

ANALISIS SENTIMEN PUBLIK PERNYATAAN 'LALAI' TRANS 7 TAYANGAN PESANTREN DENGAN SUPPORT VECTOR MACHINE

Abdulrahman Tuasamu^{1*}, Rapanca Cahya Gumilang², Mohammad Akmal Fachrian³, Daiva Rakha Krisnandi⁴, Azizah Wardah Indryani⁵

¹²³⁴⁵Informatika, Universitas Bina Sarana Informatika; Jl. Margonda No.8, Pondok Cina, Kecamatan Beji, Kota Depok, Jawa Barat 16424

Keywords:

Analisis Sentimen, Youtube, SVM, Klasifikasi.

Correspondent Email:

atuasamu@gmail.com

Abstrak. Perkembangan pesat teknologi informasi dan komunikasi telah memperluas serta meningkatkan intensitas interaksi masyarakat melalui berbagai platform media sosial. Salah satu platform yang paling populer dan banyak dimanfaatkan adalah YouTube, yang tidak hanya berperan sebagai media hiburan, tetapi juga menjadi ruang publik digital bagi masyarakat untuk menyampaikan opini, gagasan, serta beragam respons kritis terhadap isu-isu sosial, fenomena yang berkembang, dan berbagai tayangan media. Komentar yang ditinggalkan oleh pengguna YouTube dapat mencerminkan persepsi, sikap, dan sentimen publik terhadap suatu konten tertentu. Namun, analisis komentar secara manual dalam jumlah besar membutuhkan waktu dan tenaga yang tidak sedikit, sehingga dinilai kurang efisien dan kurang objektif. Dengan demikian, penelitian ini diarahkan pada pengkajian sentimen komentar pengguna YouTube terhadap tayangan “Pesantren” yang ditayangkan di Trans7 melalui penerapan algoritma Support Vector Machine (SVM). Metode penelitian yang digunakan meliputi tahap pengumpulan data komentar dari YouTube, proses pra-pengolahan data seperti pembersihan teks dan normalisasi, serta tahap klasifikasi sentimen menjadi kategori positif dan negatif menggunakan algoritma SVM. Untuk mengukur kinerja model, dilakukan pengujian menggunakan metrik evaluasi berupa akurasi, presisi, dan recall melalui metode validasi silang k-fold. Hasil penelitian menunjukkan bahwa algoritma Support Vector Machine mampu melakukan klasifikasi sentimen komentar secara optimal dengan tingkat akurasi mencapai 88,3%. Capaian ini menegaskan bahwa algoritma SVM memiliki kinerja yang efektif dan andal dalam mengolah data teks serta menganalisis sentimen pada komentar YouTube berbahasa Indonesia secara akurat dan konsisten.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. *The rapid advancement of information and communication technology has significantly increased the scale and intensity of public interaction on various social media platforms. Among these, YouTube stands out as one of the most popular, functioning not only as a source of entertainment but also as a vibrant digital public sphere where individuals share opinions, ideas, and critical responses to social issues, emerging phenomena, and media content. User comments on YouTube reflect public perceptions, attitudes, and sentiments toward specific content. However, manual analysis of large volumes of comments is time-consuming and subjective. This study addresses this challenge by analyzing the sentiment of YouTube user comments on the "Pesantren" show broadcast by Trans7 using the Support Vector Machine (SVM) algorithm. The research methodology comprises data collection from YouTube, data pre-processing—including text cleaning and normalization—and sentiment classification into positive and*

negative categories using SVM. Model performance was evaluated using metrics such as accuracy, precision, and recall, employing k-fold cross-validation. The results demonstrate that the SVM algorithm effectively classified comment sentiment, achieving an accuracy of 88.3%. These findings underscore the effectiveness and reliability of SVM in processing text data and accurately analyzing sentiment in Indonesian YouTube comments.

1. PENDAHULUAN

Televisi digemari sebagai media hiburan sekaligus sumber informasi karena kemampuannya menyampaikan konten secara audio-visual. Melalui saluran ini, berbagai jenis acara seperti film, sinetron, musik, reality show, dan program lainnya dapat dipresentasikan dengan menghadirkan artis-artis yang terkenal di kalangan publik[1]. Perkembangan teknologi informasi dan komunikasi terus meningkat secara signifikan, mencakup komputer, internet, dan berbagai sektor lainnya, yang manfaatnya kini semakin dirasakan oleh banyak orang. Kemajuan tersebut memberikan pengaruh besar pada seluruh lini kehidupan manusia, mulai dari pekerjaan, sektor bisnis, hingga aktivitas harian. Teknologi informasi dan komunikasi terus berkembang untuk memenuhi kebutuhan manusia, sehingga tidak mengherankan jika hampir semua kelompok usia—mulai dari anak-anak, remaja, hingga orang dewasa—memanfaatkannya dalam kehidupan sehari-hari. Kehadiran teknologi ini juga memudahkan akses terhadap berbagai informasi serta memungkinkan pertukaran data antar media secara lebih efisien dan fleksibel[2], serta memungkinkan perluasan jaringan sosial antarindividu dengan lebih luas dan dinamis[3].

Tayangan Trans7 yang mengangkat isu konflik pesantren menyajikan pembahasan yang mendalam mengenai dinamika yang berlangsung di lingkungan pondok pesantren. Program tersebut biasanya mengulas berbagai dimensi permasalahan, mulai dari akar konflik, aktor aktor yang terlibat, hingga konsekuensi sosial yang dirasakan masyarakat. Melalui format pemberitaan, sesi wawancara, serta cuplikan dokumentasi visual, Trans7 berusaha menyampaikan gambaran yang menyeluruh tentang penyebab serta perkembangan situasi yang terjadi. Selain memberikan informasi yang aktual, tayangan ini juga bertujuan mendorong publik untuk memahami konteks peristiwa serta mempertimbangkan langkah langkah

penyelesaian yang konstruktif terhadap isu yang muncul di lingkungan pesantren. Kemajuan teknologi informasi dan komunikasi yang sangat cepat telah membawa perubahan signifikan dalam cara masyarakat berinteraksi, menyampaikan pendapat, dan memperoleh informasi. Di era digital sekarang, media sosial seperti YouTube menjadi wadah baru bagi masyarakat untuk menyalurkan pandangan terhadap berbagai isu publik. Sebagai platform berbasis video dengan jangkauan yang luas serta kemampuan penyebaran informasi yang cepat, YouTube berperan besar dalam membentuk opini dan persepsi masyarakat. Melalui fitur komentar, unggahan, maupun diskusi daring, masyarakat dapat dengan mudah mengekspresikan pendapat, perasaan, serta penilaian mereka terhadap kebijakan pemerintah dan fenomena sosial. Sebagai salah satu platform streaming video paling populer di dunia, YouTube menyediakan beragam konten berupa video hiburan maupun edukasi[4]. Hal ini membuktikan bahwa perkembangan teknologi tidak hanya mempermudah akses terhadap informasi, tetapi juga menciptakan ruang partisipatif yang berpengaruh terhadap pola komunikasi dan keterlibatan publik dalam diskusi sosial.

2. TEORI DAN PENELITIAN RELAVAN

2.1. Analisis Sentimen

Analisis sentimen adalah metode yang dimanfaatkan untuk menggali informasi dari beragam platform yang ada di internet[5]. Analisis sentimen adalah salah satu bidang dalam *text mining* dan *natural language processing* (NLP) yang bertujuan untuk mengklasifikasikan opini publik yang terkandung dalam teks ke dalam kategori positif, negatif, maupun netral[6].

Metode ini berfokus pada identifikasi dan pemahaman emosi yang terkandung dalam teks ulasan, dengan tujuan untuk memprediksi serta

menganalisis suasana publik, mood, dan ekspresi perasaan masyarakat secara otomatis terhadap suatu isu. Analisis sentimen bertujuan mengelompokkan sentimen dalam dokumen berbasis teks baik pada komentar, ulasan produk, maupun suatu topik ke dalam kategori positif atau negative[7]. Media sosial Youtube dipilih karena saat ini menjadi salah satu platform yang paling populer di kalangan pengguna internet. Analisis sentimen terus berkembang dan menjadi topik yang banyak dibahas dalam berbagai publikasi ilmiah. Salah satu penelitian bertujuan untuk mengusulkan sistem yang mampu mendeteksi sentimen publik berdasarkan komentar pengguna di YouTube, dengan menggunakan metode Support Vector Machine (SVM).

2.2. Machine Learning

Machine learning, atau pembelajaran mesin, merupakan cabang ilmu komputer yang memungkinkan sistem bekerja tanpa harus diprogram secara eksplisit. Para peneliti terus mengeksplorasi berbagai pendekatan untuk mencapai kemajuan kecerdasan buatan yang mendekati kemampuan manusia. Bidang ini mempelajari cara mengolah serta memanfaatkan data secara efektif, sehingga sering disingkat sebagai ML. Teknologi ini diperlukan untuk menerapkan metode yang cepat dan efisien dalam mengidentifikasi serta menyelesaikan berbagai masalah baru[8]. Sebagai pendekatan alternatif, machine learning (ML) hadir sebagai solusi adaptif yang mampu mengekstraksi pola dari data kinerja masa lalu, memodelkan keterkaitan kompleks antarparameter, serta menghasilkan estimasi performa yang lebih presisi. Oleh karena itu, ML berperan penting dalam mendukung pengambilan keputusan berbasis data, perencanaan kapasitas, mitigasi risiko, serta optimasi sumber daya secara real-time[9].

Selama dua puluh tahun terakhir, Machine Learning telah berkembang menjadi salah satu pilar utama dalam teknologi informasi dan kini menjadi elemen penting meskipun sering kali tidak terlihat dalam kehidupan sehari-hari. Seiring meningkatnya ketersediaan data, muncul keyakinan kuat bahwa pemanfaatan teknik analisis data yang cerdas akan membuat data semakin berperan sebagai sumber utama dalam pengembangan teknologi modern yang terus berkembang[10].

2.3. Algorism

Algoritma menjadi komponen yang sangat krusial dalam proses pengembangan program dengan bahasa pemrograman apa pun, karena kualitas dan hasil akhir program sangat ditentukan oleh algoritma yang dirancang[11].

Algoritma merujuk pada prosedur atau tahapan yang disusun dengan sistematis dan teratur guna menyelesaikan suatu persoalan atau memperoleh hasil yang diinginkan. Gagasan mengenai algoritma pertama kali dikenalkan oleh matematikawan Persia, Abu Abdullah Muhammad Ibn Musa Al-Khawarizmi. Istilah “algoritma” berasal dari penyesuaian nama beliau, yang kemudian digunakan untuk menamai konsep penting tersebut dan tetap dikenal hingga sekarang[12].

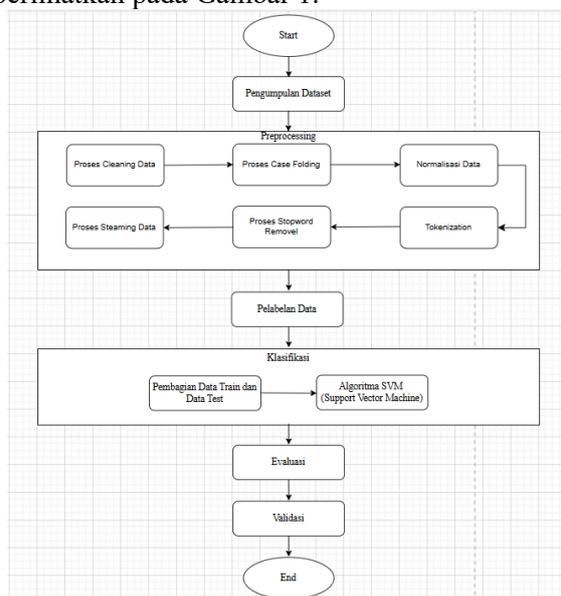
2.4. Algoritma Support Vector Machine

Support Vector Machine (SVM) adalah teknik klasifikasi yang diperkenalkan oleh Vapnik pada tahun 1998. Metode ini pada dasarnya berfungsi dengan membentuk sebuah garis atau batas pemisah antara dua kelas, yang ditentukan berdasarkan data-data yang berada paling dekat dengan batas pemisahan tersebut[13]. Algoritma ini berfungsi sebagai teknik klasifikasi dengan mempelajari pola yang terdapat pada data pelatihan untuk menentukan kelas suatu data. Dikenalkan oleh Vladimir Vapnik, Support Vector Machine beroperasi dengan mengelompokkan data ke dalam kelas yang berbeda menggunakan sebuah *hyperplane* sebagai batas pemisah, dengan tujuan menghasilkan jarak pemisahan (margin) yang maksimal antara dua kelas, yaitu positif dan negative [14]. Meskipun demikian, penggunaan kernel linear memiliki keterbatasan ketika diaplikasikan pada data yang tidak dapat dipisahkan secara linear atau memiliki pola yang lebih kompleks. Pendekatan non-linear diterapkan untuk menyelesaikan permasalahan tersebut. Umumnya, Support Vector Machine yang menggunakan metode non-linear memanfaatkan beragam jenis kernel guna meningkatkan kemampuan pemodelan dalam memisahkan data yang tidak dapat dipisahkan secara linear pada ruang aslinya[15].

3. METODE PENELITIAN

Untuk mencapai tujuan penelitian, alur rancangan penelitian analisis sentimen menggunakan Support Vector Machine

dilakukan secara terstruktur seperti yang6. diperlihatkan pada Gambar 1.



Gambar 1. Alur Perancangan

1. Tahap pengumpulan dataset atau Crawling data yaitu tahap dilakukannya penarikan data dari komentar di Youtube.
2. Pada tahap *Preprocessing* ini meliputi proses *cleaning data*, *case folding*, *normalisasi*, *tokenization*, *stopword removal*, dan *stemming data*. Serangkaian proses *Preprocessing* ini bertujuan untuk mengubah dokumen mentah menjadi representasi dokumen yang siap diproses untuk di analisis.
3. Selanjutnya tahap pelabelan data yang bertujuan untuk memberikan label atau kategori pada data komentar sebelum dimasukkan ke proses pelatihan dan klasifikasi.
4. Lalu ada tahap Classification yaitu dilakukannya pengklasifikasian data dengan menggunakan algoritma SVM (Support Vector Machine).
5. Pada tahap evaluasi, kinerja klasifikasi pada algoritma SVM (Support Vector Machine) diukur melalui perhitungan tingkat akurasi, presisi, dan *recall*. Perhitungan metrik ini didasarkan pada *confusion matrix* yang dihasilkan dari klasifikasi dokumen.

Dan yang terakhir tahap validasi, data ini dibagi menjadi 10 subset yang setara, dimana pengujian model diulang sebanyak sepuluh kali (*10-fold cross-validation*).

3.1. Pengumpulan Data

Data dalam penelitian ini adalah respons publik berupa teks yang dikumpulkan dari platform youtube terkait isu pernyataan 'lalai' Trans7 tayangan pesantren. Pengumpulan data dilakukan secara otomatis melalui teknik data crawling (atau scraping) untuk mendapatkan korpus teks yang akan dianalisis sentimennya.

3.2. Preprocessing Text

Tahapan preprocessing merupakan proses awal dalam pengolahan data teks yang bertujuan menyiapkan data agar siap digunakan sebelum masuk ke tahap klasifikasi. Tahapan ini bertujuan mengubah komentar mentah menjadi data yang lebih bersih, terorganisasi, dan sesuai dengan kebutuhan analisis. Dalam penelitian ini, tahapan *preprocessing* yang digunakan meliputi:

3.2.1. Cleaning Data

Tahap ini dilakukan untuk memurnikan data teks dengan menghilangkan karakter yang tidak relevan, seperti tanda baca, angka, emoji, simbol, maupun URL. Hasil dari proses ini adalah mendapatkan data teks murni yang siap diproses lebih lanjut. Contoh kalimat : 'KPAI harus bertanggung jawab, jangan tidur.' menjadi KPAI harus bertanggung jawab jangan tidur.

3.2.2. Case Folding

Seluruh huruf pada teks diseragamkan menjadi lowercase pada tahap ini, sehingga perbedaan kapitalisasi tidak menimbulkan makna yang berbeda. Sebagai contoh, kata "Boikot" akan diubah menjadi "boikot".

3.2.3. Normalisasi

Pada tahap ini normalisasi dilakukan untuk menyeragamkan bentuk kata tidak baku atau singkatan menjadi bentuk baku. Proses ini sangat penting karena banyak komentar di media sosial yang menggunakan bahasa tidak formal. Contoh kalimat : kebenaran yang dibungkam/memang kenyataanya begitu.

3.2.4. Tokenization

Tokenization adalah proses memecah kalimat menjadi potongan-potongan kata yang disebut token. Misalnya, komentar "Saya hormat ke siapapun saya dukung" akan

diubah menjadi ["saya", "hormat", "ke", "siapa pun", "saya", "dukung"]. Hasil tokenization ini digunakan untuk proses analisis sentimen pada tahap klasifikasi.

3.2.5. Stopword Removal

Proses ini bertujuan untuk menghilangkan kata-kata umum yang tidak memberikan kontribusi signifikan terhadap makna kalimat, seperti "yang", "dan", "di", maupun "ke". Dengan menghapus kata-kata tersebut, hanya kata-kata bermakna yang akan dipertahankan. Contoh kalimat: Kebenaran yang dibungkam memang kenyataannya begitu menjadi [kebenaran, bungkam, memang, kenyataannya]

3.2.6. Stemming Data

Stemming bertujuan mengembalikan kata ke bentuk dasarnya. Sebagai contoh, kata "Kebenaran" diubah menjadi "benar". Proses ini membantu mengurangi variasi kata yang memiliki arti serupa sehingga mempermudah tahapan klasifikasi. Contoh kalimat: *Benar bungkam memang kenyataannya aneh kriminalis.*

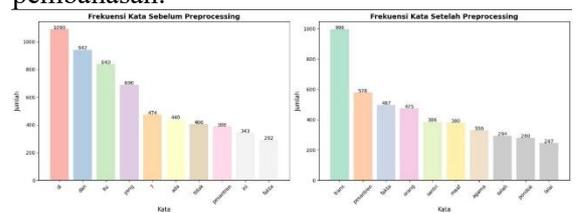
4.2. Text Processing

Setelah data berhasil dikumpulkan, data tersebut memasuki tahap pra-pemrosesan. Ini adalah fase kritis yang berfungsi sebagai fondasi sebelum data dapat digunakan dalam pemodelan machine learning. Data mentah seringkali tidak konsisten dan 'kotor' (noisy), sehingga serangkaian teknik pemrosesan teks seperti Cleaning, Case Folding, Normalisasi, Tokenisasi, Stopword Removal, dan Stemming mutlak diperlukan, seperti yang ditunjukkan pada Gambar 3. Penerapan langkah-langkah ini sangat penting untuk mereduksi kompleksitas data dan mengekstrak fitur yang relevan, sehingga memungkinkan algoritma machine learning bekerja secara lebih efisien dan menghasilkan wawasan (insight) atau prediksi yang lebih baik dan akurat.

comment	cleaning	case_folding	normalisasi	tokenize	stopword_removal	stemming_data
0 Kebenaran yg di bungkam/mnng kenyataannya bgtu...	Kebenaran yg di bungkam/mnng kenyataannya bgtu...	kebenaran yg di bungkam/mnng kenyataannya bgtu...	kebenaran yang di bungkam mng kenyataannya beg...	{kebenaran, yang, di, bungkam, mng, kenyataannya, ...}	{kebenaran, mng, kenyataannya, aneh, ...}	benar bungkam mng kenyataannya aneh kriminalis...
1 Bokok	Bokok	bokok	bokok	{bokok}	{bokok}	bokok
2 KPAI harus ikut bertanggung jawab jgn tiduu...	KPAI harus ikut bertanggung jawab jgn tiduu...	kpi harus ikut bertanggung jawab jgn tiduu...	kpi harus ikut bertanggung jawab jgn tiduu...	{kpi, harus, ikut, bertanggung, jawab, jgn, tiduu, ...}	{kpi, bertanggung, iduaur, ruang, ac, ...}	kpi tanggung iduaur ruang ac
3 Saya hormat ke siapapun/Saya dukung trans 7...	Saya hormat ke siapapun/Saya dukung trans 7...	saya hormat ke siapapun/saya dukung trans 7...	saya hormat ke siapapun saya dukung trans semo...	{saya, hormat, ke, siapapun, saya, dukung, trans, ...}	{hormat, dukung, trans, semoga, indonesia, tor, ...}	hormat dukung trans semoga indonesia tor
4 Saya gk kau berfikir kotor tapi tetap aja ya...	Saya gk kau berfikir kotor tapi tetap aja ya...	saya gk kau berfikir kotor tapi tetap aja ya...	saya gk kau berfikir kotor tapi tetap aja ya...	{saya, gk, kau, berfikir, kotor, tapi, tetap, aja, ya, ...}	{gk, kau, berfikir, kotor, demo, kpi, ...}	gk kau berfikir kotor demo kpi

Gambar 3. Hasil Pre-processing

Setelah di *preprocessing* akan menghasilkan data teks yang lebih bersih dan relevan, di mana kata-kata hubung umum (*stopwords*) seperti 'di' dan 'dan' berhasil dieliminasi, sehingga memunculkan dominasi kata kunci substantif seperti 'trans', 'pesantren', 'fakta', dan 'maaf' yang secara akurat menggambarkan inti topik pembahasan.



Gambar 4. Frekuensi Kata

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Proses pengumpulan data dilaksanakan pada 23 Oktober 2025 dengan menerapkan teknik scraping komentar dari video YouTube berjudul "Trans7 Akui Lalai dalam Tayangannya Soal Pondok Pesantren" yang diunggah pada kanal CNN. Proses ekstraksi data ini difasilitasi oleh platform Google Collab dan berhasil menghimpun total 6.686 data. Setiap data yang diperoleh memiliki empat atribut utama, yaitu author (penulis komentar), comment (isi komentar), likes (jumlah suka), dan published_at (waktu publikasi). Keseluruhan data mentah tersebut kemudian disimpan dan disusun ke dalam satu dataset tunggal bernama Data Pesantren dengan format file .csv, yang strukturnya diilustrasikan lebih lanjut pada Gambar 2.

	author	comment	likes	published_at
0	@Achie16	Kebenaran yg di bungkam/mnng kenyataannya bgtu...	0	2025-10-22T02:28:20Z
1	@Arisalwi	Bokok	0	2025-10-21T22:46:08Z
2	@rahmaheny4786	KPAI harus ikut bertanggung jawab jgn tiduu...	0	2025-10-21T22:00:13Z
3	@viralvideo-z1d	Saya hormat ke siapapun/Saya dukung trans 7, ...	0	2025-10-21T19:43:04Z
4	@SandiJoe-z8i	Saya gk kau berfikir kotor tapi tetap aja ya...	0	2025-10-21T19:37:40Z

Gambar 2. Hasil Scraping Comentar Youtube

4.3. Pelabelan Data

Pada penelitian ini, proses pelabelan data sentimen memanfaatkan pendekatan berbasis Leksikon (Lexicon-based). Metode ini bekerja dengan cara mengkuantifikasi polaritas sebuah kalimat berdasarkan kamus kata (leksikon) yang sudah memiliki bobot sentimen. Setiap kalimat dianalisis dan dihitung skor sentimen totalnya. Selanjutnya, skor tersebut digunakan untuk klasifikasi biner dengan aturan sebagai berikut: kalimat yang memperoleh skor total 0

atau negatif (misalnya, < 0) dikategorikan sebagai 'sentimen negatif'. Sebaliknya, kalimat yang memperoleh skor 1 atau nilai positif lainnya (misalnya, > 1) dikategorikan sebagai 'sentimen positif'. Gambar 3 di bawah ini menyajikan ilustrasi hasil akhir dari proses pelabelan tersebut, di mana setiap kalimat telah berhasil dipetakan ke dalam label sentimen positif atau negatif."

	steming_data	Score	Sentiment
0	benar bungkam mmng kenyataanya aneh kriminalis...	1	Positif
1	boikot	0	Negatif
2	kpai tanggung tiduur ruang ac	-2	Negatif
3	hormat dukung trasn moga indonesia buka mata r...	1	Positif
4	gknkau berfikiran kotor demo kiyai roti demo k...	0	Negatif

Gambar 5. Hasil Pelabelan Data

4.4. Klasifikasi

Pada tahap ini, dataset terlebih dahulu dibagi menjadi data untuk pelatihan dan data untuk pengujian, lalu dilanjutkan dengan mengevaluasi kinerja model Support Vector Machine. Visualisasi pada confusion matrix menunjukkan bahwa model dapat melakukan klasifikasi yang benar terhadap 539 data berlabel Negatif (True Negative) dan 115 data berlabel Positif (True Positive). Tetap ada kesalahan prediksi, yakni 35 kasus yang salah ditandai sebagai Positif padahal Negatif (False Positive) dan 52 kasus yang salah ditandai Negatif padahal seharusnya Positif (False Negative).

SVM Confusion Matrix		
	Actual Negatif	Actual Positif
Predicted Negatif	539	52
Predicted Positif	35	115

Gambar 6. Confusion Matrix

4.5. Evaluasi

Setelah melakukan klasifikasi, pada tahap ini kita akan meninjau hasil evaluasi model

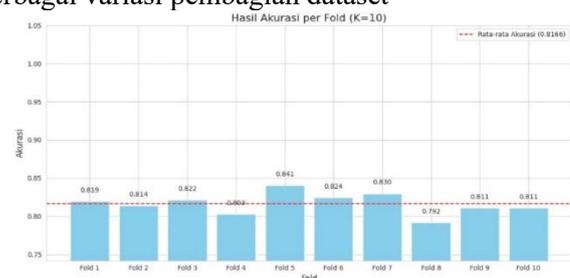
yang terangkum dalam tabel statistik. Model berhasil mencatatkan akurasi global sebesar 88,3% dengan nilai *weighted average F1-score* sebesar 0,880. Kinerja model terlihat lebih unggul pada kelas Negatif (F1-score 0,925) dibandingkan dengan kelas Positif (F1-score 0,726). Hal ini menunjukkan bahwa meskipun model memiliki kemampuan generalisasi yang baik, ia cenderung lebih sensitif dalam mendeteksi sentimen atau label Negatif.

	precision	recall	f1-score	support
Negatif	0.912	0.939	0.925	574.000
Positif	0.767	0.689	0.726	167.000
accuracy	0.883	0.883	0.883	0.883
macro avg	0.839	0.814	0.825	741.000
weighted avg	0.879	0.883	0.880	741.000

Gambar 7. Hasil Evaluasi

4.6. Validasi

Setelah melakukan evaluasi, pada tahap ini kita akan memvalidasi ketangguhan model dengan teknik *10-Fold Cross-Validation*. Dari grafik hasil akurasi per *fold*, terlihat bahwa model mampu mempertahankan performa yang konsisten dengan rata-rata akurasi akhir mencapai 0,8166. Meskipun terdapat fluktuasi nilai—mulai dari 0,792 pada *Fold 8* hingga 0,841 pada *Fold 5*—perbedaan tersebut relatif kecil, yang menegaskan bahwa algoritma mampu mengklasifikasikan data dengan baik di berbagai variasi pembagian dataset



Gambar 8. Hasil Validasi

5. KESIMPULAN

Berdasarkan pengujian model Support Vector Machine dengan teknik 10-Fold Cross Validation, diperoleh rata-rata akurasi sebesar 81,66%, dengan rentang akurasi pada tiap fold berada antara 79,2% hingga 84,1%. Temuan ini mengindikasikan bahwa model memiliki performa yang relatif stabil pada berbagai subset data serta tidak menunjukkan tanda-tanda overfitting yang berarti. Selain itu,

hasil *classification report* mencatat akurasi keseluruhan sebesar 88,3%, yang mengisyaratkan bahwa model Support Vector Machine bekerja cukup baik dalam melakukan klasifikasi sentimen.

Pada kelas Negatif, model menunjukkan kinerja yang sangat baik dengan nilai precision 0,912, recall 0,939, dan f1-score 0,925, yang menandakan bahwa model mampu mengidentifikasi data negatif dengan tingkat ketepatan serta konsistensi yang tinggi. Sebaliknya, pada kelas Positif, performa model terlihat lebih rendah dengan precision 0,767, recall 0,689, dan f1-score 0,726, sehingga menunjukkan bahwa model masih menghadapi tantangan dalam mengenali sentimen positif secara optimal. Kondisi ini diperkirakan disebabkan oleh ketidakseimbangan jumlah data antara kelas negatif (574 data) dan kelas positif (167 data).

Selanjutnya, hasil confusion matrix memperkuat temuan tersebut dengan menunjukkan bahwa model berhasil mengklasifikasikan 539 data negatif dengan benar, sementara 35 data negatif salah diprediksi sebagai positif. Untuk kelas positif, model mengklasifikasikan 115 data dengan benar, tetapi masih terdapat 52 data positif yang salah teridentifikasi sebagai negatif. Temuan ini mengindikasikan bahwa SVM memiliki kecenderungan untuk lebih baik dalam mengenali pola pada kelas mayoritas (negatif) dibandingkan kelas minoritas (positif). Secara umum, performa model dapat dikatakan baik dan stabil, dengan kemampuan generalisasi yang memadai terhadap data uji. Meskipun demikian, terdapat peluang peningkatan terutama pada aspek sensitivitas terhadap kelas positif agar model lebih seimbang dalam mengidentifikasi kedua jenis sentimen.

DAFTAR PUSTAKA

- [1] F. Bouchra, I. M. A. D. Suarjaya, and N. K. D. Rusjyanthi, "Analisis Sentimen Masyarakat terhadap Tayangan Televisi Nasional menggunakan Metode Deep Learning," *Jurnal Buana Informatika*, vol. 15, pp. 89–99, 2024, [Online]. Available: <https://ojs.uajy.ac.id/index.php/jbi/article/view/9118>
- [2] I. A. Huda, "Perkembangan Teknologi Informasi Dan Komunikasi (Tik) Terhadap Kualitas Pembelajaran Di Sekolah Dasar," *Jurnal Pendidikan dan Konseling (JPDK)*, vol. 2, no. 1, pp. 121–125, 2020, doi: 10.31004/jpdk.v1i2.622.
- [3] Yuhandra, "Dampak Positif Dan Negatif Penggunaan Gadget Dan Media Sosial," *Jurnal Pengabdian Masyarakat*, vol. 04, pp. 78–84, 2021, [Online]. Available: <https://journal.uniku.ac.id/index.php/empowerment/article/view/4028>
- [4] J. Pratama, "Klasifikasi Sentimen Terhadap Ustadz Abdul Somad Di Media Sosial Youtube Menggunakan Metode Support Vector Machine," 2021, [Online]. Available: [https://repository.uin-suska.ac.id/45664/1/Laporan Lengkap Kecuali Bab 4 & 5.pdf](https://repository.uin-suska.ac.id/45664/1/Laporan%20Lengkap%20Kecuali%20Bab%204%20&%205.pdf)
- [5] K. S. Q. Valentino, I. Ade, and D. P. Hindriyanto, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization (Online Transportation Sentiment Analysis Using Support Vector Machine Based on Particle Swarm Optimization)," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, pp. 162–170, 2020, [Online]. Available: https://www.researchgate.net/publication/342101786_Analisis_Sentimen_Transportasi_Online_Menggunakan_Support_Vector_Machine_Berbasis_Particle_Swarm_Optimization
- [6] A. D. Oktavianingsih, A. F. Fauzi, J. Immanuel, and O. D. Simatupang, "Analisis Sentimen Komentar YouTube terhadap Pengangkatan Menkeu Purbaya Menggunakan Algoritma SVM dan Naïve Bayes," vol. 5, no. 1, pp. 3702–3713, 2025.
- [7] M. P. Munthe, A. S. R. Ansori, and ..., "Analisis Sentimen Komentar Pada Saluran Youtube Food Vlogger Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes," *eProceedings of Engineering*, vol. 8, no. 6, pp. 11909–11916, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversiti.ac.id/index.php/engineering/article/view/16897>
- [8] N. HASYDNA and R. K. DINATA, "样本量估算-Machine Learning.Pdf," 2020. [Online]. Available: [https://repository.unimal.ac.id/6707/1/Machine Learning.pdf](https://repository.unimal.ac.id/6707/1/Machine%20Learning.pdf)
- [9] J. Komputer, V. No, N. Hal, and M. Alfidyah, "Optimasi Algoritma Machine Learning untuk Prediksi Kinerja Sistem Komputer," vol. 1, no. 1, pp. 1–7, 2025.
- [10] A. Aditya Permana *et al.*, *Machine learning*, vol. 45, no. 13. 2023. [Online]. Available: <https://repository.bsi.ac.id/repo/files/364253/download/Buku---Machine-Learning.pdf>

- [11] B. University, “Algoritma Pemrograman: Pengertian, Cara Kerja, dan Fungsinya,” 2024, [Online]. Available: <https://binus.ac.id/malang/2024/02/algoritma-pemrograman-pengertian-cara-kerja-dan-fungsinya/>
- [12] W. E. Susanto and A. Syukron, “Logika & Algoritma untuk Pemula,” *Graha Ilmu*, p. 76, 2020, [Online]. Available: https://repository.bsi.ac.id/index.php/unduh/item/285644/Bidang-A_Buku-Ajar-Logika-Algoritma-a.n-Wahyu-Eko-Susanto.pdf
- [13] H. Harnelia, “Analisis Sentimen Review Skincare Skintific Dengan Algoritma Support Vector Machine (Svm),” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4095.
- [14] L. Muhammad, G. Jaya, R. A. Saputra, and S. H. Idrus, “USING-SUPPORT-VECTOR-MACHINE-TO-IDENTIFY-LAND-COVER-CHANGE-DURING-COVID19-PANDEMIC-IN-KOMODO-NATIONAL-PARK-INDONESIA_2022_Russian-Geographical-Society.pdf,” no. July 2021, pp. 70–79, 2022.
- [15] M. Ma’rufudin and A. Yudhistira, “Analisis Sentimen Petani Milenial Pada Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM),” *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 3, pp. 845–857, 2025, doi: 10.52436/1.jpti.717.