BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

Analisis Sentimen Masyarakat di Twitter Mengenai Open AI CHATGPT Menggunakan Metode Support Vector Machine (SVM)

Ayu Septini[1], Susanto[2,*], Elmayati[3]

[1] Ilmu Teknik, Sistem informasi, Universitas Bina Insan, Lubuklinggau, Indonesia
[2] Ilmu Teknik, Informatika, Universitas Bina Insan, Lubuklinggau, Indonesia
[3] Ilmu Teknik, Sistem informasi, Universitas Bina Insan, Lubuklinggau, Indonesia
Email: [1]2002030032@mhs.univbinainsan.ac.id, [2,*]susanto@univbinainsan.ac.id, [3]elmayati@univbinainsan.ac.id
Email Penulis Korespondensi: susanto@univbinainsan.ac.id

Abstrak−Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap teknologi OpenAI ChatGPT di media sosial Twitter menggunakan metode *Support Vector Machine* (SVM). Latar belakang penelitian ini didasari oleh peningkatan penggunaan internet dan kecerdasan buatan (AI) secara global, serta peran media sosial sebagai platform bagi masyarakat untuk mengungkapkan pendapat mereka. Penelitian ini menggunakan penelitian kualitatif dengan menggunakan metode *Support Vector Machine*, dengan pengumpulan data melalui data primer yakni *Crawling* data di media sosial twitter. Penelitian ini menggunakan data yang diperoleh dari 4305 tweet berbahasa Indonesia sebanyak 4305 yang dikumpulkan dari bulan Januari hingga September 2023. Tweet-tweet tersebut kemudian diklasifikasikan menjadi sentimen positif, netral, dan negatif menggunakan metode SVM. Hasil penelitian menunjukkan bahwa dari total data yang terkumpul, 2196 tweet memiliki sentimen netral, 1500 tweet memiliki sentimen positif, dan 591 tweet memiliki sentimen negatif. Dalam evaluasi performa model, data *training* dengan perbandingan 80:20 menghasilkan akurasi tertinggi sebesar 94,25%, sedangkan data t*esting* dengan perbandingan 70:30 menghasilkan akurasi tertinggi sebesar 93,16%. Selain itu, penggunaan cross-validation k-10 data *training* menghasilkan akurasi sebesar 89,94% dan untuk data *testing* menghasilkan akurasi rata – rata 78,17%.

Kata Kunci: Analisis Sentimen; CHATGPT; Support Vector Machine; Media Sosial; Twitter

Abstract−This study aims to analyze public sentiment toward OpenAI ChatGPT technology on Twitter using the Support Vector Machine (SVM) method. The background of this research is based on the increasing global use of the internet and artificial intelligence (AI), as well as the role of social media as a platform for people to express their opinions. This study employs a qualitative research approach using the Support Vector Machine method, with data collection conducted through primary data obtained by crawling data from Twitter. The research uses data collected from 4,305 Indonesian-language tweets gathered between January and September 2023. These tweets were then classified into positive, neutral, and negative sentiments using the SVM method. The results indicate that out of the total collected data, 2,196 tweets had a neutral sentiment, 1,500 tweets had a positive sentiment, and 591 tweets had a negative sentiment. In the model performance evaluation, training data with an 80:20 ratio achieved the highest accuracy of 94.25%, while testing data with a 70:30 ratio achieved the highest accuracy of 93.16%. Additionally, the use of 10-fold cross-validation on training data resulted in an accuracy of 89.94%, while testing data achieved an average accuracy of 78.17%.

Kata Kunci: Sentiment Analysis; CHATGPT; Support Vector Machine; Social Media; Twitter

1. PENDAHULUAN

Jumlah pengguna internet dunia terus meningkat pesat. Pada awal 2023, tercatat ada 5,15 miliar orang yang terhubung ke internet, atau setara dengan 64,4% dari populasi global. Ini menunjukkan pertumbuhan sebesar 1,9% dibandingkan tahun sebelumnya [1]. Pertumbuhan ini dipengaruhi oleh perkembangan teknologi yang pesat, serta meningkatnya kebutuhan akan akses informasi di berbagai belahan dunia. Salah satu teknologi yang sedang tren dan banyak digunakan oleh masyarakat adalah teknologi kecerdasan buatan (AI). AI kini menjadi elemen yang tak terpisahkan dari kehidupan manusia modern dan diaplikasikan di berbagai sektor, termasuk industri, bisnis, perawatan kesehatan, pendidikan, dan lainnya [2]. Teknologi AI memberikan kemudahan dalam otomatisasi proses bisnis, analisis data, hingga pengambilan keputusan yang lebih efisien.

Salah satu contoh paling populer dari penerapan AI adalah teknologi Chat GPT dari OpenAI. Chat GPT adalah model bahasa yang dirancang untuk merespons berbagai perintah dalam bentuk pertanyaan atau pernyataan. Dengan kemampuannya dalam memahami dan menghasilkan teks yang menyerupai manusia, Chat GPT telah menjadi alat yang sangat berguna dalam berbagai bidang, mulai dari layanan pelanggan hingga pendidikan. Meskipun memiliki kemampuan luar biasa, Chat GPT juga menuai kontroversi, seperti dugaan bahwa kehadirannya dapat mengurangi minat pada forum pembelajaran, serta menggantikan pekerjaan manusia dengan AI, termasuk dalam sistem pendidikan di mana siswa dapat mencontek jawaban langsung dari ChatGPT tanpa berusaha menyelesaikan tugas secara mandiri [1]. Selain itu, kekhawatiran tentang privasi dan keamanan data juga menjadi perhatian utama dalam penggunaan teknologi AI seperti Chat GPT.

Selain AI, media sosial juga menjadi teknologi yang banyak digunakan oleh masyarakat modern. Media sosial menyediakan platform inklusif yang memungkinkan semua orang, tanpa memandang usia atau latar belakang sosial, untuk mengakses dan berbagi informasi. Pada Januari 2020, data We Are Social menunjukkan bahwa di Indonesia, Twitter telah meraih popularitas yang signifikan sebagai salah satu platform media sosial terkemuka, dengan total pengguna mencapai 10,65 juta orang. Twitter masuk dalam lima besar platform media sosial paling populer dengan persentase 56% di antara pengguna berusia 16-64 tahun [3]. Melalui Twitter, pengguna dapat berinteraksi, berbagi informasi, serta mengungkapkan pendapat tentang berbagai isu, termasuk perkembangan teknologi. Twitter juga menjadi

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

salah satu platform utama di mana pengguna mengungkapkan opini mereka tentang Chat GPT. Beberapa orang menganggap Chat GPT sebagai ancaman terhadap pekerjaan manusia, sementara yang lain melihatnya sebagai peluang untuk meningkatkan efisiensi dan produktivitas [4].

Untuk mengetahui pandangan masyarakat mengenai teknologi Chat GPT, diperlukan sebuah analisis sentimen. Analisis sentimen merupakan pendekatan yang digunakan untuk mengidentifikasi kecenderungan seseorang dalam teks untuk mendapatkan informasi tentang sentimen yang diungkapkan, baik itu positif, negatif, atau netral. Analisis ini menjadi sangat penting dalam memahami persepsi masyarakat terhadap fenomena baru, termasuk teknologi seperti Chat GPT, terutama melalui media sosial [5]. Melalui analisis sentimen, kita dapat memperoleh gambaran yang lebih jelas tentang bagaimana teknologi ini diterima oleh masyarakat luas dan apa saja kekhawatiran atau harapan yang muncul seiring dengan penggunaannya.

Dalam penelitian ini, kami menerapkan metode Support Vector Machine (SVM) untuk analisis sentimen. SVM merupakan salah satu algoritma pembelajaran mesin yang banyak digunakan dalam prediksi, baik dalam konteks regresi maupun klasifikasi. Algoritma ini bekerja dengan menemukan hyperplane yang memisahkan data ke dalam dua kategori berbeda dengan margin terbesar, sehingga meminimalkan kesalahan klasifikasi [5]. Metode ini terbukti efektif dalam menganalisis pola dalam data yang kompleks dan membantu memprediksi hasil berdasarkan karakteristik data. Dalam konteks analisis sentimen, SVM digunakan untuk mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral.

Sudah banyak penelitian yang menggunakan SVM untuk analisis sentimen. Misalnya, penelitian oleh Primandani Arsi dkk. menunjukkan bahwa SVM dapat mencapai tingkat akurasi sebesar 96,68% pada pengujian menggunakan Confusion Matrix [6]. Penelitian lain oleh Melati Indah Petiwi dkk. membandingkan metode SVM dengan Naïve Bayes dan menemukan bahwa SVM mencapai akurasi sebesar 83% dibandingkan dengan Naïve Bayes yang hanya mencapai 74,6% [7]. Kedua penelitian tersebut menunjukkan bahwa SVM memiliki performa yang lebih baik dalam analisis sentimen, terutama dalam konteks data teks yang berasal dari media sosial.

Namun, penelitian sebelumnya lebih berfokus pada tingkat akurasi secara umum tanpa mengeksplorasi lebih dalam mengenai penerapan SVM, khususnya dalam konteks opini masyarakat tentang teknologi baru seperti Chat GPT di media sosial. Oleh karena itu, penelitian ini berupaya untuk mengisi kesenjangan tersebut dengan mengkaji secara mendalam bagaimana opini publik terhadap Chat GPT, yang merupakan fenomena baru di dunia teknologi dan sosial media. Chat GPT tidak hanya mempengaruhi industri teknologi, tetapi juga berbagai sektor lain seperti pendidikan, kesehatan, dan layanan publik, sehingga penting untuk memahami bagaimana masyarakat menilai dan merespons teknologi ini.

Berdasarkan kondisi yang ada, penelitian ini berfokus mengkaji opini publik yang diungkapkan di platform Twitter mengenai teknologi OpenAI Chat GPT, dengan menggunakan metode Support Vector Machine (SVM). Studi ini berbeda dari penelitian sebelumnya karena fokus pada opini publik terhadap Chat GPT, yang merupakan fenomena baru di dunia teknologi dan sosial media. Selain itu, studi ini juga mengeksplorasi penggunaan metode SVM dalam klasifikasi sentimen secara lebih mendalam, terutama dalam konteks data yang berasal dari media sosial seperti Twitter. Data yang dikumpulkan akan dianalisis untuk memahami persepsi masyarakat, baik positif, negatif, maupun netral, terhadap Chat GPT.

Penelitian ini diharapkan dapat memberikan wawasan lebih lanjut tentang bagaimana teknologi baru seperti Chat GPT dipersepsikan oleh masyarakat, serta implikasi dari sentimen tersebut terhadap adopsi dan penggunaan teknologi di masa depan. Hasil penelitian ini juga dapat menjadi acuan bagi pengembang teknologi dalam memperbaiki dan meningkatkan kualitas produk mereka, dengan mempertimbangkan masukan dari pengguna. Selain itu, penelitian ini juga memberikan kontribusi terhadap pengembangan metode analisis sentimen yang lebih efektif, khususnya dalam konteks media sosial dan opini publik terhadap teknologi baru.

Dengan semakin berkembangnya AI dan teknologi digital lainnya, analisis sentimen menjadi semakin relevan sebagai alat untuk memahami persepsi masyarakat secara lebih komprehensif. Penerapan SVM dalam konteks ini menawarkan pendekatan yang lebih akurat dan dapat diandalkan, sehingga hasil dari penelitian ini diharapkan dapat memberikan kontribusi positif bagi pengembangan ilmu pengetahuan dan teknologi.

2. METODE PENELITIAN

2.1 Kerangka Dasar Penelitian

a. Metode penelitian
   Metode penelitian yang digunakan dalam studi ini adalah kuantitatif. Penulis menganalisis informasi yang dikumpulkan dari Twitter untuk menilai opini publik tentang open AI CHATGPT.
b. Metode Pengumpulan Data
   1. Studi pustaka
      Studi pustaka merupakan cara untuk mengumpulkan informasi kepustakaan yang dilakukan dengan mengumpulkan informasi dari jurnal, literatur, buku, dan situs internet sebagai sumber literatur yang relevan dengan subjek penelitian. Metode studi pustaka ini terutama berkaitan dengan analisis sentimen menggunakan metode vector support machine [8].

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

2. Data primer
   Data yang kami kumpulkan secara langsung untuk penelitian ini dataset yang digunakan berasal dari twitter. Melakukan pengumpulan data dilakukan menggunakan crawl data. Data yang diperoleh berupa tweets yang mencantumkan kata kunci tentang open AI chat GPT. Seluruh data dari jangka waktu bulan Januari - September 2023.

c. Metode analisa
Analisis sentimen menjadi metode utama yang digunakan penulis dalam penelitian ini dan metode klasifikasi menggunakan *support vector machine*. Analisis sentimen adalah cara untuk mengetahui apakah suatu teks itu bernada positif, negatif, atau netral [9].

d. Metode Pengujian
Pengujian dalam penelitian ini dilakukan dengan memanfaatkan pustaka yang ada Support Vector Machine yaitu confussion matrix. Menurut Prasetyo, klasifikasi adalah Pengelompokan data berdasarkan ciri-cirinya ke dalam kelompok yang sudah ditentukan [10]. Algoritma SVM merupakan metode pembelajaran mesin yang diawasi yang digunakan untuk melakukan klasifikasi data**.** karena dalam proses pelatihan memerlukan suatu tujuan atau target pembelajaran yang telah ditentukan sebelumnya. SVM adalah metode yang mengubah data menjadi bentuk yang lebih kompleks agar bisa dibedakan dengan lebih mudah [9]. Proses validasi bertujuan untuk memverifikasi akurasi data dan temuan dalam penelitian. Hasil validasi yang didapat berupa accuracy, precision, FPR, Sensitivitas dan Spesifisitas. Pada penelitian ini penulis memakai tiga nilai perbandingan dari data Training dan data Testing, yaitu: Perbandingan data pelatihan dan pengujian yang digunakan adalah 60:40, 70:30, dan 80:20 [11].
Dalam konteks penilaian dan evaluasi model klasifikasi, metrik-metrik yang umum digunakan meliputi:

a. *Accuracy*
   Accuracy merupakan efektivitas model dalam memprediksi. *Accuracy* untuk menilai sejauh mana nilai prediksi sesuai dengan nilai aktual [10]. Adapun rumus penghitungan *Accuracy* dapat dilihat pada persamaan berikut.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

b. *Precision*
   Precision merupakan ukuran seberapa akurat hasil pencarian yang diberikan oleh sistem. Dengan kata lain, precision menunjukkan seberapa tepat hasil pencarian yang diberikan sesuai dengan apa yang kita cari. Dalam kata lain, *precision* menunjukkan seberapa akurat sistem dalam memberikan jawaban yang relevan terhadap pertanyaan atau permintaan pengguna [10]. Rumus *Precision* dapat dilihat pada persamaan berikut.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

c. *FPR* (*False Positive Rate*)
   Berdasarkan analisa, akronim "FPR" merupakan singkatan dari "False Positive Rate" atau "Salah Tingkat Positif Rate". Metrik ini digunakan untuk mengoreksi beberapa hasil yang sering muncul dan jelas-jelas negatif tetapi disalah artikan sebagai positif oleh sistem atau model tertentu. FPR dihitung dengan membandingkan jumlah kasus positif dengan jumlah keseluruhan kasus negatif, yang biasanya dinyatakan dalam persentase. Semakin banyak FPR yang dibayarkan, semakin baik kinerja model dalam mengklasifikasi ulang situasi negatif menjadi situasi positif [10]. Rumus FPR dapat dilihat pada persamaan berikut.

$$FPR = \frac{FP}{FP+TN} \quad (3)$$

d. *Sensitivity*
   Sensitivity adalah ukuran sejauh mana program mampu mengklasifikasikan data ke dalam kelompok yang benar [12]. Rumus *Sensitivity* seperti di bawah ini.

$$Sensitifitas = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (4)$$

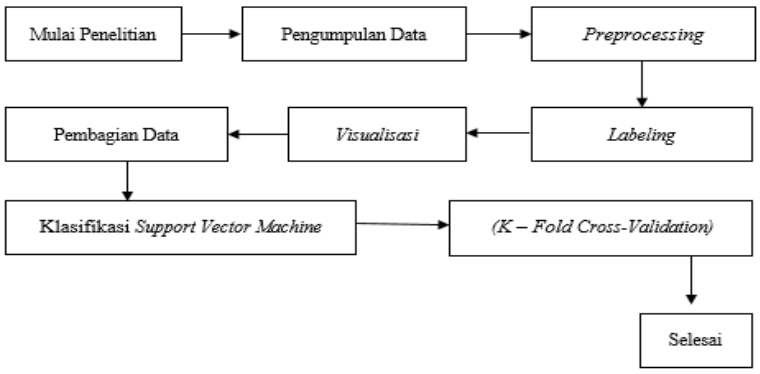
e. *Specificity*
   Specificity dapat diartikan sebagai persentase di mana Program mengklasifikasikan data ke kelas yang salah [12]. Rumus Spesifisitas dapat dilihat pada persamaan berikut.

$$Specificity = \frac{\mathrm{TN}}{\mathrm{TN+FP}} \quad (5)$$

2.2 Tahapan Penelitian

Berikut Gambar 1 merupakan tahapan dari penelitian:

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475



Gambar 1. Tahap Penelitian

Dari Gambar 1 tersebut dapat dijelaskan:

a. Pengumpulan data
Studi ini memakai dua macam data. Data primer diperoleh langsung dari platform Twitter, sedangkan data sekunder diambil dari sumber-sumber lain. Melakukan pengumpulan data dilakukan menggunakan crawl data. Data yang Data yang diperoleh berupa tweets yang mencantumkan kata kunci tentang open AI chat GPT. Seluruh data dari jangka waktu bulan Januari - September 2023. Kemudian untuk data sekunder dilakukan dengan mengumpulkan informasi dari jurnal, literatur, buku, dan situs internet sebagai sumber literatur yang relevan dengan subjek penelitian. Metode studi pustaka ini terutama berkaitan dengan analisis sentimen menggunakan metode vector support machine

b. Tahap *Preprocessing*
Preprocessing merupakan langkah menormalisasi istilah dalam kalimat agar data latih lebih baik dan fitur yang diekstrak sesuai keinginan [13]. Tujuannya adalah membuat pemrosesan data lebih mudah dan sederhana. Tahap *Preprocessing* sebagai berikut*:*
   1. *Cleaning* adalah Pada tahap preprocessing, dilakukan pembersihan teks dengan menghilangkan karakter non-alfanumerik, termasuk tanda baca dan simbol-simbol khusus, untuk menghasilkan corpus teks yang lebih murni [14].
   2. *Case folding* adalah Merubah teks menjadi huruf kecil (*lowercase*) bertujuan untuk menghapus perbedaan yang disebabkan oleh variasi penulisan huruf besar dan kecil dalam dokumen teks [4].
   3. Tokenisasi adalah Tahap ini berperan sebagai pembagi kalimat berdasarkan setiap kata yang membentuknya, Tokenizing dilakukan dengan memisahkan kata-kata berdasarkan spasi [14].
   4. *Stopword removal* merupakan teknik pra-pemrosesan teks yang bertujuan untuk menghilangkan kata-kata yang sering terlihat namun tidak memiliki dampak yang besar terhadap makna keseluruhan suatu teks. Kata-kata seperti "apakah", "ke", "luar", "biasanya" termasuk dalam kategori ini[4].
   5. *Stemming* adalah cara untuk menyederhanakan kata dengan menghapus bagian-bagian yang tidak penting seperti kata sambung, kata ganti, dan kata depan. Proses ini dilakukan dengan menghapus awalan atau akhiran dari kata tersebut[6].

c. *Labeling*
*Labelin**g*** adalah proses pemberian kategori atau label pada data untuk tujuan analisis atau pelatihan model. Dalam konteks pembelajaran mesin dan analisis data, labeling umumnya mengacu pada penandaan data dengan informasi tambahan yang menunjukkan kategori atau kelas yang relevan.

d. *Visualisasi*
Pada tahap visualisasi, analisis data dilakukan menggunakan *library* matplotlib untuk menghasilkan *pie chart* yang menunjukkan proporsi kategori sentimen yang ditemukan dalam data. Selain itu, *wordcloud* digunakan untuk menunjukkan frekuensi kata-kata yang sering terlihat dalam tweet, memberikan gambaran visual tentang istilah-istilah kunci yang berhubungan dengan topik CHAT GPT.

e. Pembagian data
Pada penelitian memakai tiga nilai perbandingan data *Training* dan data *Testing* Perbandingan data pelatihan dan pengujian yang digunakan adalah 60:40, 70:30, dan 80:20.

f. Klasifikasi Menggunakan Metode *Support Vector Machine*
SVM dipakai untuk menentukan sentimen suatu teks. dalam data tweet, membagi teks menjadi kategori positif, negatif, dan netral. SVM bekerja dengan mendapati hyperplane optimal yang memisahkan kelas-kelas data. Model dilatih dan diuji dengan metrik seperti akurasi dan precision untuk mengevaluasi kinerjanya,memungkinkan analisis mendalam terhadap opini masyarakat mengenai CHAT GPT di Twitter.

g. *K-Fold Cross Validation*

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

Pada tahap validasi ini dilakukan validasi model k-fold validastion. K-fold cross-validation merupakan cara mengevaluasi model yang membagi dataset menjadi k sub-set yang sama besar, lalu secara bergantian menggunakan satu sub-set sebagai data uji dan sisanya sebagai data latih.

3. HASIL DAN PEMBAHASAN

3.1 Pembahasan

a. Tahap Pengumpulan data
Pengumpulan data dikerjakan melalui crawl data memakai bahasa pemrograman Python. Langkah pertama adalah menginstal pustaka Pandas di platform Google Colab dengan menjalankan perintah '!pip install pandas'. Proses pengambilan data kemudian dilakukan dengan menyesuaikan kata kunci dan rentang waktu sesuai kebutuhan penelitian. Dalam penelitian ini, kami berhasil mengumpulkan sebanyak 4.287 data dengan menggunakan kata kunci 'CHAT GPT' sebagai fokus pencarian dan jangka waktu pengumpulan data adalah dari 01-01-2023 sampai 30-09-2023.

b. Tahap *Preprocessing*
1. *Cleaning*
Pada tahap preprocessing, dilakukan pembersihan teks dengan menghilangkan karakter non-alfanumerik, termasuk tanda baca dan simbol-simbol khusus, untuk menghasilkan corpus teks yang lebih murni [15]. Hasil dari proses ini pada data tweet ditampilkan dalam Gambar 2 berikut:

	full_text	cleaning
0	Gila betul chatGPT ni 😵	Gila betul chatGPT ni 😵
1	Microsoft tambah fitur chatgpt openai #faktas...	Microsoft tambah fitur chatgpt openai
2	ChatGpt mengingatkan kepada simi-simi	ChatGpt mengingatkan kepada simi simi
3	@maeksfantasy Ngeplot sama human aja mikir ker...	Ngeplot sama human aja mikir keras confessnya ...
4	chatgpt ngelag bgt asuu	chatgpt ngelag bgt asuu
...	...	...
4300	bahas jawaban chat gpt di breakout room: https...	bahas jawaban chat gpt di breakout room
4301	@selfolrps Yang upchar gampang kok bedainnya. ...	Yang upchar gampang kok bedainnya Tim chatgpt ...
4302	@Nathan_sslvr chat gpt ajuda?	chat gpt ajuda
4303	@selfolrps WN. Aku karena dari awal join RP ma...	WN Aku karena dari awal join RP masuk di ENG j...
4304	Yang mau cek turnitin akun turnitin unlock doc...	Yang mau cek turnitin akun turnitin unlock doc...

Gambar 2. Hasil cleaning

2. *Case folding*
Case folding adalah Merubah teks menjadi huruf kecil (*lowercase*) bertujuan untuk menghapus perbedaan yang disebabkan oleh variasi penulisan huruf besar dan kecil dalam dokumen teks [9]**.** Hasil dari proses ini pada data tweet ditampilkan dalam Gambar 3 berikut :

cleaning	case_folding
Gila betul chatGPT ni 😵	gila betul chatgpt ni 😵
Microsoft tambah fitur chatgpt openai	microsoft tambah fitur chatgpt openai
ChatGpt mengingatkan kepada simi simi	chatgpt mengingatkan kepada simi simi
Ngeplot sama human aja mikir keras confessnya ...	ngeplot sama human aja mikir keras confessnya ...
chatgpt ngelag bgt asuu	chatgpt ngelag bgt asuu
...	...
bahas jawaban chat gpt di breakout room	bahas jawaban chat gpt di breakout room
Yang upchar gampang kok bedainnya Tim chatgpt ...	yang upchar gampang kok bedainnya tim chatgpt ...
chat gpt ajuda	chat gpt ajuda
WN Aku karena dari awal join RP masuk di ENG j...	wn aku karena dari awal join rp masuk di eng j...
Yang mau cek turnitin akun turnitin unlock doc...	yang mau cek turnitin akun turnitin unlock doc...

Gambar 3. Hasil case folding

3. Tokenisasi
Tokenisasi merupakan teknik dalam pemrosesan bahasa alami yang digunakan untuk memisahkan teks menjadi unit-unit terkecil yang bermakna, yaitu token. *Tokenizing* dilakukan dengan memisahkan kata-kata berdasarkan spasi [16]. Hasil dari proses ini pada data tweet ditampilkan dalam Gambar 4 berikut:

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

case_folding	tokenisasi
gila betul chatgpt ni 😵	[gila, betul, chatgpt, ni, 😵]
microsoft tambah fitur chatgpt openai	[microsoft, tambah, fitur, chatgpt, openai]
chatgpt mengingatkan kepada simi simi	[chatgpt, mengingatkan, kepada, simi, simi]
ngeplot sama human aja mikir keras confessnya ...	[ngeplot, sama, human, aja, mikir, keras, conf...
chatgpt ngelag bgt asuu	[chatgpt, ngelag, bgt, asuu]
...	...
bahas jawaban chat gpt di breakout room	[bahas, jawaban, chat, gpt, di, breakout, room]
yang upchar gampang kok bedainnya tim chatgpt ...	[yang, upchar, gampang, kok, bedainnya, tim, c...
chat gpt ajuda	[chat, gpt, ajuda]
wn aku karena dari awal join rp masuk di eng j...	[wn, aku, karena, dari, awal, join, rp, masuk,...
yang mau cek turnitin akun turnitin unlock doc...	[yang, mau, cek, turnitin, akun, turnitin, unl...

Gambar 4. Hasil Tokenisasi

4. *Stopword Removal*
   Stopword removal merupakan teknik pra-pemrosesan teks yang bertujuan untuk menghilangkan kata-kata yang sering terlihat namun tidak memiliki dampak yang besar terhadap makna keseluruhan suatu teks. Kata-kata seperti "apakah", "ke", "luar", "biasanya" termasuk dalam kategori ini[4]. Hasil penerapan teknik ini pada data tweet yang telah kita kumpulkan adalah sebagai berikut:

tokenisasi	stopword
[gila, betul, chatgpt, ni, 😵]	[gila, chatgpt, ni, 😵]
[microsoft, tambah, fitur, chatgpt, openai]	[microsoft, fitur, chatgpt, openai]
[chatgpt, mengingatkan, kepada, simi, simi]	[chatgpt, simi, simi]
[ngeplot, sama, human, aja, mikir, keras, conf...	[ngeplot, human, aja, mikir, keras, confessnya...
[chatgpt, ngelag, bgt, asuu]	[chatgpt, ngelag, asuu]
...	...
[bahas, jawaban, chat, gpt, di, breakout, room]	[bahas, chat, gpt, breakout, room]
[yang, upchar, gampang, kok, bedainnya,	[upchar, gampang, bedainnya, tim, chatgpt,

Gambar 5. Hasil Stopword Removal

5. *Stemming*
   Stemming adalah cara untuk menyederhanakan kata dengan menghapus bagian-bagian yang tidak penting seperti kata sambung, kata ganti, dan kata depan [6]. Tahap ini dikerjakan dengan menghapus awalan atau akhiran dari kata tersebut. Dengan memanfaatkan library sastrawi di Python, kata-kata dalam teks diubah menjadi bentuk dasarnya melalui tahap stemming. Hasil dari proses ini pada data tweet ditampilkan dalam Gambar 6 berikut:

stopword	stemming
[gila, chatgpt, ni, 😵]	gila chatgpt ni
[microsoft, fitur, chatgpt, openai]	microsoft fitur chatgpt openai
[chatgpt, simi, simi]	chatgpt simi simi
[ngeplot, human, aja, mikir, keras, confessnya...	ngeplot human aja mikir keras confessnya ga di...
[chatgpt, ngelag, asuu]	chatgpt ngelag asuu
...	...
[bahas, chat, gpt, breakout, room]	bahas chat gpt breakout room
[upchar, gampang, bedainnya, tim, chatgpt, tet...	upchar gampang bedainnya tim chatgpt tetep
[chat, gpt, ajuda]	chat gpt ajuda
[wn, join, rp, masuk, eng, udah, kebiasa, bang...	wn join rp masuk eng udah biasa banget nulis m...
[cek, turnitin, akun, turnitin, unlock, docs, ...	cek turnitin akun turnitin unlock docs quillbo...

Gambar 6. Hasil Stemming

c. *Labeling*
   Pasca preprocessing, proses pelabelan dilakukan. Pelabelan ini memanfaatkan library TextBlob yang terintegrasi dengan bahasa pemrograman Python. TextBlob hanya bisa membaca data yang berbahasa Inggris, oleh karena itu data harus di translate ke bahasa Inggris terlebih dahulu. Hasil dari proses ini pada data tweet ditampilkan dalam Gambar 7 berikut :

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

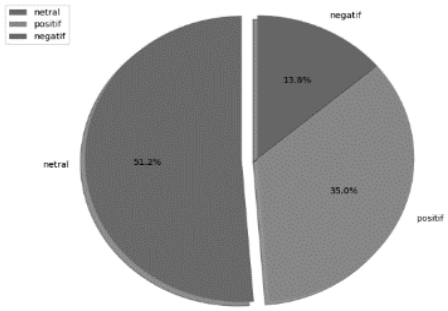
	stemming	English	polarity	TextBlob
0	gila chatgpt ni	This chatgpt is really crazy	-0.600000	negatif
1	microsoft fitur chatgpt openai	Microsoft adds openai chatgpt feature	0.000000	netral
2	chatgpt simi simi	chatgpt reminds simsimi	0.000000	netral
3	ngeplot human aja mikir keras confessnya ga di...	Just plotting with humans, thinking hard about...	-0.291667	negatif
4	chatgpt ngelag asuu	chatgpt is lagging really bad	-0.700000	negatif
...	...	...	...	...
4295	bahas chat gpt breakout room	discuss gpt chat answers in the breakout room	0.000000	netral
4296	upchar gampang bedainnya tim chatgpt tetep	the upchar is easy, the difference is the chat...	0.433333	positif
4297	chat gpt ajuda	chat gpt ajuda	0.000000	netral
4298	wn join rp masuk eng udah biasa banget nulis m...	wn me, because from the start of joining rp, I...	0.200000	positif
4299	cek turnitin akun turnitin unlock docs quillbo...	If you want to check your Turnitin account, Tu...	0.000000	netral

Gambar 7. Hasil labeling textblob

Setelah dilakukan proses *labeling* pada dataset yang berisi 4287 tweet, data tersebut berhasil dikelompokkan ke dalam tiga kategori sentimen utama. Kategori pertama adalah sentimen netral, yang mencakup 2196 tweet atau sekitar 51.22% dari total data. Tweet dalam kategori ini tidak menunjukkan emosi yang jelas, baik positif maupun negatif, sehingga dianggap netral. Kategori kedua adalah sentimen positif, dengan jumlah 1500 tweet, yang mewakili sekitar 35.00% dari keseluruhan dataset. Tweet dengan sentimen positif ini menunjukkan pandangan atau opini yang bersifat mendukung, optimis, atau menyenangkan terhadap topik yang dibahas. Terakhir, terdapat 591 tweet yang termasuk dalam kategori sentimen negatif, mewakili sekitar 13.78% dari total data, yang berisi pendapat atau pandangan kritis, kurang puas, atau bersifat negatif terhadap topik yang dianalisis. Pembagian ini menunjukkan bahwa mayoritas tweet dalam dataset cenderung netral, sementara sentimen positif lebih banyak dibandingkan sentimen negatif.

d. Visualisasi

Visualisasi data merupakan teknik yang digunakan untuk menyederhanakan data kompleks menjadi representasi visual yang lebih mudah dicerna, sekaligus meningkatkan daya tarik sehingga audiens lebih tertarik untuk menjelajahi informasi yang disajikan. Proses visualisasi data pada penelitian ini dilakukan menggunakan library matplotlib, menggunakan matplotlib, data akan divisualisasi menjadi bentuk pie chart. Hasil dari proses ini ditampilkan dalam Gambar 8 berikut:



Gambar 8. Hasil visualisasi data

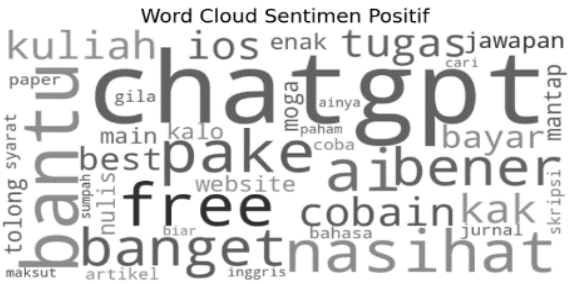
Gambar 8 menampilkan hasil visualisasi distribusi data sentimen dalam bentuk persentase. Visualisasi ini memudahkan untuk melihat seberapa besar proporsi masing-masing kategori sentimen dalam dataset yang dianalisis. Dari gambar tersebut, terlihat bahwa sentimen netral mendominasi dengan persentase sebesar 51,2%. Ini menunjukkan bahwa sebagian besar masyarakat memiliki pandangan yang cenderung netral terhadap topik yang sedang dibahas, dalam hal ini terkait dengan teknologi Chat GPT. Selain itu, sebanyak 35,0% dari data yang dianalisis menunjukkan sentimen positif. Hal ini mengindikasikan bahwa cukup banyak masyarakat yang melihat Chat GPT secara positif, mungkin karena manfaat dan kemajuan yang dihadirkan oleh teknologi ini. Di sisi lain, sentimen negatif hanya berjumlah 13,8%, menandakan bahwa hanya sebagian kecil dari masyarakat yang memandang Chat GPT secara kritis atau memiliki kekhawatiran tertentu terhadap penggunaannya.

Selain menggunakan visualisasi data berupa diagram batang atau pie chart yang dihasilkan melalui matplotlib, penelitian ini juga memanfaatkan word cloud sebagai salah satu cara untuk menggambarkan frekuensi kata-kata yang sering muncul pada kategori sentimen tertentu. Word cloud adalah teknik visualisasi yang menampilkan kata-kata dalam ukuran yang bervariasi sesuai dengan frekuensi kemunculannya dalam teks [17].

Pada Gambar 9, hasil visualisasi word cloud untuk sentimen positif ditampilkan. Word cloud ini memudahkan untuk melihat kata-kata yang paling sering muncul dalam tweet yang berisi sentimen positif terhadap Chat GPT. Kata-kata dengan ukuran font yang lebih besar menandakan frekuensi kemunculannya yang lebih tinggi, sehingga memberikan

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

gambaran sekilas mengenai aspek-aspek yang dianggap penting atau positif oleh masyarakat. Visualisasi ini sangat efektif dalam menyederhanakan data teks yang kompleks, sehingga membantu pembaca dengan cepat memahami tema dan fokus utama dari opini positif yang diungkapkan oleh masyarakat di media sosial Twitter.Kombinasi antara grafik presentasi dan word cloud memberikan dua sudut pandang yang saling melengkapi. Grafik memberikan gambaran kuantitatif mengenai distribusi sentimen, sementara word cloud memberikan wawasan kualitatif tentang konten spesifik yang sering muncul dalam setiap kategori sentimen.



Gambar 9. Word cloud Sentimen Positif

Berikut hasil visualisasi data menggunakan *word cloud* matplotlib sentimen netral dapat dilihat pada Gambar 10 berikut ini :

Gambar 10. Word cloud Sentimen Netral

Berikut hasil visualisasi data menggunakan *word cloud* matplotlib sentimen negatif dapat dilihat pada Gambar 11 berikut ini :

Gambar 11. Word cloud Sentimen Negatif

e. Pembagian data

Langkah awal menggunakan SVM adalah memisahkan data menjadi dua bagian. Bagian pertama untuk melatih model (data pelatihan), dan bagian kedua untuk menguji kinerja model (data pengujian). dalam penelitian ini, kami membandingkan tiga proporsi pembagian yang berbeda antara data pelatihan dan data pengujian.

Tabel 1. Hasil pembagian data *Training* 60% dan data *Testing* 40%

Jenis data	Presentase	Jumlah
Data *Training*	60%	2572
Data *Testing*	40%	1715
Jumlah	100%	4287

Tabel 1 menunjukkan hasil pembagian data dengan proporsi 60% untuk data Training dan 40% untuk data Testing. Dari total 4287 data yang ada, sebanyak 2572 data (60%) dialokasikan untuk data Training, sedangkan 1715 data (40%) digunakan untuk data Testing. Pembagian ini memberikan kesempatan bagi model untuk belajar dari data yang lebih besar (data Training), sehingga diharapkan dapat meningkatkan kemampuan model dalam mengenali pola dan karakteristik data. Dengan jumlah data Testing yang cukup signifikan, model juga dapat diuji secara efektif untuk mengukur kinerjanya dalam mengklasifikasikan data yang belum pernah dilihat sebelumnya. Pembagian ini

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

merupakan langkah penting dalam proses pengembangan model, memastikan bahwa ada keseimbangan antara data yang digunakan untuk melatih model dan data yang digunakan untuk evaluasi.

Tabel 2. Hasil pembagian Data *Training* 70% dan data *Testing* 30%

Jenis data	Presentase	Jumlah
Data *Training*	70%	3000
Data *Testing*	30%	1287
Jumlah	100%	4287

Tabel 2 menunjukkan hasil pembagian data dengan proporsi 70% untuk data Training dan 30% untuk data Testing. Dalam tabel ini, 3000 data (70%) dialokasikan untuk data Training, sedangkan 1287 data (30%) digunakan untuk data Testing. Dengan menggunakan proporsi yang lebih besar untuk data Training dibandingkan dengan Tabel 1, model memiliki lebih banyak data untuk mempelajari pola yang ada. Ini diharapkan dapat meningkatkan akurasi dan performa model dalam mengklasifikasikan data Testing. Jumlah data Testing yang masih cukup banyak (1287) memastikan bahwa evaluasi model tetap dapat dilakukan secara menyeluruh, memberikan gambaran yang jelas mengenai kemampuannya dalam mengenali pola di luar data Training.

Tabel 3. Hasil pembagian Data *Training* 80% dan data *Testing* 20%

Jenis data	Presentase	**Jumlah**
Data *Training*	80%	3429
Data *Testing*	20%	858
Jumlah	100%	4287

Tabel 3 menunjukkan hasil pembagian data dengan proporsi 80% untuk data Training dan 20% untuk data Testing. Dalam tabel ini, sebanyak 3429 data (80%) dialokasikan untuk data Training, sementara 858 data (20%) digunakan untuk data Testing. Dengan proporsi 80% untuk data Training, model memiliki kesempatan terbaik untuk belajar dari data yang lebih banyak, yang dapat meningkatkan kemampuannya dalam mengenali pola dan karakteristik dalam dataset. Meskipun jumlah data Testing lebih kecil dibandingkan dengan tabel sebelumnya, yakni 858, model masih diuji dengan cukup efektif untuk memastikan kinerjanya. Tabel ini menunjukkan strategi yang berfokus pada optimalisasi data Training untuk meningkatkan akurasi model, dengan harapan dapat menghasilkan performa yang tinggi dalam klasifikasi sentimen.

Ketiga tabel tersebut menggambarkan berbagai pembagian data Training dan Testing yang berbeda dalam proporsi 60:40, 70:30, dan 80:20. Setiap pembagian memiliki tujuan untuk memberikan pelatihan yang cukup bagi model SVM serta memungkinkan evaluasi kinerjanya yang tepat. Variasi dalam pembagian data ini membantu dalam mengevaluasi bagaimana perubahan dalam proporsi data dapat memengaruhi akurasi dan kinerja model dalam analisis sentimen.

f. Klasifikasi dan evaluasi menggunakan metode *Support Vector Machine*

Support Vector Machine akan membangun model klasifikasi dengan mengidentifikasi fitur-fitur pembeda pada setiap kelas dalam data pelatihan. Model yang terbentuk kemudian dievaluasi performansinya menggunakan data pengujian. Evaluasi dikerjakan dengan memanfaatkan confusion matrix untuk memperoleh matrik-matrik seperti nilai *accuracy*, *presisi*, *false positif rate*, *sensitvitas dan spesifisitas*. Adapun hasil dari proses klasifikasi data *Training* dan data *Testing* menggunakan metode *SVM* dapat dilihat pada Gambar 12 hingga Gambar 17 berikut:

1. Data *Training*

```
Akurasi untuk metode SVM : 0.9405


Confusion Matrix
[[ 275   58   22]
 [   0 1313    4]
 [   0   69  831]]

              precision    recall  f1-score   support

     negatif       1.00      0.77      0.87       355
      netral       0.91      1.00      0.95      1317
     positif       0.97      0.92      0.95       900

    accuracy                           0.94      2572
   macro avg       0.96      0.90      0.92      2572
weighted avg       0.94      0.94      0.94      2572

Negatif: Specificity = 1.0000, FPR = 0.0000
Netral: Specificity = 0.8988, FPR = 0.1012
Positif: Specificity = 0.9844, FPR = 0.0156
```

Gambar 12. Hasil klasifikasi metode SVM untuk perbandingan 60:40

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

```
Akurasi untuk metode SVM : 0.9407


Confusion Matrix
[[ 323   66   24]
 [   1 1531    5]
 [   0   82  968]]

              precision    recall  f1-score   support

     negatif       1.00      0.78      0.88       413
      netral       0.91      1.00      0.95      1537
     positif       0.97      0.92      0.95      1050

    accuracy                           0.94      3000
   macro avg       0.96      0.90      0.92      3000
weighted avg       0.94      0.94      0.94      3000

Negatif: Specificity = 0.9996, FPR = 0.0004
Netral: Specificity = 0.8988, FPR = 0.1012
Positif: Specificity = 0.9851, FPR = 0.0149
```

Gambar 13. Hasil klasifikasi metode SVM untuk perbandingan 70:30

```
Akurasi untuk metode SVM : 0.9425


Confusion Matrix
[[ 379   70   24]
 [   1 1751    4]
 [   0   98 1102]]

              precision    recall  f1-score   support

     negatif       1.00      0.80      0.89       473
      netral       0.91      1.00      0.95      1756
     positif       0.98      0.92      0.95      1200

    accuracy                           0.94      3429
   macro avg       0.96      0.91      0.93      3429
weighted avg       0.95      0.94      0.94      3429

Negatif: Specificity = 0.9997, FPR = 0.0003
Netral: Specificity = 0.8996, FPR = 0.1004
Positif: Specificity = 0.9874, FPR = 0.0126
```

Gambar 14. Hasil klasifikasi metode SVM untuk perbandingan 80:30

2. Data *Testing*

```
Akurasi untuk metode SVM : 0.9306


Confusion Matrix
[[186  38  12]
 [  0 878   1]
 [  0  68 532]]

              precision    recall  f1-score   support

     negatif       1.00      0.79      0.88       236
      netral       0.89      1.00      0.94       879
     positif       0.98      0.89      0.93       600

    accuracy                           0.93      1715
   macro avg       0.96      0.89      0.92      1715
weighted avg       0.94      0.93      0.93      1715

Negatif: Specificity = 1.0000, FPR = 0.0000
Netral: Specificity = 0.8732, FPR = 0.1268
Positif: Specificity = 0.9883, FPR = 0.0117
```

Gambar 15. Hasil klasifikasi metode SVM untuk perbandingan 60:40

```
Akurasi untuk metode SVM : 0.9316


Confusion Matrix
[[139  31   8]
 [  0 659   0]
 [  0  49 401]]

              precision    recall  f1-score   support

     negatif       1.00      0.78      0.88       178
      netral       0.89      1.00      0.94       659
     positif       0.98      0.89      0.93       450

    accuracy                           0.93      1287
   macro avg       0.96      0.89      0.92      1287
weighted avg       0.94      0.93      0.93      1287

Negatif: Specificity = 1.0000, FPR = 0.0000
Netral: Specificity = 0.8726, FPR = 0.1274
Positif: Specificity = 0.9904, FPR = 0.0096
```

Gambar 16. Hasil klasifikasi metode SVM untuk perbandingan 70:30

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 2, February 2025 | Hal 138-149
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i2.475

```
Akurasi untuk metode SVM : 0.9207

Confusion Matrix
[[ 84  25   9]
 [  0 440   0]
 [  0  34 266]]

              precision    recall  f1-score   support

     negatif       1.00      0.71      0.83       118
      netral       0.88      1.00      0.94       440
     positif       0.97      0.89      0.93       300

    accuracy                           0.92       858
   macro avg       0.95      0.87      0.90       858
weighted avg       0.93      0.92      0.92       858

Negatif: Specificity = 1.0000, FPR = 0.0000
Netral: Specificity = 0.8589, FPR = 0.1411
Positif: Specificity = 0.9839, FPR = 0.0161
```

Gambar 17. Hasil klasifikasi metode SVM untuk perbandingan 80:20

Dari Gambar 12 sampai Gambar 17, terlihat bahwa hasil pengujian untuk data Training dengan perbandingan 80% untuk training dan 20% untuk testing menunjukkan tingkat akurasi tertinggi sebesar 94,25%. Hal ini menunjukkan bahwa model SVM yang diterapkan sangat efektif dalam mempelajari pola-pola yang ada dalam dataset ketika 80% dari data digunakan untuk pelatihan. Dengan proporsi yang lebih besar ini, model memiliki lebih banyak informasi untuk menganalisis dan memahami karakteristik data, yang berkontribusi pada kemampuan prediktifnya yang kuat. Akurasi yang tinggi ini menunjukkan bahwa model tidak hanya mampu mengenali pola dalam data training, tetapi juga berhasil menggeneralisasi informasi tersebut ketika digunakan untuk mengklasifikasikan data baru.Di sisi lain, hasil pengujian untuk data Testing dengan perbandingan 70% untuk training dan 30% untuk testing mencatat tingkat akurasi tertinggi sebesar 93,16%. Meskipun akurasinya sedikit lebih rendah dibandingkan dengan pengujian data Training pada 80%, nilai 93,16% masih menunjukkan kinerja yang baik. Ini berarti bahwa model tetap mampu melakukan prediksi yang akurat terhadap data yang belum pernah dilihat sebelumnya. Dengan 30% dari dataset digunakan untuk pengujian, model diuji dalam kondisi yang lebih ketat karena memiliki lebih sedikit data untuk dipelajari. Meskipun ada penurunan akurasi, hasil ini menandakan bahwa model SVM tetap memiliki kemampuan yang solid dalam mengklasifikasikan data testing, menunjukkan bahwa model ini dapat diandalkan dalam analisis sentimen.Secara keseluruhan, perbandingan ini menunjukkan bahwa proporsi data training yang lebih besar dapat meningkatkan akurasi model, tetapi model tetap menunjukkan performa yang baik meskipun dalam pengujian dengan proporsi yang berbeda. Hal ini mengindikasikan bahwa model SVM yang digunakan dalam penelitian ini memiliki potensi untuk memberikan hasil yang akurat dalam klasifikasi sentimen masyarakat terhadap teknologi, terlepas dari variasi dalam pembagian data.

4. KESIMPULAN

Pengumpulan data dilakukan menggunakan metode crawl data, dengan kata kunci pencarian “CHAT GPT” dari periode bulan Januari hingga September 2023. Dari hasil crawling, terkumpul sebanyak 4305 data tweet yang relevan dengan topik Chat GPT. Setelah dilakukan proses labeling, data tersebut terbagi menjadi tiga kategori, yaitu 2196 tweet dengan sentimen netral, 1500 tweet dengan sentimen positif, dan 591 tweet dengan sentimen negatif. Proses analisis dilakukan menggunakan algoritma Support Vector Machine (SVM) dengan membandingkan beberapa rasio data *training* dan *testing*. Data *training* perbandingan 60:40 menghasilkan nilai akurasi sebesar 94,05% dan untuk perbandingan 70:30 menghasilkan nilai akurasi sebesar 94,07% serta pada perbandingan data *training* 80:20, hasil menunjukkan akurasi tertinggi sebesar 94,25% untuk data training. Untuk Data *testing* perbandingan 60:40 menghasilkan nilai akurasi sebesar 93,06% dan untuk perbandingan 70:30 menghasilkan nilai akurasi sebesar 93,13% serta pada perbandingan data *testing* 80:20, hasil menunjukkan akurasi tertinggi sebesar 92,07%. pada rasio 70:30, akurasi tertinggi tercatat sebesar 93,16% untuk data testing. Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa pandangan masyarakat terhadap kemajuan teknologi OpenAI Chat GPT cenderung netral, di mana banyak yang mengakui manfaat positifnya, seperti meningkatkan produktivitas dan efisiensi, tetapi juga memahami potensi dampak negatifnya, seperti risiko penyalahgunaan informasi. Masyarakat Indonesia dinilai kritis dalam menyikapi perkembangan teknologi ini, dengan tetap berhati-hati terhadap risiko yang mungkin timbul di masa depan. Secara umum, teknologi ini diterima dengan baik namun tetap dipandang secara hati-hati.

REFERENCES

[1] A. Erfina, “Implementation of Naive Bayes classification algorithm for Twitter user sentiment analysis on ChatGPT using Python programming language,” Data & Metadata, pp. 2–11, 2023, doi: 10.56294/dm202345.
[2] R. Kajayaan, T. Putra, F. R. Saputro, L. Hakim, and Y. Ramadhan, “Fenomena ChatGPT : Peningkatkan civic skill digital native generation,” Nautical, vol. 2, no. 2, pp. 140–147, 2023.
[3] A. Bin, M. Alkatiri, and A. N. S. Nasution, “Opini Publik Terhadap Penerapan New Normal Di Media Sosial Twitter,” CoverAge, vol. 11, no. 1, 2020.
[4] M. Muadin, and H. Asnal, “Implementasi Metode Support Vector Machine Pada Opinion,” JOISIE, vol. 7, no. 1, 2023.
[5] A. Witanti, “Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma

Support Vector Machine (Svm),” Simika, vol. 5, no. 1, pp. 59–67, 2022.
[6] P. Arsi and R. Waluyo, “Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
[7] M. I. Petiwi, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine,” MIB, vol. 6, pp. 542–550, 2022, doi: 10.30865/mib.v6i1.3530.
[8] M. R. Fauzan, H. O. L. Wijaya, S. Satrianansyah, and J. Karman, “Analisis Sentimen Masyarakat Terhadap Kenaikan Harga Bbm Di Media Sosial Twitter Menggunakan Metode Support Vector Machine,” in *Prosiding Seminar Riset Mahasiswa*, 2023.
[9] D. Atika, “Term Frequency-Inverse Document Frequency Support Vector Machine Untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Mental Pada Media Sosial Twitter,” Jurnal Teknologi dan Sistem Informasi, vol. 3, no. 4, 2022.
[10] A. Alwasi’a, “Analisis Sentimen Pada Review Aplikasi Berita Online Menggunakan Metode Maximum Entropy (Studi Kasus: Review Detikcom pada Google Play 2019),” *Skripsi*, 2020.
[11] M. Rahman Fauzan, H. Oktafia Lingga Wijaya, and J. Karman, “Analisis Sentimen Masyarakat Terhadap Kenaikan Harga Bbm Di Media Sosial Twitter Menggunakan Metode Support Vector Machine,” *Semin. Ris. Mahasiswa-Computer Electr. (SERIMA-CE*, vol. 1, no. 1, p. 82, 2023.
[12] M. H. Rasyadi “Analisis sentimen pada twitter menggunakan metode naïve bayes (studi kasus pemilihan gubernur dki jakarta 2017),” 2017.
[13] R. Vindua and A. U. Zailani, “Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 10, no. 2, p. 479, 2023, doi: 10.30865/jurikom.v10i2.5945.
[14] O. I. Gifari, M. Adha, I. R. Hendrawan, F. Freddy, and S. Durrand, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” JifoTech, vol. 2, no. 1, pp. 36–40, 2022.
[15] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, “Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine,” *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021.
[16] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, “Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm,” *J. Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi: 10.14710/j.gauss.v9i3.28932.
[17] A. W. V. Hutabarat, N. L. S. S. Adnyani, and K. Suryadi, “Analisis Sentimen Data Ulasan Pengguna MyPertamina di Twitter dengan Metode Text Mining,” *J. Rekayasa Sist. Ind.*, vol. 13, no. 1, pp. 145–154, 2024, doi: 10.26593/jrsi.v13i1.6958.145-154.