

ARTICLE

Perbandingan Algoritma Machine Learning dan Deep Learning Dalam Analisis Sentimen Komentar Video YouTube Berjudul Hidup Tanpa Sosial Media

Comparison of Machine Learning and Deep Learning Algorithms in Sentiment Analysis of YouTube Video Comments Titled Life Without Social Media

Eveline Valencia Anthony* dan Simon Prananta Barus

Teknik Informatika, Fakultas Sains, Komputer, dan Matematika, Universitas Matana, Tangerang, Indonesia

*Penulis Korespondensi: eveline.anthony@student.matanauniversity.ac.id

(Disubmit 22-12-24; Diterima 07-04-25; Dipublikasikan online pada 20-06-25)

Abstrak

Penelitian ini membandingkan kinerja model deep learning (LSTM) dan machine learning (Logistic Regression, Naive Bayes, KNN) dalam analisis sentimen pada komentar YouTube, dengan fokus pada persepsi pengguna Indonesia terkait pengaruh konten terhadap kesehatan mental. Dataset terdiri dari komentar yang dilabeli sebagai positif, netral, dan negatif. Hasil menunjukkan bahwa Logistic Regression mencapai akurasi validasi tertinggi (84,0%), mengungguli LSTM yang kesulitan melakukan generalisasi (akurasi validasi 70,94%). Mayoritas komentar menunjukkan sentimen netral, dengan sebagian kecil positif dan negatif. Temuan ini menyoroti trade-off antara kompleksitas model dan stabilitas performa, di mana Logistic Regression menawarkan pendekatan yang lebih cepat dan stabil. Penelitian ini juga menekankan pentingnya kesadaran akan dampak negatif media sosial terhadap kesehatan mental, merekomendasikan pengembangan fitur yang mendukung penggunaan sehat dan pengaturan waktu konsumsi konten. Menunjukkan performa yang lebih baik dengan akurasi validasi sebesar 84,0%. Meskipun LSTM memiliki potensi untuk menangani data yang lebih kompleks, Logistic Regression menawarkan pendekatan yang lebih cepat dan stabil untuk tugas analisis sentimen. Perbandingan ini menyoroti trade-off antara model deep learning dan machine learning dalam aplikasi dunia nyata, dengan penekanan pada pemilihan model berdasarkan karakteristik dataset dan keterbatasan komputasi.

Kata kunci: Analisis Sentimen; Kesehatan Mental; Logistic Regression; LSTM; Machine Learning; Deep Learning

Abstract

This study compares deep learning (LSTM) and machine learning (Logistic Regression, Naive Bayes, KNN) models for sentiment analysis of YouTube comments, focusing on Indonesian users' perceptions of content's impact on mental health. The dataset includes comments labeled as positive, neutral, or negative. Results show Logistic Regression achieved the highest validation accuracy (84.0%), outperforming LSTM, which struggled with generalization (70.94% accuracy). Most comments were neutral, with fewer positive and negative ones. The findings highlight the trade-off between model complexity and performance stability, with Logistic Regression offering a faster, more stable approach. Although LSTM can handle complex data, Logistic Regression is more efficient for sentiment analysis. The study also stresses the need for awareness of social media's negative effects on mental health, recommending features that promote healthy usage and regulated content consumption. This comparison emphasizes selecting models based on dataset characteristics and computational limits.

KeyWords: Sentiment Analysis; Mental Health; Logistic Regression; LSTM; Machine Learning; Deep Learning

This is an Open Access article - copyright on authors, distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY SA) (<http://creativecommons.org/licenses/by-sa/4.0/>)

How to Cite: E. V. Anthony *et al.*, "Perbandingan Algoritma Machine Learning dan Deep Learning Dalam Analisis Sentimen Komentar Video YouTube Berjudul Hidup Tanpa Sosial Media", *JIKO (JURNAL INFORMATIKA DAN KOMPUTER)*, Volume: 9, No.2, Pages 370–380, Juni 2025, doi: 10.26798/jiko.v9i2.1619.

1. Pendahuluan

YouTube, sebagai platform berbagi video terbesar di dunia, telah mengalami pertumbuhan yang luar biasa pesat dalam beberapa tahun terakhir. Dengan jumlah pengunjung bulanan sebesar 1,7 milyar, YouTube adalah situs web berbagi video terbesar di dunia[1]. Di Indonesia, platform ini telah berkembang menjadi bagian tak terpisahkan dari kehidupan sehari-hari jutaan orang, terutama individu berusia 13-35 tahun. Melalui YouTube, pengguna muda dapat menikmati beragam konten mulai dari video musik, tutorial kecantikan, edukasi, vlog kehidupan sehari-hari influencer, hingga konten yang mempromosikan citra tubuh ideal. Fitur-fitur interaktif seperti komentar, like, dislike, dan subscription memungkinkan pengguna untuk berinteraksi langsung dengan kreator maupun sesama penonton, sehingga YouTube telah berkembang menjadi ruang publik tempat individu dapat saling berpendapat, berbagi pengalaman, dan membangun komunitas. Namun, di balik berbagai interaksi positif yang tercipta, platform ini juga menghadapi tantangan dalam mengatasi komentar negatif, ujaran kebencian, dan perilaku toksik yang kerap dialami oleh kreator.

Peningkatan penggunaan media sosial, termasuk YouTube, di kalangan individu muda sering kali dikaitkan dengan perasaan tidak mampu, perbandingan sosial yang tidak realistis, dan ekspektasi yang berlebihan terhadap pencapaian diri. Fenomena ini berpotensi memperburuk kondisi kesehatan mental individu, yang dapat memicu depresi, kecemasan, hingga gangguan harga diri. Dalam konteks ini, kesehatan mental didefinisikan sebagai keadaan kesejahteraan di mana individu menyadari kemampuannya sendiri, dapat mengatasi tekanan kehidupan sehari-hari, bekerja secara produktif, dan mampu berkontribusi pada masyarakat. Perbandingan sosial merujuk pada proses membandingkan diri sendiri dengan orang lain, terutama melalui media sosial, yang dapat memicu perasaan rendah diri dan tidak puas.

Salah satu pendekatan untuk memahami tanggapan masyarakat di platform YouTube adalah melalui analisis sentimen. Ulasan pendapat orang tentang suatu topik dapat menjadi sumber data yang berguna untuk mengetahui minat dan pendapat masyarakat tentang topik tersebut[2]. Analisis sentimen adalah studi komputasi untuk mengenali serta mengekspresikan opini, sentimen, evaluasi, sikap, emosi, subjektivitas, penilaian, atau pandangan yang terdapat dalam suatu teks. Ini juga dikenal sebagai penambangan opini[3]. Analisis sentimen berfungsi untuk mengolah dan menganalisis dokumen atau teks guna mengkategorikan opini menjadi positif, netral, atau negatif secara otomatis. Meskipun analisis sentimen lebih umum diterapkan pada platform media sosial seperti Twitter dan Instagram, YouTube menawarkan potensi yang lebih besar untuk pemanfaatan metode ini. Dengan pertumbuhan pengguna yang signifikan dan jumlah subscriber yang besar, YouTube menjadi media yang semakin relevan untuk analisis sentimen dibandingkan dengan platform lainnya. Selain itu, masih terbatasnya penggunaan bahasa Indonesia dalam analisis sentimen di YouTube memperbesar peluang untuk memanfaatkan platform ini secara lebih efektif.

Meskipun beberapa studi sebelumnya telah mengungkapkan adanya korelasi antara penggunaan media sosial yang berlebihan dengan peningkatan risiko gangguan kesehatan mental, penelitian mengenai persepsi pengguna YouTube sendiri terkait dampak platform ini terhadap kesehatan mental, terutama dalam konteks perbandingan sosial yang dipicu oleh konten-konten seperti video tentang gaya hidup influencer dan tutorial kecantikan, masih sangat terbatas. Penelitian ini bertujuan untuk mengisi kekosongan tersebut dengan menganalisis sentimen pengguna YouTube di Indonesia terhadap pengaruh konten YouTube, khususnya yang berkaitan dengan kesehatan mental.

Penting untuk dicatat bahwa penelitian ini memiliki beberapa keterbatasan. Analisis sentimen terhadap data teks sering kali terbatas dalam menangkap nuansa kompleks dari pengalaman emosional pengguna. Misalnya, sentimen positif atau negatif yang direkam dalam data teks mungkin tidak selalu menunjukkan intensitas emosional atau faktor-faktor psikologis di balik komentar pengguna. Selain itu, sampel penelitian yang digunakan mungkin tidak sepenuhnya representatif dari populasi pengguna YouTube secara keseluruhan.

Penelitian ini membandingkan pendekatan Deep Learning dan Machine Learning dalam klasifikasi sentimen pada komentar video YouTube. Pada sisi Deep Learning, digunakan arsitektur *Long Short-Term Memory* (LSTM) dengan pendekatan *Bidirectional LSTM*, yang dipilih karena kemampuannya menangani data sekuensial seperti dengan mempertimbangkan konteks dari kedua arah (maju dan mundur), sehingga mampu memahami dependensi jangka panjang dan nuansa dalam teks. Sementara itu, pada sisi Machine

Learning, digunakan empat algoritma klasifikasi, yaitu *Naïve Bayes* (NB), *K-Nearest Neighbors* (KNN), *Logistic Regression* (LR). NB dipilih karena kesederhanaan dan kecepatannya dalam menangani teks, meski mengasumsikan independensi antar fitur. KNN digunakan karena fleksibilitasnya dalam menangani data non-linear dengan membandingkan titik data baru berdasarkan tetangga terdekat, meski cenderung lambat pada data besar. Sementara itu, LR dipilih karena interpretabilitasnya yang baik, ketahanan terhadap overfitting, dan kesesuaiannya untuk klasifikasi biner. Dengan membandingkan kedua pendekatan ini, penelitian bertujuan untuk mengevaluasi keakuratan, kecepatan, dan ketepatan masing-masing metode. Bidirectional LSTM memiliki kemampuan untuk mempertimbangkan informasi dari kedua arah dalam urutan data, yang membuatnya potensial dalam menangani konteks kompleks. Sementara itu, algoritma Machine Learning seperti Naive Bayes (NB) dan Logistic Regression (LR) umumnya lebih cepat dan efisien dalam menangani dataset yang lebih sederhana karena kompleksitas modelnya yang lebih rendah.

Dengan memahami persepsi masyarakat, penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan strategi pencegahan dan penanganan masalah kesehatan mental yang terkait dengan penggunaan media sosial, terutama di kalangan remaja. Hasil penelitian ini juga dapat menjadi dasar bagi pengembangan program literasi digital yang lebih efektif untuk membantu remaja menggunakan media sosial secara sehat dan produktif.

2. Metode

2.1 Pengumpulan Data

Tahap pengumpulan data dilakukan dengan mengakses Google YouTube Data API. Dalam Google Developers Console sendiri, API & Services Tools tersedia. Untuk mengaktifkan API & Services ini, kita harus mengarahkannya ke API Library dengan menggunakan Search YouTube Data API v3 dan membuat credentials untuk mendapatkan API Key[4]. API dapat mengumpulkan data komentar secara langsung dari server YouTube. Melalui API ini, dataset berhasil di kumpulkan dari komentar pada video YouTube berjudul "Hidup Tanpa Social Media Dan Bahagia" sejumlah lebih dari 1000 data[5]. Data komentar ini kemudian dijadikan sebagai bahan utama dalam analisis sentimen.

Dalam penelitian ini, video "Hidup Tanpa Social Media Dan Bahagia" dipilih berdasarkan kriteria tertentu, yaitu:

1. Memiliki jumlah penonton yang signifikan.
2. Relevan dengan topik kesehatan mental.
3. Memiliki jumlah komentar yang cukup untuk analisis sentimen.

Data dikumpulkan dengan memanggil endpoint API menggunakan parameter yang sesuai, seperti `videoId`, `part=snippet`, dan `maxResults`. Hanya komentar dalam bahasa Indonesia yang diambil dengan filter berbasis deteksi bahasa. Sebanyak lebih dari 1000 komentar berhasil dikumpulkan. Setiap entri dalam dataset mencakup:

1. Comment ID: Identitas unik komentar.
2. User ID: Identitas pengguna (disamarkan untuk privasi).
3. Comment Text: Isi komentar asli.

2.2 Preprocessing Data

Pre-processing adalah proses mengubah data teks yang tidak terstruktur menjadi data yang dapat dianalisis [6]. Proses membersihkan set data mentah dari simbol, emoji, tanda baca, link url, tanda baca, dan angka dikenal sebagai *cleaning text*[7]. Pada tahap *cleaning text*, dilakukan penghapusan tag HTML seperti `
`, dan karakter khusus seperti simbol, emoji. Selain itu, kata-kata yang tidak memberikan makna seperti imbuhan dan kata penghubung juga dihilangkan. Proses case folding diterapkan untuk mengubah semua kata menjadi huruf kecil, misalnya kata "Jakarta" menjadi "jakarta". Tokenizing dilakukan menggunakan library NLTK. Tujuan dari proses tokenizing ini adalah untuk membagi teks yang berupa kalimat menjadi

token atau membaginya menjadi kata perkata[8]. Selanjutnya menghapus kata-kata yang tidak penting dan tidak memengaruhi proses perolehan informasi atau biasa dikenal sebagai stopword removal[9]. Singkatnya, stopword adalah kata yang seringkali dianggap tidak bermakna[10]. Kata-kata stop seperti "dan", "atau", "yang" dihilangkan menggunakan daftar stopwords bahasa Indonesia dan english. Terakhir, proses stemming menggunakan library dari Sastrawi. Tujuan stemming adalah untuk mengembalikan kata dasar dengan menghilangkan awalan, akhiran, dan sisipan[11]. Setelah melalui proses preprocessing, data teks yang telah bersih dan terstruktur kemudian dirangkai kembali menjadi kalimat untuk analisis sentimen. Hasil preprocessing data komentar kemudian disimpan dalam format CSV dengan dua kolom utama, yaitu "comment" yang berisi teks komentar asli dan "preprocessed_comment" yang berisi teks komentar setelah melalui proses pembersihan dan normalisasi.

2.3 Pelabelan

Labeling adalah tahap di mana hasil dari tahap sebelumnya dipertimbangkan terhadap polarity ulasan, yang menghasilkan 3 kategori—label positif, negatif dan netral[12]. Pelabelan dilakukan dengan menggunakan data lexicon yang berisi kata-kata positif dan negatif. Setiap komentar dianalisis berdasarkan skor yang diberikan pada kata-kata yang ada dalam lexicon tersebut. Jika total skor positif lebih tinggi, komentar diberikan label positif, jika lebih rendah diberikan label negatif, dan jika skor seimbang, komentar diberi label netral.

2.4 Pemodelan

Pada tahap pemodelan, tiga algoritma klasifikasi *machine learning*, yaitu *Naive Bayes*, KNN, dan Regresi Logistik, diterapkan pada data yang telah diproses. *Naive Bayes* adalah algoritma yang digunakan untuk mengidentifikasi dan menghitung kemungkinan anggota kelas[13]. Algoritma *Naive Bayes* akan digunakan untuk mengklasifikasikan sentimen berdasarkan probabilitas kemunculan kata pada kelas tertentu, KNN akan mengklasifikasikan sentimen berdasarkan kemiripan dengan data terdekat, sedangkan *Logistic regression* adalah teknik statistik yang bergantung pada hubungan antara satu atau lebih variabel[14]. Selain itu, model deep learning berbasis *Long Short-Term Memory* (LSTM) dengan pendekatan *Bidirectional LSTM* juga diterapkan. Pemilihan model ini didasarkan pada karakteristik data yang bersifat teks, dimana pendekatan *deep learning* umumnya lebih unggul dalam menangani data sekuensial yang kompleks. Setelah pemilihan model, tahap selanjutnya adalah pelatihan model menggunakan data yang telah dilabeli. Model *machine learning* dilatih menggunakan data yang telah dibersihkan dan diproses, sementara model *deep learning* dilatih menggunakan Arsitektur LSTM dan *Bidirectional LSTM* untuk memanfaatkan informasi konteks sekuensial dalam komentar. Proses pelatihan dilakukan dengan membagi data menjadi data pelatihan dan data uji untuk menghindari *overfitting*.

2.5 Pembagian Data Latih dan Data Uji

Untuk mengevaluasi kinerja model secara komprehensif, pembagian data dilakukan dengan beberapa variasi perbandingan, yaitu 60:40, 70:30, dan 80:20. Pada setiap perbandingan, data yang telah dilabeli dibagi secara acak menjadi dua bagian. Bagian yang lebih besar digunakan sebagai data latih untuk melatih model, sedangkan bagian yang lebih kecil digunakan sebagai data uji untuk mengukur akurasi model.

2.6 Evaluasi Model

Tahap evaluasi bertujuan untuk mengukur kemampuan model dalam menggeneralisasi pada data baru. Model yang telah dilatih akan diaplikasikan pada data uji, dan hasilnya akan dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score.

Akurasi adalah keakuratan kumpulan data yang diprediksi. Akurasinya adalah keakuratan prediksi kasus positif dengan prediksi benar. Recall mengacu pada keakuratan kasus positif yang teridentifikasi dengan benar. Skor F1 adalah ukuran kinerja yang mengambil presisi p dan mengembalikan r untuk menghitung skor[15].

Selain itu, Nilai *Confusion Matrix* digunakan untuk melakukan evaluasi kinerja. Ini dilakukan untuk mengetahui keakuratan metode dan kesesuaian sistem yang dibuat[16]. Hasil evaluasi ini akan menjadi dasar untuk memilih model terbaik dan melakukan perbaikan jika diperlukan.

2.7 Analisis Sentimen

Setelah evaluasi, hasil dari kedua pendekatan tersebut dianalisis untuk melihat kekuatan dan kelemahan masing-masing metode dalam konteks analisis sentimen komentar YouTube. Berdasarkan analisis ini, akan ditarik kesimpulan mengenai kelebihan dan kekurangan model machine learning dibandingkan dengan model deep learning dalam menangani data komentar yang tidak terstruktur.

Analisis sentimen, atau biasa disebut penelitian opini, adalah studi tentang opini, sentimen, ulasan, penilaian, sikap, dan perasaan masyarakat terhadap suatu hal tertentu, seperti produk, layanan, organisasi, individu, isu, peristiwa, atau topik. Ini adalah bidang. Analisis sentimen terdiri dari tiga level analisis dan aspek: level dokumen, level kalimat, dan level entitas[17].

Setelah model terbaik terpilih, tahap selanjutnya adalah melakukan analisis sentimen terhadap seluruh dataset. Tujuannya adalah untuk mengidentifikasi sentimen dominan terkait pengaruh konten YouTube terhadap kesehatan mental. Hasil prediksi sentimen akan divisualisasikan dalam bentuk grafik distribusi sentimen. Analisis ini diharapkan dapat memberikan gambaran yang jelas mengenai persepsi publik terhadap pengaruh konten YouTube terhadap kesehatan mental dan mendukung pengambilan keputusan yang lebih baik.

3. Hasil dan Pembahasan

3.1 Machine Learning

Pembersihan data dilakukan dalam tahap pra-pemrosesan teks. Dalam data tersebut, terdapat komentar yang ditinggalkan oleh penonton di video YouTube. Komentar ini terbagi menjadi tiga kategori: positif, netral, dan negatif. Data tersebut terdiri dari 1.000 komentar yang membahas pengaruh media sosial.

Pada analisis sentimen ini dilakukan tiga model eksperimen. Eksperimen dilakukan menggunakan tiga pembagian data yaitu 80:20, 70:30, 60:40. Hasil eksperimen yang diukur berdasarkan akurasi dan F1-score dari masing-masing model dengan pembagian data tersebut dapat dilihat pada Tabel 1 berikut:

Tabel 1. Evaluasi Model Naive Bayes 60:40

	Naive Bayes 60:40			
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	20
Netral	0.85	0.97	0.91	330
Positif	0.61	0.23	0.33	48
Accuracy			0.83	398
Micro Avg	0.49	0.40	0.41	398
Weighted Avg	0.78	0.83	0.79	398

Model *Naive Bayes* dengan skema pembagian data 60:40 pada Tabel 1 menunjukkan bahwa kelas netral memiliki performa terbaik dengan F1-score 0.91. Model ini mencapai akurasi keseluruhan sebesar 83%.

Tabel 2. Evaluasi Model KNN 60:40

	KNN 60:40			
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	20
Netral	0.84	1.00	0.91	330
Positif	1.00	0.08	0.15	48
Accuracy			0.84	398
Micro Avg	0.61	0.36	0.36	398
Weighted Avg	0.82	0.84	0.77	398

Model KNN dengan pembagian data 60:40 pada Tabel 2 memiliki akurasi sebesar 84%, sedikit lebih ting-

gi dari Naive Bayes. Namun, kelas positif memiliki nilai recall yang sangat rendah (0.08), menunjukkan bahwa model kesulitan dalam mengenali kelas positif. Kelas netral tetap menjadi kelas yang paling akurat diklasifikasikan.

Tabel 3. Evaluasi Model *Logistic Regression* 60:40

Logistic Regression 60:40				
	Precision	Recall	F1-Score	Support
Negatif	0.67	0.10	0.17	20
Netral	0.85	0.97	0.91	330
Positif	0.61	0.23	0.33	48
Accuracy			0.83	398
Micro Avg	0.49	0.40	0.41	398
Weighted Avg	0.78	0.83	0.79	398

Logistic Regression pada skema 60:40 seperti yang ditunjukkan pada Tabel 3 memiliki akurasi tertinggi di antara ketiga model, yaitu 86.4%. Meskipun kelas negatif masih kurang terprediksi dengan baik, kelas netral memiliki F1-score tertinggi (0.91). Secara keseluruhan, model ini menunjukkan keseimbangan yang lebih baik dibandingkan dengan *Naive Bayes* dan KNN.

Tabel 4. Evaluasi Model *Naive Bayes* 60:40

Naive Bayes 70:30				
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	13
Netral	0.85	0.97	0.91	247
Positif	0.62	0.21	0.31	38
Accuracy			0.83	298
Micro Avg	0.49	0.39	0.41	298
Weighted Avg	0.78	0.83	0.79	298

Pada pembagian data 70:30, model *Naive Bayes* pada Tabel 4 tetap mengalami kesulitan dalam mengklasifikasikan kelas negatif dengan F1-score 0.00. Akurasi keseluruhan model adalah 83%, yang sejalan dengan hasil dari skema 60:40.

Tabel 5. Evaluasi Model KNN 70:30

KNN 70:30				
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	13
Netral	0.84	1.00	0.91	247
Positif	1.00	0.08	0.15	38
Accuracy			0.84	298
Micro Avg	0.61	0.36	0.35	298
Weighted Avg	0.82	0.84	0.77	298

Model KNN dengan pembagian 70:30 pada Tabel 5 memiliki akurasi sebesar 84%. Sama seperti sebelumnya, model ini hampir sempurna dalam mengklasifikasikan kelas netral, tetapi memiliki kelemahan dalam mengenali kelas positif dengan recall hanya 0.08.

Logistic Regression kembali unggul dalam eksperimen ini dengan akurasi 85.9%, seperti yang tampak pada Tabel 6. Model ini menunjukkan peningkatan dalam klasifikasi kelas negatif dibandingkan dengan skema 60:40, meskipun kelas positif masih memiliki recall yang cukup rendah (0.29).

Pada pembagian data 80:20, performa model *Naive Bayes* sedikit menurun dengan akurasi 82%, sesuai

Tabel 6. Evaluasi Model *Logistic Regression* 70:30

Logistic Regression 70:30				
	Precision	Recall	F1-Score	Support
Negatif	0.67	0.15	0.25	13
Netral	0.86	0.98	0.92	247
Positif	0.79	0.29	0.42	38
Accuracy			0.86	298
Micro Avg	0.77	0.48	0.53	298
Weighted Avg	0.85	0.86	0.83	298

Tabel 7. Evaluasi Model Naive Bayes 80:20

Naive Bayes 80:20				
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	8
Netral	0.84	0.98	0.90	162
Positif	0.56	0.17	0.26	29
Accuracy			0.82	199
Micro Avg	0.46	0.38	0.39	199
Weighted Avg	0.76	0.82	0.77	199

Tabel 7. Kelas negatif masih tidak dapat diklasifikasikan dengan baik, sementara kelas netral tetap memiliki performa terbaik.

Tabel 8. Evaluasi Model KNN 80:20

KNN 80:20				
	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	8
Netral	0.83	1.00	0.91	162
Positif	1.00	0.10	0.19	29
Accuracy			0.83	199
Micro Avg	0.61	0.37	0.36	199
Weighted Avg	0.82	0.83	0.76	199

Model KNN pada skema 80:20 mempertahankan akurasi sebesar 83%, sesuai Tabel 8. Seperti pada skema sebelumnya, model ini sangat baik dalam mengklasifikasikan kelas netral tetapi buruk dalam mengenali kelas positif.

Tabel 9. Evaluasi Model *Logistic Regression* 80:20

Logistic Regression 80:20				
	Precision	Recall	F1-Score	Support
Negatif	0.60	0.38	0.46	8
Netral	0.86	0.98	0.91	162
Positif	0.80	0.28	0.41	29
Accuracy			0.85	199
Micro Avg	0.75	0.54	0.60	199
Weighted Avg	0.84	0.85	0.82	199

Logistic Regression kembali menunjukkan performa terbaik dengan akurasi 85%, sesuai Tabel 9. Meskipun ada sedikit penurunan dibandingkan skema 70:30, model ini tetap unggul dalam menangani berbagai kelas

dibandingkan model lainnya.

Berdasarkan tiga eksperimen diatas maka eksperimen ketiga menghasilkan nilai *accuracy* paling besar yaitu 86.4%, nilai akurasi ini didapatkan di model *Logistic Regression* di pembagian data 60:40.

Tabel 10. Hasil Evaluasi Model

No	Model	Accuracy	F1-Score
0	Naive Bayes (0.6:0.4)	0.831658	0.821827
1	KNN (0.6:0.4)	0.839196	0.821827
2	Logistic Regression (0.6:0.4)	0.864322	0.821827
3	Naive Bayes (0.7:0.3)	0.832215	0.821827
4	KNN (0.7:0.3)	0.838926	0.821827
5	Logistic Regression (0.7:0.3)	0.859060	0.821827
6	Naive Bayes (0.8:0.2)	0.819095	0.821827
7	KNN (0.8:0.2)	0.829146	0.821827
8	Logistic Regression (0.8:0.2)	0.849246	0.821827

Model terbaik: Logistic Regression dengan F1-Score: 0.821827, Accuracy: 0.864322, 0.859060, 0.849246

Berdasarkan hasil eksperimen dari ketiga model pada Tabel 10, model *Logistic Regression* terbukti sebagai model terbaik dengan akurasi tertinggi (86.4%) pada skema pembagian data 60:40. F1-score keseluruhan tetap konsisten di angka 0.82, menunjukkan keseimbangan yang baik antara *precision* dan *recall*. Dengan demikian, *Logistic Regression* adalah model yang paling andal untuk analisis sentimen dalam penelitian ini.

3.2 Deep Learning

Model yang digunakan dalam eksperimen Deep Learning adalah model LSTM (Long Short-Term Memory) yang dilatih dengan menggunakan data komentar YouTube. Model ini terdiri dari beberapa lapisan seperti embedding, bidirectional LSTM, dense, dan dropout, dengan total parameter sebanyak 2,167,995. Berikut adalah ringkasan hasil dari pelatihan model:

3.2.1 Akurasi Pelatihan

Model ini mencapai akurasi sebesar 84.82% pada epoch terakhir dari 10 epoch pelatihan. Meskipun akurasi pelatihan terus meningkat, model tidak menunjukkan peningkatan signifikan pada akurasi validasi setelah epoch ke-7.

3.2.2 Performa pada Data Validasi

Akurasi terbaik yang dicapai pada data validasi adalah 70.94% pada epoch ke-10, yang menunjukkan bahwa model mampu mengenali pola dalam data sentimen dengan performa yang cukup baik meskipun ada ketidakseimbangan kelas.

3.2.3 Loss dan Accuracy

Nilai loss pada pelatihan menunjukkan penurunan yang signifikan sepanjang epoch, menandakan bahwa model mengalami pembelajaran yang baik.

3.3 Perbandingan Model Deep Learning dan Machine Learning

Dalam perbandingan antara model LSTM (Deep Learning) dan *Logistic Regression* (Machine Learning), ditemukan beberapa poin penting yang dapat disoroti:

3.3.1 Akurasi

Meskipun model LSTM mencapai akurasi pelatihan yang tinggi (84.82%), performa pada data validasi tidak meningkat secara signifikan setelah epoch ke-7. Akurasi terbaik pada data validasi adalah 70.94%. Di sisi lain, *Logistic Regression* dengan akurasi 86.4% pada data validasi menunjukkan kinerja yang sangat baik, meskipun tidak sebaik LSTM pada data pelatihan.

3.3.2 Keterbatasan pada Machine Learning

Model machine learning, seperti Logistic Regression, lebih cepat untuk dilatih dan tidak memerlukan parameter yang sebanyak LSTM, namun performanya dapat terbatas pada pemrosesan data yang lebih kompleks. Hal ini terlihat pada kinerja Logistic Regression yang lebih stabil meskipun tidak memiliki arsitektur yang dalam seperti LSTM.

3.3.3 Kelebihan Deep Learning

LSTM, dengan kemampuannya untuk menangkap dependensi jangka panjang dalam urutan data, menunjukkan keunggulan dalam pengolahan teks yang lebih panjang atau lebih kompleks. Meskipun performa validasi LSTM tidak lebih baik dari Logistic Regression, potensi untuk memperbaiki akurasi bisa lebih besar dengan lebih banyak data dan tuning model yang lebih lanjut.

3.3.4 Waktu Pelatihan

Salah satu kelemahan dari model Deep Learning adalah waktu pelatihannya yang lebih lama. LSTM memerlukan waktu lebih lama dibandingkan model Logistic Regression, meskipun memiliki kemampuan untuk menangani data dengan lebih banyak fitur dan kompleksitas.

4. Simpulan

Analisis sentimen terhadap komentar pengguna YouTube di Indonesia terkait pengaruh konten terhadap kesehatan mental menunjukkan bahwa mayoritas komentar memiliki sentimen netral. Hal ini mengindikasikan bahwa pengguna YouTube memiliki pandangan yang beragam mengenai dampak platform ini terhadap kesejahteraan mental. Meskipun demikian, hasil analisis juga menunjukkan adanya proporsi kecil komentar dengan sentimen negatif, yang mengindikasikan adanya kekhawatiran terkait potensi dampak negatif YouTube terhadap kesehatan mental, terutama terkait perbandingan sosial dan tekanan untuk mencapai citra ideal.

Pengujian tiga algoritma klasifikasi (Naive Bayes, KNN, dan Regresi Logistik) menunjukkan bahwa Regresi Logistik memberikan akurasi tertinggi dalam mengklasifikasikan sentimen. Hal ini mengindikasikan bahwa Regresi Logistik merupakan model machine learning yang paling cocok untuk data yang digunakan dalam penelitian ini.

Berdasarkan hasil percobaan, meskipun model LSTM menunjukkan akurasi pelatihan yang lebih tinggi, Logistic Regression memberikan akurasi validasi yang lebih stabil dan lebih tinggi pada data yang diuji. Oleh karena itu, model Logistic Regression dapat dianggap lebih efektif dalam aplikasi analisis sentimen pada komentar YouTube dalam konteks ini. Namun, untuk dataset yang lebih besar atau kompleks, model Deep Learning seperti LSTM memiliki potensi untuk memberikan hasil yang lebih baik, terutama jika dioptimalkan lebih lanjut.

Perbandingan ini juga menegaskan bahwa pemilihan model sangat bergantung pada sifat dan ukuran data yang digunakan serta tujuan aplikasi yang diinginkan. Dalam kasus ini, meskipun model machine learning dapat memberikan hasil yang cukup baik dengan lebih cepat, model deep learning mungkin lebih efektif dalam menangani data yang lebih kompleks atau besar.

Hasil penelitian ini memberikan wawasan mengenai performa berbagai model dalam analisis sentimen terhadap komentar YouTube. Berdasarkan evaluasi yang telah dilakukan, model Logistic Regression menunjukkan hasil terbaik dengan akurasi tertinggi dibandingkan model lainnya.

Namun, penelitian ini masih memiliki keterbatasan dalam menghubungkan hasil analisis sentimen dengan dampak nyata terhadap kesehatan mental pengguna YouTube. Oleh karena itu, pernyataan mengenai dampak negatif penggunaan media sosial terhadap kesehatan mental belum dapat disimpulkan secara langsung dari hasil penelitian ini.

Meskipun demikian, temuan ini dapat menjadi langkah awal bagi penelitian lebih lanjut yang mengeksplorasi hubungan antara sentimen dalam komentar YouTube dan kesejahteraan mental pengguna. Selain itu, hasil ini dapat memberikan wawasan bagi pengembang algoritma moderasi konten serta kreator YouTube dalam memahami reaksi audiens mereka secara lebih mendalam.

Penelitian selanjutnya dapat memperluas cakupan analisis dengan menganalisis konten video YouTube secara lebih mendalam atau dengan melibatkan sampel yang lebih besar dan lebih representatif. Selain itu, penelitian juga dapat dilakukan untuk mengidentifikasi faktor-faktor yang mempengaruhi sentimen pengguna, seperti jenis konten yang ditonton, durasi penggunaan YouTube, dan karakteristik demografis pengguna.

Pustaka

- [1] G. Sanjaya and K. M. Lhaksana, "Analisis sentimen komentar youtube tentang terpilihnya menteri kabinet indonesia maju menggunakan lexicon based," *e-proceeding of engineering*, vol. 7, no. 3, p. 9698, 2020.
- [2] R. Fatmasari, V. M. Ayu, H. Anto, W. Gata, and L. D. Yulianto, "Analisis sentimen dalam pengkategorian komentar youtube terhadap layanan akademik dan non-akademik universitas terbuka untuk prediksi kepuasan," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 2, pp. 395–404, Sep 2022.
- [3] H. A. R. Harpizon *et al.*, "Analisis sentimen komentar di youtube tentang ceramah ustadz abdul somad menggunakan algoritma naïve bayes," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.
- [4] A. A. Ningtyas, A. Solichin, and R. Pradana, "Analisis sentimen komentar youtube tentang prediksi resesi ekonomi tahun 2023 menggunakan algoritme naïve bayes," 2023.
- [5] F. Irwandi, "Hidup tanpa social media dan bahagia," <https://youtu.be/JE4lT3e3xFU?si=iPLjZYRqSjAv45NE>, accessed: Oct. 29, 2024.
- [6] S. F. Huwaida, R. Kusumawati, and B. Isnaini, "Analisis sentimen komentar youtube terhadap pemin-dahan ibu kota negara menggunakan metode naïve bayes," *Jambura Journal of Informatics*, vol. 6, no. 1, pp. 26–39, Apr 2024.
- [7] M. Hudha, E. Supriyati, and T. Listyorini, "Analisis sentimen pengguna youtube terhadap tayangan #matanajwamenantiterawan dengan metode naïve bayes classifier," *Jurnal Informatika dan Komputer (JIKO)*, vol. 5, no. 1, pp. 2614–8897, 2022.
- [8] A. Wahid and G. Saputri, "Analisis sentimen komentar youtube tentang relawan patwal ambulancce menggunakan algoritma naïve bayes dan decision tree," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 319, Dec 2022.
- [9] D. N. Larasakti, A. Aziz, and D. Aditya, "Analisis sentimen komentar video youtube dengan metode k-nearest neighbor," *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 5, pp. 132–142, 2023.
- [10] E. F. Saputra and M. R. Pribadi, "Analisis sentimen komentar pada kanal youtube the lazy monday menggunakan algoritma naïve bayes," in *MDP Student Conference (MSC)*, 2023, pp. 17–23.
- [11] C. A. Misrun, E. Haerani, M. Fikry, and E. Budianita, "Analisis sentimen komentar youtube terha-dap anies baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 1, pp. 207–215, Apr 2023.
- [12] D. Mualfah, Ramadhoni, R. Gunawan, and D. M. Suratno, "Analisis sentimen komentar youtube tvone tentang ustadz abdul somad dideportasi dari singapura menggunakan algoritma svm," *Jurnal Fasil-kom*, vol. 13, no. 1, pp. 72–80, 2023.
- [13] C. Ayunda *et al.*, "Analisis komparasi algoritma machine learning untuk sentiment analysis (studi kasus: Komentar youtube 'kekerasan seksual')," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 7, no. 2, pp. 80–84, 2022.
- [14] B. Ramadhani and R. R. Suryono, "Komparasi algoritma naïve bayes dan logistic regression untuk analisis sentimen metaverse," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 2, p. 714, Apr 2024.

- [15] A. N. Badri, N. Noviandi, F. Anastya, and M. Roland, "Sentiment analisis untuk identifikasi kepuasan masyarakat terhadap kenaikan bbm menggunakan algoritma naïve bayes," *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, p. 287, Sep 2023.
- [16] M. P. Munthe, A. Siswo, R. Ansori, and R. R. Septiawan, "Analisis sentimen komentar pada saluran youtube food vlogger berbahasa indonesia menggunakan algoritma naïve bayes," 2021.
- [17] R. Sanusi, F. D. Astuti, D. Indra, and Y. Buryadi, "Analisis sentimen pada twitter terhadap program kartu pra kerja dengan recurrent neural network," 2021.