informal.

Tabel 4.1 menyajikan contoh hasil *preprocessing* terhadap lima komentar YouTube yang mewakili tiga kelas sentimen, yaitu positif, negatif, dan netral. Kolom teks awal menunjukkan komentar sebelum dilakukan *preprocessing*, sedangkan kolom hasil *preprocessing* menunjukkan teks setelah melalui seluruh tahapan pembersihan dan normalisasi.

Tabel 4.1 Hasil Preprocessing Komentar YouTube

No	Teks Awal Komentar	Hasil Eprocessing	Sentimen
1	KALAU TIDAK MAU DIKRITIK JANGAN MENJADI PEJABAT PUBLIK PAK ,, Lebay	kritik pejabat publik lebay	Negatif
2	Pak Bahlil orang baik dan bijaksana dan waktunya kebanyakan di habiskan bersama rakyat, sehat selalu rakyat bersamamu 🩶	bahlil orang bijaksana banyak habis rakyat sehat rakyat sama	
3	Sigoblok meme doang langsung lapor!! Bahlul bahlull! Woi polisi, urus yg lebih penting!!	sigoblok meme doang langsung lapor bahlil bahlull woi polisi urus	Negatif
4	Perbaiki saja kinerjanya dan buktikan dengan prestasi yang pro rakyat. Pasti hilang semua meme meme itu. Bahkan bisa jadi berubah menjadi lebih baik.	baik kerja bukti prestasi pro rakyat hilang meme meme ubah	Netral
5	Klau di proses hukum ya di buktikan dulu betul gak pak bahlil kerja sdh bener apa belum.	kalau proses hukum iya bukti bahlil kerja	Netral

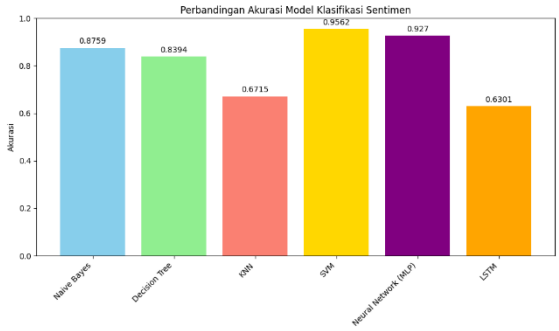
4.2 Hasil Pengujian Model Klasifikasi Sentimen

Setelah melalui tahapan *preprocessing* teks, model *machine learning* dilatih menggunakan representasi fitur TF-IDF menggunakan 80% data latih dan diuji menggunakan 20% data uji, sedangkan model LSTM dilatih menggunakan representasi berbasis tokenisasi dan embedding. Evaluasi awal dilakukan menggunakan metrik akurasi untuk memberikan gambaran umum performa masing-masing model. Hasil pengujian ditunjukkan pada Tabel 4.2:

Tabel 4.2 Hasil Pengujian Model Klasifikasi Sentimen

Model	**Akurasi**
Naive Bayes	0.8759
Decision Tree	0.8394
K-Nearest Neighbor	0.6715

SVM (Linear)	0.9562
Neural Network (MLP)	0.9270
LSTM	0.6301



Gambar 4.2 Diagram Batang Perbandingan Akurasi Model

Berdasarkan Tabel 4.2, model Support Vector Machine (SVM) dengan kernel linear menghasilkan nilai akurasi tertinggi, diikuti oleh model Neural Network (NN) dan Naive Bayes. Sementara itu, model LSTM menunjukkan nilai akurasi terendah dibandingkan model lainnya.

Meskipun demikian, perlu ditekankan bahwa penggunaan akurasi sebagai satu-satunya indikator kinerja ***tidak cukup representatif***, terutama pada dataset dengan distribusi kelas yang tidak seimbang seperti pada penelitian ini, di mana kelas sentimen netral mendominasi data. Oleh karena itu, evaluasi lanjutan menggunakan metrik lain diperlukan untuk memperoleh gambaran kinerja model yang lebih adil dan komprehensif.

4.3 Evaluasi Performa Model Menggunakan Metrik Klasifikasi

Untuk menilai performa model secara lebih menyeluruh, evaluasi lanjutan dilakukan menggunakan metrik precision, recall, dan F1-score. Metrik-metrik ini lebih informatif dalam kondisi distribusi kelas yang tidak seimbang karena memperhatikan kemampuan model dalam mengklasifikasikan masing-masing kelas sentimen. Hasil evaluasi ditunjukkan pada Tabel 4.3

Tabel 4.3 Hasil Evaluasi metrix

Model	**Precision**	**Recall**	**F1-Score**
Naive Bayes	0.90	0.88	0.87
Decision Tree	0.84	0.84	0.84
KNN	0.77	0.67	0.59
SVM	**0.96**	**0.96**	**0.96**
MLP	0.93	0.93	0.93

Pada penelitian ini, evaluasi lanjutan menggunakan precision, recall, dan F1-score belum diterapkan pada model LSTM, karena pipeline evaluasi yang digunakan pada model *deep learning* hanya menghasilkan metrik akurasi. Selain itu, penelitian ini bersifat eksploratif dan menggunakan pelabelan otomatis dengan distribusi kelas yang tidak seimbang, sehingga evaluasi mendalam terhadap model LSTM belum menjadi fokus utama. Oleh karena itu, performa LSTM hanya dianalisis berdasarkan nilai akurasi dan tidak dibandingkan secara langsung dengan model *machine learning* klasik menggunakan metrik precision, recall, dan F1-score. Keterbatasan ini menjadi salah satu arah pengembangan penelitian selanjutnya.

4.4 Pembahasan Hasil Eksperimen

Hasil eksperimen menunjukkan bahwa model *machine learning* tradisional, khususnya SVM dan MLP, mampu memberikan performa klasifikasi sentimen yang lebih stabil dibandingkan *model deep learning* LSTM. Keunggulan SVM dapat dikaitkan dengan kemampuannya dalam menangani data berdimensi tinggi yang dihasilkan oleh representasi fitur TF-IDF, serta karakteristiknya yang relatif robust terhadap ukuran dataset yang terbatas.

Model Naive Bayes juga menunjukkan performa yang cukup baik, meskipun menggunakan asumsi independensi antar fitur yang sederhana. Sebaliknya, model KNN dan Decision Tree menunjukkan performa yang lebih rendah, yang dapat disebabkan oleh sensitivitas terhadap distribusi data dan karakteristik ruang fitur.

Model LSTM menunjukkan performa yang paling rendah dalam penelitian ini. Hal ini mengindikasikan bahwa pendekatan *deep learning* memerlukan jumlah data yang lebih besar dan konfigurasi pelatihan yang lebih kompleks agar dapat mempelajari pola bahasa secara efektif. Dengan ukuran dataset yang relatif kecil dan distribusi kelas yang tidak seimbang, LSTM belum mampu menunjukkan keunggulannya dibandingkan model *machine learning* tradisional. Temuan ini sejalan dengan karakteristik umum model *deep learning*

yang sangat bergantung pada ketersediaan data dalam jumlah besar.

5. Simpulan

Penelitian ini menyimpulkan bahwa pendekatan *machine learning* dan *deep learning* dapat digunakan untuk klasifikasi sentimen komentar YouTube berbahasa Indonesia ke dalam kategori positif, negatif, dan netral, dengan *preprocessing* teks serta representasi fitur TF-IDF yang terbukti efektif. Hasil eksperimen menunjukkan bahwa model *machine learning* tradisional, khususnya SVM kernel linear dan Neural Network (MLP), memiliki performa yang lebih stabil dan unggul, terutama SVM yang mampu memberikan nilai precision, recall, dan F1-score yang tinggi meskipun terdapat ketidakseimbangan kelas. Sebaliknya, model *deep learning* LSTM menunjukkan performa yang kurang optimal akibat keterbatasan ukuran dan distribusi dataset. Temuan ini menegaskan bahwa pemilihan algoritma harus mempertimbangkan karakteristik data, serta menunjukkan bahwa *machine learning* tradisional masih menjadi solusi yang efektif dan efisien untuk analisis sentimen berbasis teks pada kondisi data terbatas.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada pihak-pihak yang telah mendukung tersusunnya penelitian ini. Penelitian ini disusun sebagai bagian dari pemenuhan tugas Ujian Akhir Semester pada mata kuliah Big Data Analysis. Penulis juga mengucapkan terima kasih kepada platform YouTube sebagai sumber data yang digunakan dalam penelitian ini. Semoga penelitian ini dapat memberikan manfaat dan menjadi referensi bagi pengembangan penelitian selanjutnya di bidang analisis sentimen dan pengolahan bahasa alami.

Daftar Pustaka

Artikel jurnal:

Fahrezi, M., Pratama, Y. B., & Pramudiyantoro, A. (2025). *Analisis Sentimen Debat Publik Pilpres 2024 Menggunakan Metode Algoritma LSTM dan IndoBERT Pada Platform Youtube*. *02*(03), 1936–1961.

Maulana, B. A., & Fahmi, M. J. (2024). *Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)*. *4*(April), 375–384.

Misrun, C. A., Haerani, E., Fikry, M., & Budianita, E. (2024). *Jurnal Computer Science and Information Technology (CoSciTech) naive bayes classifier method*. *4*(1), 207–215.

Mufriz, F. (2024). *Analisis Sentimen Pada Komentar Youtube Untuk Mengetahui Pandangan Masyarakat Kepada Calon Presiden Indonesia 2024 Menggunakan Algoritma Support Vector Machine*. *11*(4), 4257–4263.

Naquitasia, Risca; Fudholi, Dhomas Hatta; Iswari, L. (2022). *Analisis Sentimen Berbasis Aspek Pada Wisata Halal Dengan Metode Deep Learning*. *16*(2), 156–164.

Nursyanti, R., Alamsyah, N., & Yusuf, T. M. (2024). *Perbandingan Dan Efektivitas Kinerja Algoritma Svm Dan Bi-Lstm Dengan Tf-Idf Dan Smote Untuk Analisis Sentimen Kelestarian Hewan Liar*. *7*(1), 137–145.

Pandie, E. S. Y., Fanggidae, A., & Lona, R. (2025). *Perbandingan Naïve Bayes dan K-NN dalam Analisis Sentimen Aplikasi X*. *22*(2), 97–107.

Sutriawan; Mutmainnah, Siti; Lorosae, Teguh Ansyor; Ramadhan, S. (2025). *Model Text Embedding dan TF-IDF + Ngram untuk Meningkatkan Kinerja Algoritma Binary Classifier pada Klasifikasi SMS Palsu*. *4*(1), 55–64.

Utari, E. L., & Wibowo, S. H. (2025). *Analisis komparatif algoritma SVM, Naive Bayes, dan LSTM pada sentimen komentar lagu Labour*. *7*(3), 1276–1286. https://doi.org/10.51401/jinteks.v7i3.6213

Wijaya, F., Ramadhani, R. A., & Setiawan, A. B. (2025). *Analisis Sentimen Komentar Video Youtube Dengan Leksikon Textblob Menggunakan Algoritma Svm Berdasarkan Rasio Data Uji*. *9*, 1894–1902.