

Analisis Sentimen Komentar Twitter Tentang Kemajuan Kecerdasan Buatan (Ai) Menggunakan Metode Support Vector Maching (SVM)

Chalista Thesa Amanda^{1*}, Destiarini², dan Anggraeni Agustin Muris³

^{1,2,3} Informatika, Fakultas Teknik dan Komputer, Universitas Baturaja, Indonesia

*E-mail: kalistatesaamanda@gmail.com

ABSTRAK

Perkembangan kecerdasan buatan (Artificial Intelligence/AI) yang pesat, khususnya melalui kemunculan ChatGPT, telah memunculkan berbagai tanggapan di media sosial seperti Twitter. Penelitian ini bertujuan untuk menganalisis sentimen pengguna Twitter terhadap kemajuan AI ChatGPT dengan menggunakan metode Support Vector Machine (SVM). Data dikumpulkan dari sejumlah cuitan yang mengandung kata kunci "ChatGPT" dan "AI" selama periode tertentu. Setelah melalui tahap preprocessing seperti pembersihan teks, tokenisasi, stopword removal, dan stemming, data diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Hasil analisis menunjukkan bahwa metode SVM mampu mengklasifikasikan sentimen dengan tingkat akurasi yang tinggi dan stabil. Secara umum, mayoritas pengguna Twitter memberikan tanggapan positif terhadap kemajuan AI ChatGPT, terutama terkait kemampuannya dalam membantu pekerjaan dan pendidikan. Namun, terdapat pula sentimen negatif yang menyoroti kekhawatiran terhadap etika, privasi, dan potensi penggantian peran manusia. Penelitian ini diharapkan dapat memberikan gambaran umum tentang persepsi masyarakat terhadap perkembangan AI, sekaligus menjadi dasar bagi pengembangan kebijakan dan pemanfaatan teknologi secara lebih bijak.

Kata kunci: Analisis sentimen, Twitter, ChatGPT, kecerdasan buatan, Support Vector Machine

ABSTRACT

The rapid development of Artificial Intelligence (AI), particularly through the emergence of ChatGPT, has sparked various responses on social media platforms such as Twitter. This study aims to analyze Twitter users' sentiments toward the advancement of AI ChatGPT using the Support Vector Machine (SVM) method. Data were collected from tweets containing the keywords "ChatGPT" and "AI" during a specific period. After going through preprocessing stages such as text cleaning, tokenization, stopword removal, and stemming, the data were classified into three sentiment categories: positive, negative, and neutral. The analysis results show that the SVM method can classify sentiments with high and stable accuracy. In general, most Twitter users expressed positive sentiments toward the advancement of AI ChatGPT, particularly regarding its ability to assist in work and education. However, negative sentiments were also found, highlighting concerns about ethics, privacy, and the potential replacement of human roles. This study is expected to provide an overview of public perceptions regarding AI development and serve as a basis for developing policies and utilizing technology more wisely.

Keywords: Sentiment analysis, Twitter, ChatGPT, artificial intelligence, Support Vector Machine

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi pada era digital saat ini berlangsung sangat pesat, khususnya dalam bidang kecerdasan buatan (Artificial Intelligence/AI). Kemajuan teknologi ini telah membawa perubahan besar dalam berbagai aspek kehidupan manusia, mulai dari bidang pendidikan, ekonomi, kesehatan, hingga komunikasi [1]. Salah satu inovasi AI yang paling banyak menarik perhatian masyarakat adalah ChatGPT, yaitu sistem berbasis Natural Language Processing

(NLP) yang mampu menghasilkan teks secara otomatis dan menyerupai percakapan manusia [2]. ChatGPT dapat membantu pengguna dalam menulis, mencari informasi, menjawab pertanyaan, menerjemahkan bahasa, hingga membantu menyelesaikan berbagai tugas akademik maupun profesional [3]. Kehadiran teknologi ini memberikan kemudahan dan efisiensi dalam berbagai aktivitas, sehingga banyak pihak yang menyambutnya dengan antusias dan menjadikannya sebagai bagian dari kehidupan sehari-hari.

Namun demikian, di balik berbagai manfaat yang ditawarkan, perkembangan AI juga menimbulkan beragam tanggapan dan perdebatan di tengah masyarakat. Sebagian orang memandang AI sebagai inovasi yang dapat meningkatkan produktivitas dan kualitas hidup, sementara sebagian lainnya mengkhawatirkan dampak negatifnya, seperti potensi penggantian tenaga kerja manusia, penyalahgunaan teknologi untuk penyebaran informasi palsu, permasalahan etika, serta isu privasi dan keamanan data [4]. Perbedaan persepsi tersebut menunjukkan bahwa kemajuan teknologi tidak hanya berdampak secara teknis, tetapi juga secara sosial dan psikologis [5]. Oleh karena itu, penting untuk mengetahui bagaimana sebenarnya pandangan masyarakat terhadap perkembangan AI, khususnya ChatGPT, agar dapat menjadi bahan evaluasi, pertimbangan kebijakan, serta dasar pengembangan teknologi yang lebih bertanggung jawab di masa depan.

Media sosial, terutama Twitter (X), menjadi salah satu platform yang paling aktif digunakan masyarakat untuk menyampaikan pendapat dan berdiskusi mengenai berbagai isu, termasuk perkembangan kecerdasan buatan. Twitter menyediakan ruang publik digital yang memungkinkan pengguna untuk mengekspresikan opini secara bebas, cepat, dan terbuka. Komentar-komentar yang muncul pada sebuah postingan sering kali mencerminkan respons spontan masyarakat terhadap suatu topik tertentu [6]. Data komentar tersebut memiliki nilai yang sangat penting karena dapat menggambarkan kecenderungan opini publik secara luas dan real-time. Akan tetapi, jumlah komentar yang sangat banyak membuat proses analisis secara manual menjadi tidak efektif dan membutuhkan waktu yang lama serta berpotensi menimbulkan subjektivitas [7]. Oleh sebab itu, diperlukan pendekatan berbasis teknologi untuk mengolah dan menganalisis data teks secara otomatis dan objektif.

Salah satu pendekatan yang dapat digunakan adalah analisis sentimen berbasis machine learning. Analisis sentimen merupakan teknik untuk mengidentifikasi dan mengklasifikasikan opini atau emosi dalam suatu teks ke dalam kategori tertentu, seperti positif, negatif, dan netral. Dalam penelitian ini, metode yang digunakan untuk melakukan klasifikasi sentimen adalah Support Vector Machine (SVM) [8]. SVM dipilih karena dikenal memiliki kemampuan yang baik dalam menangani data berdimensi tinggi seperti teks, mampu bekerja secara efektif pada jumlah data yang relatif terbatas, serta menghasilkan tingkat akurasi yang stabil melalui pencarian hyperplane terbaik sebagai pemisah antar kelas [9]. Dengan penerapan metode SVM, komentar-komentar Twitter dapat diklasifikasikan secara sistematis sehingga menghasilkan informasi yang lebih terstruktur mengenai kecenderungan opini masyarakat terhadap kemajuan AI.

Berdasarkan latar belakang tersebut, penulisan ini bertujuan untuk menganalisis sentimen komentar Twitter terkait kemajuan kecerdasan buatan (AI ChatGPT), mengklasifikasikan komentar ke dalam kategori positif, negatif, dan netral menggunakan metode Support Vector Machine (SVM), serta mengukur tingkat akurasi model dalam melakukan klasifikasi sentimen. Selain itu, penelitian ini juga bertujuan untuk mengetahui distribusi sentimen yang paling dominan dan memahami kecenderungan respons masyarakat

terhadap perkembangan AI di ruang digital. Ruang lingkup tulisan ini dibatasi pada penggunaan data komentar Twitter yang membahas tentang kemajuan AI ChatGPT dalam periode waktu tertentu dengan jumlah data yang telah ditentukan. Proses analisis meliputi tahapan pengambilan data (crawling), pembersihan dan pemrosesan teks (preprocessing), pelabelan sentimen, penerapan metode SVM untuk klasifikasi, serta evaluasi model menggunakan metrik akurasi, confusion matrix, precision, recall, dan f1-score. Dengan batasan tersebut, diharapkan tulisan ini dapat memberikan hasil analisis yang terarah, sistematis, objektif, dan dapat dipertanggungjawabkan secara akademis, serta menjadi referensi bagi penelitian selanjutnya dalam bidang analisis sentimen dan pemanfaatan kecerdasan buatan di era digital.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis machine learning untuk mengkaji persepsi masyarakat terhadap kemajuan kecerdasan buatan (AI ChatGPT) melalui komentar pada media sosial Twitter. Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada pengolahan data numerik hasil klasifikasi sentimen serta pengukuran tingkat akurasi model yang digunakan. Metode pengembangan sistem dalam penelitian ini mengadopsi model waterfall, yaitu model yang dilakukan secara bertahap dan sistematis mulai dari tahap perencanaan hingga evaluasi akhir.

Tahap pertama dalam penelitian ini adalah perencanaan, yang meliputi penentuan topik, identifikasi masalah, studi literatur, serta penentuan metode yang akan digunakan. Studi literatur dilakukan dengan mengkaji berbagai jurnal, buku, dan penelitian terdahulu yang berkaitan dengan analisis sentimen, kecerdasan buatan, Twitter sebagai sumber data, serta metode Support Vector Machine (SVM). Tahap ini bertujuan untuk memperkuat landasan teori dan memastikan metode yang digunakan relevan dengan permasalahan penelitian.

Tahap kedua adalah pengumpulan data (crawling data). Data penelitian diperoleh dari komentar pengguna Twitter pada dua postingan berbeda yang membahas tentang kemajuan AI ChatGPT. Pengambilan data dilakukan dengan metode by URL menggunakan layanan berbasis web ExportComments. Proses ini menghasilkan data mentah dalam bentuk file CSV yang berisi informasi seperti nama pengguna, isi komentar (tweet text), dan sumber tweet. Dalam penelitian ini digunakan sebanyak 600 komentar yang diambil secara proporsional dari dua postingan berbeda untuk menjaga keseimbangan data dan menghindari bias dari satu sumber saja.

Setelah data berhasil dikumpulkan, tahap berikutnya adalah preprocessing atau pembersihan data. Tahap ini bertujuan untuk menghilangkan noise atau gangguan dalam teks sehingga data menjadi lebih bersih dan siap dianalisis. Proses preprocessing meliputi beberapa langkah, yaitu penghapusan URL, emoji, simbol, angka, dan karakter khusus menggunakan regular expression. Selanjutnya dilakukan case folding untuk mengubah seluruh huruf menjadi huruf kecil agar tidak terjadi perbedaan token akibat perbedaan kapitalisasi. Kemudian dilakukan normalisasi dengan mengganti kata tidak baku menjadi kata baku menggunakan kamus kata baku yang telah disiapkan. Tahap berikutnya adalah tokenisasi, yaitu memecah kalimat menjadi kata per kata. Setelah itu dilakukan penghapusan stopword menggunakan library NLTK untuk menghilangkan kata-kata yang tidak memiliki makna penting dalam analisis sentimen. Tahap terakhir dari preprocessing adalah stemming menggunakan library Sastrawi untuk mengubah kata menjadi bentuk dasarnya sehingga variasi kata dapat diseragamkan.

Sebelum dilakukan pelabelan sentimen, data hasil preprocessing diterjemahkan terlebih dahulu ke dalam bahasa Inggris menggunakan library Googletrans. Proses translasi ini dilakukan karena metode pelabelan otomatis menggunakan TextBlob lebih optimal dalam bahasa Inggris. Setelah proses translasi selesai, setiap komentar dianalisis menggunakan TextBlob untuk mendapatkan nilai polarity. Nilai polarity tersebut kemudian dikategorikan menjadi tiga kelas sentimen, yaitu positif jika polarity bernilai lebih dari nol, negatif jika kurang

dari nol, dan netral jika bernilai nol. Hasil pelabelan ini digunakan sebagai data target (label) dalam proses klasifikasi menggunakan SVM.

Tahap selanjutnya adalah implementasi metode Support Vector Machine (SVM) sebagai algoritma klasifikasi. Data yang telah memiliki label dibagi menjadi data latih (80%) dan data uji (20%) menggunakan teknik train-test split dengan stratifikasi untuk menjaga proporsi kelas tetap seimbang. Untuk mendapatkan performa model yang optimal, dilakukan proses pencarian parameter terbaik menggunakan GridSearchCV dengan beberapa kombinasi parameter seperti nilai C, gamma, dan jenis kernel (linear dan rbf). Parameter terbaik yang diperoleh dari proses ini digunakan untuk membangun model SVM final yang kemudian diuji menggunakan data uji.

Tahap evaluasi dilakukan untuk mengukur kinerja model klasifikasi. Evaluasi dilakukan dengan menghitung nilai akurasi sebagai persentase prediksi yang benar terhadap seluruh data uji. Selain itu, ditampilkan confusion matrix untuk melihat distribusi prediksi benar dan salah pada masing-masing kelas sentimen. Evaluation juga diperkuat dengan classification report yang mencakup precision, recall, dan f1-score pada setiap kategori sentimen. Precision menunjukkan ketepatan prediksi model, recall menunjukkan kemampuan model dalam mendeteksi data yang sebenarnya termasuk dalam suatu kelas, sedangkan f1-score merupakan rata-rata harmonis dari precision dan recall.

Sebagai pendukung analisis, hasil penelitian juga divisualisasikan dalam bentuk wordcloud untuk mengetahui kata-kata yang paling sering muncul dalam komentar, serta diagram batang untuk menggambarkan distribusi jumlah sentimen positif, negatif, dan netral. Visualisasi ini membantu memberikan gambaran yang lebih jelas mengenai kecenderungan opini masyarakat terhadap perkembangan AI ChatGPT.

Dengan tahapan penelitian yang sistematis mulai dari pengumpulan data, preprocessing, pelabelan, klasifikasi, hingga evaluasi, diharapkan hasil penelitian ini memiliki tingkat validitas dan reliabilitas yang baik serta dapat memberikan gambaran yang objektif mengenai persepsi publik terhadap kemajuan kecerdasan buatan di media sosial Twitter.

3. HASIL DAN PEMBAHASAN

Setelah melakukan pengumpulan data berdasarkan tujuan dan prosedur penelitian yaitu untuk menganalisis sentimen komentar Twitter mengenai kemajuan kecerdasan buatan (AI ChatGPT) menggunakan metode Support Vector Machine (SVM). Proses analisis dilakukan melalui beberapa tahapan yang sistematis, mulai dari pengambilan data, preprocessing, pelabelan sentimen, implementasi model klasifikasi, hingga evaluasi dan interpretasi hasil.

3.1. Hasil Pengambilan dan Pengolahan Data

Data dalam penelitian ini diperoleh melalui proses crawling komentar Twitter menggunakan layanan berbasis web ExportComments dengan metode pengambilan berdasarkan URL (by URL). Sumber data berasal dari dua postingan Twitter yang membahas tentang kemajuan dan perkembangan kecerdasan buatan (AI ChatGPT). Pemilihan dua sumber postingan dilakukan untuk memperoleh variasi opini yang lebih luas serta menghindari bias dari satu akun atau satu kelompok pengguna saja. Dari proses crawling tersebut diperoleh total 600 komentar yang selanjutnya digunakan sebagai dataset penelitian.

Data mentah yang berhasil diunduh dalam format CSV berisi beberapa atribut seperti nama pengguna (Name), isi komentar (Tweet Text), serta sumber tweet (Tweet Source). Namun, pada tahap awal analisis hanya atribut yang relevan dengan isi teks komentar yang digunakan, yaitu kolom yang berisi teks komentar. Hal ini dilakukan agar proses analisis lebih terfokus pada konten opini yang disampaikan oleh pengguna.

Sebelum dilakukan analisis lebih lanjut, data mentah tersebut diperiksa terlebih dahulu untuk memastikan tidak terdapat duplikasi komentar atau data yang kosong. Proses pengecekan ini penting untuk menjaga validitas dan kualitas dataset. Komentar yang kosong atau tidak terbaca dihapus agar tidak mempengaruhi hasil analisis. Selain itu, dilakukan

pengecekan terhadap distribusi jumlah komentar dari masing-masing sumber untuk memastikan keseimbangan data tetap terjaga.

Tahap berikutnya adalah preprocessing atau pembersihan data. Pada tahap ini dilakukan proses cleaning data untuk menghilangkan elemen-elemen yang tidak diperlukan dalam analisis sentimen, seperti URL, emoji, simbol, angka, dan karakter khusus lainnya menggunakan regular expression. Penghapusan elemen tersebut bertujuan untuk mengurangi noise yang dapat mengganggu proses pembentukan fitur teks pada model klasifikasi.

Setelah proses cleaning selesai, dilakukan case folding dengan mengubah seluruh huruf dalam komentar menjadi huruf kecil (lowercase). Tahap ini bertujuan untuk menyeragamkan bentuk kata sehingga tidak terjadi perbedaan token akibat perbedaan kapitalisasi, misalnya antara "ChatGPT" dan "chatgpt" yang seharusnya dianggap sebagai kata yang sama.

Tahap selanjutnya adalah normalisasi kata tidak baku menjadi kata baku menggunakan kamus kata tidak baku yang telah disiapkan dalam format file eksternal. Proses ini sangat penting mengingat banyak komentar di media sosial menggunakan bahasa tidak formal, singkatan, atau kata gaul. Dengan normalisasi, kata-kata tersebut diubah menjadi bentuk standar sehingga lebih mudah dipahami oleh sistem.

Setelah normalisasi, dilakukan tokenisasi, yaitu proses memecah kalimat menjadi kata per kata. Tokenisasi mempermudah proses penghapusan stopword dan stemming pada tahap berikutnya. Kemudian dilakukan stopword removal menggunakan library NLTK untuk menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen, seperti "dan", "yang", "di", dan kata penghubung lainnya. Penghapusan stopword membantu sistem lebih fokus pada kata-kata yang mengandung makna emosional atau opini.

Tahap terakhir dalam preprocessing adalah stemming menggunakan library Sastrawi. Proses stemming bertujuan untuk mengubah kata menjadi bentuk dasarnya, misalnya "membantu", "dibantu", dan "membantunya" menjadi "bantu". Dengan demikian, variasi bentuk kata dapat diseragamkan sehingga mengurangi kompleksitas fitur dan meningkatkan efisiensi model dalam mengenali pola.

3.2. Hasil Translasi dan Pelabelan Sentimen

Setelah melalui tahap preprocessing, data komentar yang telah bersih kemudian diterjemahkan ke dalam bahasa Inggris menggunakan library Googletrans. Proses translasi ini dilakukan karena metode pelabelan otomatis yang digunakan dalam penelitian ini, yaitu TextBlob, memiliki performa yang lebih optimal dalam menganalisis teks berbahasa Inggris dibandingkan bahasa Indonesia. Proses penerjemahan dilakukan secara otomatis pada seluruh komentar hasil preprocessing sehingga setiap teks dapat dianalisis nilai polaritasnya secara konsisten. Meskipun penerjemahan otomatis berpotensi menyebabkan sedikit pergeseran makna, secara umum hasil translasi tetap mampu mempertahankan konteks utama dari setiap komentar.

Setelah proses translasi selesai, tahap selanjutnya adalah pelabelan sentimen menggunakan TextBlob. Pelabelan dilakukan berdasarkan nilai polarity yang dihasilkan dari setiap komentar. Nilai polarity berada pada rentang -1 hingga 1, di mana nilai lebih dari 0 dikategorikan sebagai sentimen positif, nilai kurang dari 0 sebagai sentimen negatif, dan nilai sama dengan 0 sebagai sentimen netral. Berdasarkan hasil analisis terhadap 600 komentar, diperoleh 264 komentar (44,07%) termasuk dalam kategori sentimen positif, 255 komentar (42,57%) termasuk dalam kategori sentimen netral, dan 80 komentar (13,36%) termasuk dalam kategori sentimen negatif.

Hasil tersebut menunjukkan bahwa mayoritas pengguna Twitter memberikan tanggapan positif terhadap kemajuan AI ChatGPT, sementara sebagian besar lainnya memberikan tanggapan yang bersifat netral atau informatif tanpa menunjukkan emosi yang kuat. Jumlah sentimen negatif relatif lebih sedikit, yang menunjukkan bahwa meskipun terdapat kekhawatiran terhadap perkembangan AI, proporsinya tidak dominan dibandingkan pandangan positif dan netral. Distribusi sentimen ini memberikan gambaran awal mengenai

kecenderungan opini publik yang selanjutnya digunakan sebagai dasar dalam proses klasifikasi menggunakan metode Support Vector Machine (SVM).

3.3. Hasil Visualisasi Data

3.3.1. Wordcloud

Hasil visualisasi wordcloud menunjukkan bahwa kata "ChatGPT" merupakan kata yang paling sering muncul dalam komentar. Selain itu, muncul kata-kata seperti "AI", "teknologi", "kerja", "belajar", dan "bantu". Dominasi kata-kata tersebut menunjukkan bahwa pembahasan publik banyak berfokus pada fungsi dan manfaat AI dalam kehidupan sehari-hari. Visualisasi ini memperkuat hasil distribusi sentimen yang menunjukkan kecenderungan opini positif dan netral terhadap kemajuan AI ChatGPT.

3.3.2. Implementasi SVM

Setelah seluruh tahapan sebelumnya selesai, langkah berikutnya adalah mengimplementasikan metode klasifikasi Support Vector Machine (SVM). Proses ini diawali dengan membagi dataset menjadi data latih dan data uji dengan perbandingan 80:20. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur dan mengevaluasi kemampuan

model dalam mengklasifikasikan sentimen secara akurat. Pembagian data dilakukan secara terstratifikasi agar proporsi masing-masing kelas sentimen tetap seimbang pada data latih dan data uji.

Untuk memperoleh performa model yang optimal, dilakukan proses pencarian parameter terbaik menggunakan teknik Grid Search dengan cross-validation sebanyak lima kali. Parameter yang diuji meliputi nilai C (yang mengatur tingkat kompleksitas model), parameter gamma (yang mempengaruhi pola pemisahan data pada kernel tertentu), serta jenis kernel yang digunakan, yaitu linear dan radial basis function (RBF). Seluruh kombinasi parameter diuji secara sistematis, kemudian dipilih kombinasi yang menghasilkan performa terbaik berdasarkan hasil validasi silang. Model dengan konfigurasi parameter terbaik inilah yang selanjutnya digunakan sebagai model akhir dalam penelitian.

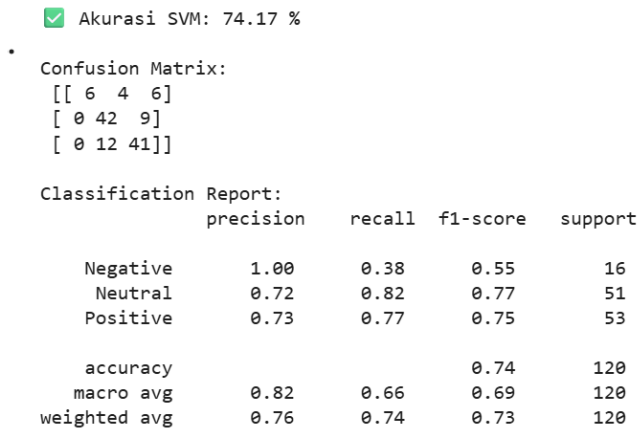
Model terbaik kemudian digunakan untuk memprediksi label sentimen pada data uji. Hasil prediksi tersebut dibandingkan dengan label sebenarnya untuk menghitung nilai akurasi. Berdasarkan hasil evaluasi, model SVM memperoleh tingkat akurasi sebesar 74%, yang berarti 74% dari seluruh data uji berhasil diklasifikasikan dengan benar sesuai dengan label aslinya.

Selain akurasi, evaluasi performa model juga dilakukan menggunakan confusion matrix dan classification report. Confusion matrix digunakan untuk melihat distribusi prediksi benar dan salah pada masing-masing kelas sentimen. Sementara itu, classification report menampilkan tiga metrik utama, yaitu precision, recall, dan F1-score. Precision menunjukkan tingkat ketepatan prediksi model dalam suatu kelas, recall menggambarkan kemampuan model dalam mengenali seluruh data yang termasuk dalam kelas tersebut, sedangkan F1-score merupakan rata-rata harmonis antara precision dan recall yang memberikan gambaran keseimbangan performa model.

Hasil evaluasi menunjukkan bahwa kelas Negative memiliki nilai precision yang sangat tinggi, yang berarti ketika model memprediksi sentimen negatif, hasilnya hampir selalu benar. Namun, nilai recall pada kelas ini relatif rendah, yang menunjukkan bahwa model belum mampu mendeteksi seluruh komentar negatif secara optimal. Sebaliknya, kelas Neutral dan Positive menunjukkan performa yang lebih stabil dengan nilai precision dan recall yang relatif seimbang, sehingga menghasilkan F1-score yang cukup baik. Selain itu, nilai macro average menunjukkan rata-rata performa seluruh kelas tanpa mempertimbangkan jumlah data, sedangkan weighted average memperhitungkan distribusi jumlah data pada masing-masing kelas sehingga memberikan gambaran yang lebih proporsional terhadap kondisi dataset.

3.3.3. Hasil Klasifikasi SVM

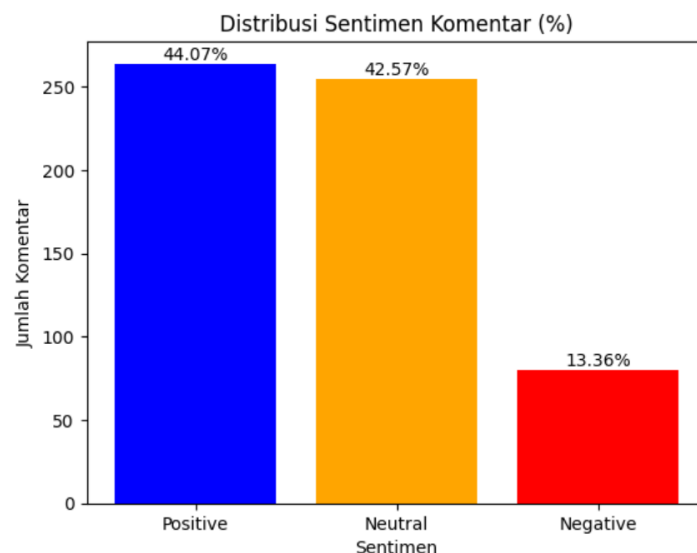
Hasil evaluasi model klasifikasi sentimen yang menggunakan algoritma SVM (Support Vector Machine). Dari data yang digunakan, terdapat 255 komentar netral, 264 komentar positif, dan 80 komentar negatif. Setelah dilakukan pelatihan dan pengujian, model SVM menghasilkan akurasi sebesar 74,17%, yang berarti bahwa dari seluruh data uji, sekitar tiga perempatnya berhasil diprediksi dengan benar. Data ini dapat dilihat pada gambar di bawah ini:



Gambar 1. Hasil kasivikasi svm

Pada bagian confusion matrix, terlihat bagaimana model menebak setiap kategori. Misalnya, untuk komentar negatif, hanya 6 yang berhasil diprediksi dengan benar, sementara sisanya keliru diklasifikasikan sebagai netral atau positif. Untuk kelas netral dan positif, model menunjukkan performa yang lebih baik, terutama karena jumlah datanya lebih besar sehingga model dapat belajar lebih banyak pola.

Kemudian, pada classification report, ditampilkan nilai precision, recall, dan f1-score untuk tiap sentimen. Kategori negatif memiliki recall yang rendah (0.38), artinya model masih sering gagal mengenali komentar yang sebenarnya negatif. Sebaliknya, kategori netral dan positif memiliki keseimbangan performa yang lebih baik, dengan f1-score masing-masing sekitar 0.77 dan 0.75. Secara keseluruhan, model bekerja cukup baik, terutama untuk komentar netral dan positif, tetapi masih perlu diperbaiki dalam mendeteksi komentar negatif yang jumlah datanya lebih sedikit.



Gambar 2. Gambar diagram batang

Gambar tersebut memperlihatkan sebaran jenis sentimen dari komentar yang telah dianalisis. Dari grafik, terlihat bahwa sebagian besar komentar bersifat **positif**, yaitu sekitar **44, 07%**, yang berarti banyak orang memberikan komentar positif dan mendukung tentang adanya AI. Jumlah komentar **netral** juga hampir sama, sekitar **42, 57%**, menunjukkan bahwa banyak pengguna menyampaikan tanggapan yang biasa saja tanpa ekspresi sedikitpun. Sementara itu, komentar **negatif** hanya sekitar **13, 36%**, sehingga dapat disimpulkan bahwa hanya sedikit pengguna yang menuliskan komentar bernada tidak menyenangkan atau kritis. Secara keseluruhan, grafik ini menunjukkan bahwa komentar yang masuk cenderung lebih netral dan positif daripada negative.

Berdasarkan hasil penelitian, metode Support Vector Machine (SVM) menunjukkan kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar Twitter terkait kemajuan AI ChatGPT dengan tingkat akurasi sebesar 74,17%. Nilai ini menunjukkan bahwa model mampu mengenali pola teks dan mengelompokkan opini publik secara cukup efektif. Performa yang stabil pada kelas positif dan netral mengindikasikan bahwa model mampu menangkap karakteristik bahasa yang umum digunakan dalam komentar yang bersifat apresiatif maupun informatif. Namun demikian, performa pada kelas negatif masih relatif rendah, khususnya pada nilai recall, yang menunjukkan bahwa model belum mampu mendeteksi seluruh komentar negatif secara optimal. Kondisi ini kemungkinan dipengaruhi oleh ketidakseimbangan jumlah data antar kelas, di mana komentar negatif memiliki jumlah yang lebih sedikit dibandingkan kelas lainnya.

Dominasi sentimen positif dalam hasil penelitian menunjukkan bahwa mayoritas masyarakat memiliki pandangan yang baik terhadap perkembangan AI ChatGPT [10]. Hal ini dapat diinterpretasikan sebagai bentuk penerimaan terhadap inovasi teknologi yang dianggap mampu memberikan kemudahan, efisiensi, serta dukungan dalam berbagai bidang seperti pendidikan,

pekerjaan, dan pencarian informasi [11]. Sementara itu, tingginya jumlah sentimen netral menunjukkan bahwa sebagian masyarakat masih bersikap objektif dan informatif dalam menyikapi perkembangan AI, tanpa menunjukkan sikap dukungan atau penolakan secara tegas [12]. Di sisi lain, keberadaan sentimen negatif tetap menjadi aspek penting yang perlu diperhatikan, karena mencerminkan adanya kekhawatiran terkait isu etika, keamanan data, potensi penyalahgunaan teknologi, serta dampaknya terhadap lapangan pekerjaan manusia.

Dari sisi metodologi, penggunaan translasi sebelum pelabelan sentimen memberikan kemudahan dalam penerapan TextBlob, namun berpotensi menyebabkan sedikit perubahan makna akibat proses penerjemahan otomatis [13]. Hal ini dapat memengaruhi akurasi pelabelan awal yang menjadi dasar dalam pelatihan model SVM [14]. Selain itu, karakteristik bahasa pada media sosial yang cenderung tidak formal, menggunakan singkatan, ironi, atau sarkasme, juga menjadi tantangan tersendiri dalam analisis sentimen berbasis machine learning [15]. Model klasifikasi berbasis teks seperti SVM umumnya bekerja berdasarkan pola kata, sehingga sulit mengenali konteks implisit atau makna tersirat dalam suatu komentar.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa analisis sentimen berbasis machine learning dapat menjadi alat yang efektif untuk memahami opini publik secara cepat dan sistematis. Penelitian ini juga membuktikan bahwa media sosial dapat dimanfaatkan sebagai sumber data yang relevan untuk mengkaji respons masyarakat terhadap perkembangan teknologi. Ke depan, performa model dapat ditingkatkan dengan menyeimbangkan jumlah data pada setiap kelas, menggunakan metode pelabelan manual untuk meningkatkan kualitas label, atau membandingkan SVM dengan algoritma lain seperti Naïve Bayes, Random Forest, maupun metode berbasis deep learning. Dengan pengembangan lebih lanjut, analisis sentimen dapat memberikan gambaran yang lebih akurat dan mendalam mengenai persepsi masyarakat terhadap kemajuan kecerdasan buatan.

4. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen komentar Twitter mengenai kemajuan kecerdasan buatan (AI ChatGPT) menggunakan metode Support Vector Machine (SVM), dapat disimpulkan bahwa Twitter merupakan media yang efektif untuk mengetahui

pandangan dan respons masyarakat terhadap perkembangan teknologi AI. Komentar pengguna mencerminkan beragam persepsi, mulai dari dukungan, kekhawatiran, hingga sikap netral. Metode SVM terbukti mampu mengklasifikasikan sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral, dengan tingkat akurasi sebesar 74,17%, yang menunjukkan bahwa model memiliki performa yang cukup baik dalam klasifikasi teks berbasis media sosial.

Hasil klasifikasi menunjukkan terdapat 80 komentar bersentimen negatif, 255 komentar netral, dan 364 komentar positif, sehingga sentimen positif menjadi yang paling dominan. Hal ini menunjukkan bahwa sebagian besar masyarakat memiliki pandangan yang baik terhadap kemajuan AI ChatGPT, terutama dari segi manfaat dan kemudahan yang diberikan. Namun demikian, masih terdapat sentimen negatif yang mencerminkan kekhawatiran terkait isu etika, privasi, dan dampak terhadap lapangan pekerjaan, serta sentimen netral yang menunjukkan sikap informatif tanpa kecenderungan tertentu. Secara keseluruhan, penelitian ini memberikan gambaran komprehensif mengenai persepsi publik terhadap AI dan dapat menjadi bahan pertimbangan bagi pengembang teknologi maupun peneliti selanjutnya.

DAFTAR PUSTAKA

- Jusman, J., Hajar, A., & Habibi, A. (2024). Analisis pemanfaatan kecerdasan buatan berbasis Chat GPT untuk membantu mahasiswa jurusan Teknologi Pendidikan di Universitas Muhammadiyah Bone. *Jurnal Ilmiah Wahana Pendidikan*, 10(11), 791-798. <https://doi.org/10.5281/zenodo.12888797>
- Koonchanok, R., Pan, Y., & Jang, H. (2023). Public Attitudes Toward ChatGPT on Twitter: Sentiments, Topics, and Occupations. *arXiv*. <https://arxiv.org/abs/2306.12951>
- Septini, A., Susanto, & Elmayati. (2024). Analisis Sentimen Masyarakat di Twitter Mengenai Open AI ChatGPT Menggunakan Metode Support Vector Machine (SVM). *Bulletin of Computer Science Research*, 5(2). <https://doi.org/10.47065/bulletincsr.v5i2.475>
- Andini, D., & Sari, N. A. (2023). Analisis Sentimen Masyarakat terhadap Kebijakan Pemerintah Menggunakan Metode Support Vector Machine (SVM). *Jurnal Teknologi Informasi dan Komputer*, 9(1), 45-52. <https://journal.unp.ac.id>
- Simon, Herbert A. 1987. *Artificial Intelligence: An Empirical Science*. Massachusetts: MIT Press.
- Knight, Kevin dan Rich, Elaine. 1991. *Artificial Intelligence*. New York: McGraw-Hill.
- Badan Pengembangan dan Pembinaan Bahasa. 2024. *Kamus Besar Bahasa Indonesia (KBBI): Kecerdasan Buatan*. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia. Diakses dari <https://kbbi.kemdikbud.go.id>
- Van Rossum, Guido. 1991. *Python Programming Language*. Python Software Foundation. Diakses dari <https://www.python.org>
- Export Comments. 2024. *Export Comments – Social Media Data Extractor*. Diakses dari <https://exportcomments.com>
- Google Research. 2024. *Google Colaboratory*. Diakses dari <https://colab.research.google.com>
- Saputra, H. (2025). Penerapan teknik crawling data untuk analisis sentimen media sosial Twitter. *Jurnal Sains Data Indonesia*, 3(2), 45–53. 13, 2019).
- Kumar, A., & Garg, N. (2022). *Sentiment analysis on Twitter data using machine learning*

techniques. *Procedia Computer Science*, 218, 1310–1319.
<https://doi.org/10.1016/j.procs.2022.06.171>

C. Low, "NSL-KDD Dataset," 2015. https://github.com/defcom17/NSL_KDD (accessed Sep. 13, 2019).

Amrullah, M. S., & Pane, S. F. (2023). Analisis sentimen masyarakat terhadap kebijakan polisi tilang manual di Indonesia [E-book]. Penerbit Buku Pedia. https://www.google.co.id/books/edition/Analisis_Sentimen_Masyarakat_Terhadap_Ke/_Uq5EAAAQBAJ?hl=id&gbpv=1

Urva, G., Desyanti, & Albanna, I., Sungkar, M. S., Gunawan, I. A. O., Adhicandra, I., Ramadhan, S., Rahardian, R. L., Herlawati, & Handayanto, R. T., Ariana, A. A. G. B., Hartatik, Atika, P. D., & Junaidi, S. (2023). Penerapan data mining di berbagai bidang: Konsep, metode, dan studi kasus [E-book]. PT. Sonpedia Publishing Indonesia. https://www.google.co.id/books/edition/PENERAPAN_DATA_MINING_DI_BERBAGAI_BIDANG/uq6-EAAAQBAJ?hl=id&gbpv=1