

ANALISIS SENTIMEN DATA TWITTER TERKAIT CHATGPT MENGGUNAKAN ORANGE DATA MINING

Tri Yuli Pahtoni*¹, Handaru Jati²

^{1,2}Universitas Negeri Yogyakarta, Yogyakarta
Email: ¹triyuli.2022@sudent.uny.ac.id, ²handaru@uny.ac.id

*Penulis Korespondensi

(Naskah masuk: 17 Mei 2023, diterima untuk diterbitkan: 04 April 2024)

Abstrak

Perkembangan teknologi bergerak begitu cepat, diikuti dengan popularitas media sosial yang semakin meluas. Platform media sosial mampu membangun profil *big-data* pengguna, dengan melacak setiap aktivitas seperti partisipasi, pengiriman pesan, dan kunjungan situs Web. Saat ini banyak orang sering membagikan kritik terhadap sesuatu melalui platform media sosial seperti Facebook, Twitter, Instagram, dan lainnya. Sehingga perlu diketahui bagaimana komentar dari pengguna media sosial yang menghasilkan reaksi masyarakat terhadap chatGPT yang dirilis oleh OpenAI. Banyaknya komentar di Twitter menyebabkan sulitnya mengetahui kecenderungan respon masyarakat. Tujuan dari penelitian ini yaitu melakukan analisis sentimen postingan publik di Twitter untuk memberikan wawasan tentang sikap dan persepsi orang tentang suatu peristiwa. Penelitian ini memberikan ilustrasi peran Twitter dalam menampung postingan pengguna Twitter terkait chatGPT. Hasil penelitian ini dapat digunakan oleh pemangku kepentingan untuk menentukan kebijakan dalam penggunaan chatGPT. Penelitian ini menganalisis sebanyak 5.192 postingan tweet bahasa Inggris dan 641 tweet bahasa Indonesia, mulai dari tanggal 27 April hingga 8 Mei 2023. Tanggapan positif, negatif, dan netral diolah menggunakan perangkat lunak Orange Data Mining dengan algoritma Ekman untuk menganalisis emosi data tweet yang diperoleh. Hasil menunjukkan bahwa chatGPT mendapatkan tanggapan netral berbahasa Inggris dengan nilai sebesar 54,72%, tanggapan positif sebesar 31,64%, dan tanggapan negatif sebesar 13,64%. Hasil analisis sentimen berbahasa Indonesia tidak jauh berbeda, dengan nilai tanggapan netral sebesar 63,96%, tanggapan positif 23,56%, dan tanggapan negatif 12,48%. Sehingga dapat disimpulkan bahwa, rilisnya chatGPT mayoritas publik memberikan tanggapan netral atau tidak terdapat penolakan.

Kata kunci: algoritma ekman, chatGPT, orange data mining, sentimen, text mining, twitter

SENTIMENT ANALYSIS OF TWITTER DATA REGARDING CHATGPT USING ORANGE DATA MINING

Abstract

Technological developments move so fast, followed by the increasingly widespread popularity of social media. Social media platforms can build big-data profiles of users by tracking every activity such as participation, messaging, and website visits. Currently, many people often share criticism of something through social media platforms, such as Facebook, Twitter, Instagram, and others. So it is necessary to know how comments from social media users generate public reactions to chatGPT released by OpenAI. A lot of comments on Twitter make it difficult to know the trend of people's responses. The purpose of this research is to conduct sentiment analysis of public posts on Twitter to provide insights into people's attitudes and perceptions regarding a particular event. This research illustrates Twitter's role in accommodating Twitter user posts regarding chatGPT. The results of this study can be used by stakeholders in making policies on the use of chatGPT. This study analyzed 5,192 posts in English and 641 tweets in Indonesian from April 27 to May 8, 2023. The positive, negative, and neutral responses are processed using Orange Data Mining software with the Ekman algorithm to analyze the emotional content of the acquired tweet data. The results show that chatGPT received neutral responses in English with a value of 54.72%, positive responses of 31.64%, and negative responses of 13.64%. The results of sentiment analysis in Indonesian were not much different, with neutral responses of 63.96%, positive responses of 23.56%, and negative responses of 12.48%. So it can be concluded that after the release of chatGPT, the majority of the public gave neutral responses or no rejection.

Keywords: chatGPT, ekman algorithm, orange data mining, sentiment, text mining, twitter

1. PENDAHULUAN

Pertumbuhan media sosial yang sangat cepat dalam beberapa tahun terakhir telah menjadikan internet sebagai *platform* yang layak untuk memberikan sebuah pengetahuan. Penggunaan *big-data* dari media sosial dapat digunakan untuk prediksi kejadian di dunia nyata (Sekioka, Hatano and Niahiyama, 2023). *Platform* media sosial mampu membangun profil *big-data* pengguna dengan melacak setiap aktivitas seperti partisipasi, pengiriman pesan, dan kunjungan situs Web. Pendapat tentang suatu keadaan tersebut dapat diambil menggunakan teknik teks mining agar dapat menarik kesimpulan tentang keadaan emosional individu.

Diantara media sosial yang sangat beragam, twitter menjadi salah satu jejaring sosial yang paling banyak dipelajari untuk penelitian analisis (Aminimotlagh, Sakhoseini dan Fatehi, 2023). Twitter memungkinkan pengguna untuk membuat jejaring sosial yang luas dalam menyebarkan informasi dan komunikasi (Boukes, 2019). Twitter menjadi jejaring sosial yang memungkinkan pengguna untuk berbagi emosi dan pendapat sehari-hari, dan dianggap sebagai *platform* yang nyaman bagi pengguna untuk berbagi pesan, gambar, dan video pribadi. Hal ini menjadikan Twitter sebagai sumber referensi bagi ilmuwan sosial dan akademisi yang tertarik untuk mempelajari nuansa emosi manusia dari berbagai sudut (Öztürk dan Ayvaz, 2018). Hal ini dapat dilakukan melalui analisis sentimen.

Analisis sentimen merupakan suatu istilah yang biasa digunakan untuk memberikan gambaran topik secara faktual atau non faktual dengan opini positif ataupun negatif (Wongkar dan Angdresey, 2019). Analisis ini menggunakan suatu pendekatan dalam menganalisis teks untuk mengetahui subjektivitas pendapat/opini postingan tweet (Agustina and Ihsan, 2023). Karena masyarakat cenderung banyak memberi opini berupa tanggapan atau kritik terhadap suatu situasi melalui media sosial (Tri Saputra et al., 2021). Analisis sentimen memungkinkan untuk membaca dan memahami emosi orang dan memberikan kesimpulan berdasarkan hasil deteksi bahasa postingan tweet. Banyak tweet yang dibuat oleh orang-orang di seluruh dunia sebagai tanggapan atas peristiwa tertentu, salah satunya mengenai *artificial intelligence*.

Artificial intelligence adalah istilah umum yang mengacu pada penciptaan sistem cerdas yang dapat melakukan tugas-tugas dan membutuhkan kecerdasan manusia (Holzinger et al., 2023), dapat melakukan tugas-tugas yang kaitan dengan kecerdasan manusia, seperti belajar, berpikir, memahami bahasa, dan menyelesaikan masalah. *Artificial Intelligence* (AI) memainkan peran penting dalam peningkatan penggunaan digitalisasi dalam kehidupan sosial (Cooper, 2023). Kemampuan AI untuk mengotomatiskan tugas, memproses data

dalam jumlah besar, dan memberikan wawasan prediktif akan semakin merevolusi berbagai aspek kehidupan kita sehari-hari (Yang, 2022). ChatGPT menjadi salah satu produk yang dirilis oleh OpenAI (Tong dan Zhang, 2023).

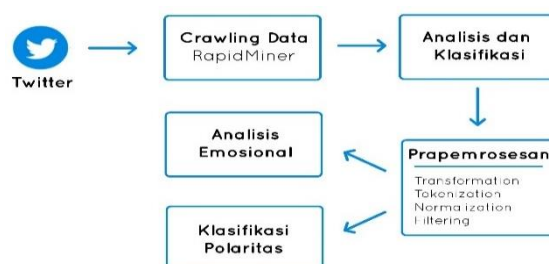
ChatGPT merupakan model bahasa yang dilatih untuk dapat memahami dan menghasilkan teks dalam berbagai bahasa. ChatGPT dapat menulis cerita, memberikan sebuah nasihat, bahkan menulis puisi dan membuat kode program komputer (Schart, 2022). OpenAI kedepan memungkinkan akan merevolusi banyak bidang termasuk bidang ilmiah (Aghemo, Forner dan Valenti, 2023). Secara teori, ulasan deskriptif, revisi naskah yang dikirimkan sebagian besar dapat disusun oleh ChatGPT, dengan sedikit atau tanpa input/kontrol oleh pakar atau ilmuwan.

Pengenalan ChatGPT pada November 2022 yang tersedia secara gratis, telah menimbulkan kekhawatiran di antara beberapa pendidik tentang kelayakan dan manfaat mengajarkan keterampilan menulis dasar dan berpikir kritis kepada siswa dalam konteks kecurangan, dengan mudahnya mekanisme akses berbasis AI (Karaali, 2023). Kekhawatiran tersebut menyebabkan publik menyampaikan opini melalui media sosial.

Penelitian ini bertujuan untuk mempelajari lebih lanjut pandangan publik mengenai chatGPT. Kategorisasi tweet dilakukan menggunakan teknologi *deep learning*. Fokus utama penelitian ini untuk menguji sentimen publik terhadap rilisnya chatGPT melalui *platform* media sosial twitter. Kontribusi dalam penelitian ini mengusulkan metode yang efektif dalam menganalisis sentimen publik melalui twitter menggunakan orange data mining. Metode yang diusulkan terbukti valid dan dapat digunakan setelah melalui percobaan beberapa kali menggunakan data tweet.

2. METODE PENELITIAN

Penelitian ini dilaksanakan dengan beberapa tahapan, di setiap tahapan memiliki tujuan dalam menghasilkan luaran yang dibutuhkan. Tahapan tersebut mencakup: i) crawling data, ii) analisis dan klasifikasi konten tweet, iii) prapemrosesan, iv) analisis emosional, v) klasifikasi polaritas. Metode dapat dilihat pada gambar 1.



Gambar 1. Metode Penelitian

2.1 Crawling Data

Tahap pertama yang dilakukan yaitu mengambil tweet dari API Twitter. API Twitter bisa didapatkan dengan membuat akun terlebih dahulu pada Twitter Developer. *Consumer key* dan *access token* akan tersedia setelah pengguna berhasil mendaftar akun pada Twitter *developer*. Pastikan *consumer key* dan *access token* yang didapatkan tidak diketahui oleh pengguna lain. Tahap ini dilakukan menggunakan RapidMiner. Rapidminer adalah aplikasi *open-source* untuk pengambilan data yang dapat digunakan sebagai kerangka kerja yang berdiri sendiri dalam analisis data (Ghous and Kovacs, 2020). RapidMiner menjadi salah satu aplikasi yang sangat baik untuk integrasi data, *analytical extract transform*, dan analisis data.

2.2 Analisis dan Klasifikasi Konten Tweet

Tahap analisis dan klasifikasi konten tweet bertujuan untuk menyaring dan mengelompokkan tweet berdasarkan karakter sentimen postingan tweet dan berdasarkan bahasa. Dalam memahami karakter sentimen tweet dan menentukan format tweet (yaitu teks, gambar, atau video), dalam sebuah penelitian dapat menggunakan pendekatan naturalistik (Bashir et al., 2021). Pendekatan naturalistik mengarah pada proses analisis konten tweet yang dipilih secara random dengan jumlah tertentu.

Pengelompokan konten tweet berdasarkan bahasa dilakukan secara manual. Dalam penelitian ini terdapat dua kelompok bahasa yang akan dianalisis, yaitu bahasa Indonesia dan bahasa Inggris.

2.3 Prapemrosesan

Tahap ini dilakukan sebelum hasil *text mining* dianalisis. Prapemrosesan bertujuan untuk membuat huruf kapital menjadi kecil, menghapus tanda baca, menghapus *mention* (@), *hashtag* (#), *retweet* atau kata yang tidak dibutuhkan. Dalam tugas sebuah *machine learning*, pembersihan dan prapemrosesan data sama pentingnya dalam membuat suatu model.

Bagian-bagian yang dilaksanakan pada tahap ini mencakup: *i) Transformation*, *ii) Tokenization*, *iii) Normalization*, dan *iv) Filtering* (T. Elshoush and A. Dinar, 2019). *Orange data mining* sudah menyediakan widget *Preprocess Text* dalam menjalankan tugas tersebut.

2.4 Analisis Emosional

Emosional merupakan keadaan psikologi kompleks yang mencakup suatu pengalaman subjektif, respon fisiologis, dan respon perilaku. Hal ini dapat terlihat pada ekspresi wajah seseorang. Terdapat enam kategori emosi yang sering muncul dalam mengkarakterisasi emosi dasar seseorang, kemarahan, penghinaan, ketakutan, kebahagiaan, kesedihan, dan terkejut. Terkejut menjadi ekspresi yang paling ambigu, karena dapat memberikan arti emosi positif dan negatif (Pang dan Lee, 2008).

Tahap ini dapat dilakukan pada perangkat orange data mining menggunakan widget profiler tweet, algoritma yang digunakan Ekman (multiclass) untuk menganalisis emosi dari tweet yang telah didapatkan. Kemudian dihubungkan dengan *widget select columns* untuk memilih kolom emosi yang selanjutnya terhubung dengan *widget distributions* dengan *output* diagram hasil analisis emosi.

2.5 Klasifikasi polaritas

Polaritas menjadi bagian yang menentukan kondisi emosional, polaritas sentimen yang mengklasifikasikan tweet sebagai opini positif dan negatif (Imran et al., 2020). Polaritas sering dikenal dengan emosi yang direpresentasikan dalam sebuah pernyataan. Pada tahap ini mencakup polaritas yang dikategorikan sebagai tanggapan positif, negatif, dan netral.

Orange data mining telah menyediakan enam jenis alat untuk analisis sentimen yang dikemas dalam *widget Sentiment Analysis*. Alat tersebut yaitu *valence aware dictionary* dan *sentiment reasoner* (VADER), *Liu-hu*, *LilaH*, *Multilingual*, *custom dictionary* dan *SentiArt*. *Liu-hu* dan *Multilingual* merupakan alat yang digunakan dalam penelitian ini.

Metode *Liu-Hu* melibatkan pemeriksaan leksikon (Hu dan Liu, 2004). Proses berbasis leksikon adalah metode pembelajaran mesin tanpa pengawasan yang menggunakan kamus atau daftar leksikon (Abayomi-Alli et al., 2022). Setiap leksikon dikaitkan dengan kekuatan sentimen yang mewakili orientasi positif atau negatif (Labille, Gauch dan Alfarhood, 2017). Daftar leksikon merupakan sebuah daftar atau kumpulan kata-kata yang ada dalam sebuah kamus atau buku referensi yang kerap digunakan sebagai sumber informasi atau panduan dalam mengartikan kata-kata atau istilah-istilah tertentu.

Multilingual sentiment merupakan analisis sentimen yang dilakukan terhadap teks yang ditulis dalam beberapa jenis bahasa yang berbeda. *Multilingual model* terdiri lebih dari 100 bahasa (Alammary, 2022). Dalam konteks *multilingual sentiment*, analisis sentimen dapat dilakukan pada teks yang ditulis dalam berbagai jenis bahasa, termasuk bahasa Inggris, Indonesia, Prancis, dan bahasa lainnya.

3. TINJAUAN PUSTAKA

Sosial media menjadi salah satu platform yang memberikan fasilitas bagi pengguna untuk membagikan pemikiran dan informasi pribadi mereka secara *real time*. Hal ini menyebabkan peningkatan jumlah informasi yang dibagikan melalui sosial media. Informasi ini tentu menjadi sumber yang bagus bagi seorang peneliti dalam pengambilan keputusan. Hal ini bisa dilakukan dengan menganalisis opini publik yang dibagikan dalam media sosial.

Analisis sentimen menjadi salah satu subjek penelitian yang menarik di bidang pemrosesan bahasa alami. Tindakan menganalisis sikap atau perasaan seseorang berdasarkan kata-kata yang tertulis pada sosial media yang disebut dengan sentimen (Alsaedi dan Khan, 2019). Teknik ini seperti rekayasa fitur untuk menganalisis karakter konten tekstual seperti postingan atau komentar pengguna (Carrillo-de-Albornoz, Vidal and Plaza, 2018; Nawangsari, Kusumaningrum and Wibowo, 2019). Analisis polaritas memungkinkan penggalian sentimen terkait dengan kata tertentu (Mumtaz and Ahuja, 2018). Kemudian, metode analisis sentimen dapat diterapkan untuk memberikan hasil akhir analisis (Li et al., 2018). Analisis sentimen memiliki keterkaitan dengan kondisi emosional, karena emosional sekarang dapat dijadikan sebagai cara umum untuk memberikan gambaran sentimen, dan emosional dapat mewakili keadaan sentimen positif atau negatif.

Sumber daya internet telah diketahui dapat menawarkan dukungan dalam menggambarkan keadaan emosional dan sosial (Bashir et al., 2021). Twitter menjadi salah satu layanan media sosial dengan *platform* online yang memungkinkan pengguna mengirimkan pesan secara *realtime*. Twitter telah menjadi bagian penting dari rutinitas masyarakat modern karena menyediakan forum publik, dimana individu dapat dengan bebas memberikan pandangan pemikiran terhadap berbagai macam topik (Sajwan et al., 2023). Beberapa penelitian telah dilakukan menggunakan Twitter sebagai platform yang memungkinkan dalam penyedia informasi, mengukur opini publik, dan mengekspresikan emosi.

Seperti penelitian yang dilakukan oleh Patterwar dan Jain (2022) melakukan analisis sentimen dengan mengevaluasi polaritas tweet untuk memprediksi saham. Hasilnya, berdasarkan analisis sentimen, detail saham berhasil diprediksi dan mendapatkan keefektifan akurasi mencapai 97,06%. Penelitian oleh Wadhvani et al., (2023) yang bertujuan untuk melihat persepsi publik tentang krisis antara Rusia dan Ukraina dengan memanfaatkan data tweet sebanyak 11.250 tentang perang antara Rusia dan Ukraina, dan berhasil mengembangkan tiga buah model klasifikasi teknik ekstraksi fitur. Penelitian

Lacatus dan Meibauer (2023) menjelaskan data tweet dapat membantu dalam menentukan kebijakan pemerintah dalam menangani krisis, dengan menganalisis tweet Trump selama bulan pertama pandemi COVID-19.

4. HASIL DAN PEMBAHASAN

Hasil klasifikasi dari masing-masing tahapan untuk analisis sentimen chatGPT dilakukan menggunakan orange data mining. Model analisis sentimen yang ditawarkan menggunakan orange data mining dapat dilihat pada gambar 2.

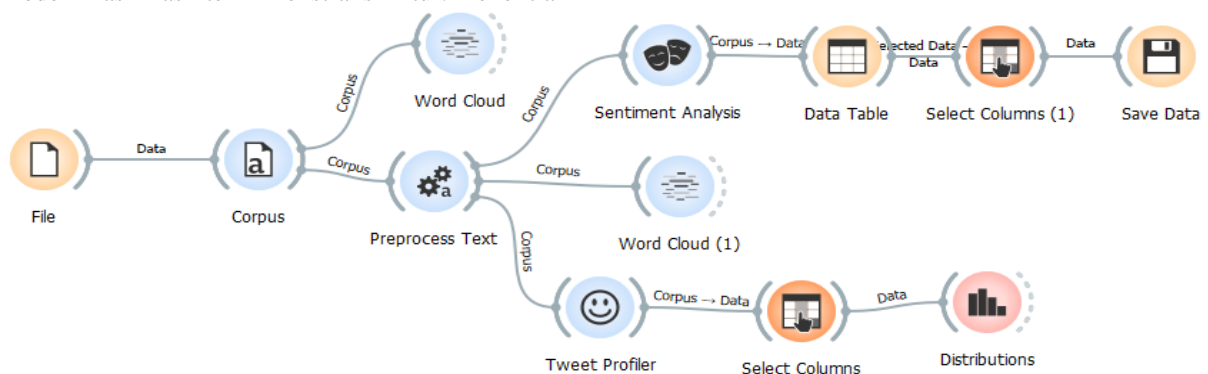
Model tersebut dapat digunakan setelah tahap *crawling* data dan analisis/klasifikasi tweet sudah dilakukan. Beberapa widget digunakan dalam proses analisis sentiment data Twitter terkait ChatGPT. Pada setiap widget terdapat fungsi, tujuan dan hasilnya masing-masing. Model ini ditawarkan agar ke depan, proses analisis sentimen terkait berbagai topik dari sebuah platform media sosial dapat dilakukan dengan lebih mudah dan efisien. Model ini menggunakan algoritma Ekman yang terdapat pada widget tweet profiler, berfungsi untuk menganalisis emosi data tweet. Terdapat beberapa algoritma yang dapat digunakan pada widget tweet profiler yaitu Ekman, Plunchik, dan Profile Of Mood States (POMS). Algoritma tersebut bisa dipilih dan digunakan untuk proses analisis sentimen menggunakan orange data mining.

4.1 Crawling Data

API Twitter bisa diperoleh setelah membuat akun pada *Twitter developer*. *Tools* yang digunakan dalam penelitian ini yaitu RapidMiner. Tweet yang didapatkan sebanyak 6.000 lebih tweet dengan bahasa Inggris dan bahasa Indonesia.

4.2 Analisis dan Klasifikasi Tweet

Sebanyak 6.000 lebih tweet didapatkan terkait dengan #chatGPT. Tweet tersebut selanjutnya dianalisis agar tidak terjadi duplikasi tweet, menghasilkan sebanyak 5.833 terdiri dari bahasa Inggris dan bahasa Indonesia.



Gambar 2. Model analisis sentimen

Sebanyak 100 tweet dipilih secara random dan dianalisis sesuai dengan karakteristik sentimennya. Seratus tweet tersebut dipilih sehingga kategori sentimen yang direpresentasikan melalui 100 tweet tersebut dapat dianalisis secara komprehensif. Setiap tweet dibaca dan dipahami dengan teliti sehingga sentimen yang diwakili melalui tweet tersebut dapat didokumentasikan dengan baik.

Dengan demikian, tweet dikategorikan berdasarkan konteks subjek yang dihasilkan. Analisis konten dari 100 tweet tersebut dilakukan secara manual. Tweet diklasifikasikan dalam lima kategori sentimen, dapat dilihat pada Tabel 1. Studi sebelumnya (Sajwan et al., 2023) juga membantu mengeksplorasi kategori sentimen tambahan yang kemudian diperiksa dan dimodifikasi.

Analisis dilakukan menggunakan data tweet sebanyak 5.833 dan mengklasifikasikannya dalam lima kategori, berdasarkan ekspresi yang dibuat dalam postingan setiap tweet yang didapatkan.

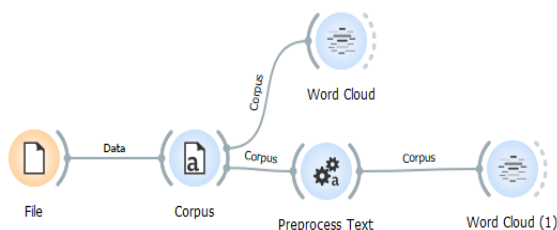
Tabel 1. Klasifikasi kelompok sentimen tweet

Kategori	Deskripsi
Bertanya	Postingan dengan pertanyaan terkait chatGPT
Mengungkapkan emosi	Ekspresi kebahagiaan, kesedihan, enjoy, terkejut, dan emosi positif atau negatif
Perbedaan pendapat	Postingan yang mengungkapkan perbedaan pandangan
Distribusi berita dan data	Berupa artikel berita, informasi, dan item relevan yang berkaitan dengan chatGPT
Tidak relevan	Postingan yang tidak ada keterkaitan dengan chatGPT

Dilakukan klasifikasi juga terhadap 5.833 tweet tersebut untuk melihat respon pengguna Twitter terhadap chatGPT berdasarkan bahasa. Terhitung sebanyak 641 tweet berbahasa Indonesia dan 5.192 tweet berbahasa Inggris.

4.3 Prapemrosesan

Tahap prapemrosesan menjadi bagian yang penting untuk dilakukan. Orange data mining menyediakan *widget preprocess text* dalam melakukan tahap prapemrosesan yang mencakup *i) Transformation, ii) Tokenization, iii) Normalization, dan iv) Filtering*. Selanjutnya *widget* yang digunakan yaitu *Word Cloud*, berfungsi untuk memberikan visualisasi kata sebelum dan setelah proses prapemrosesan. Gambar 3. menunjukkan alur prapemrosesan menggunakan orange data mining.



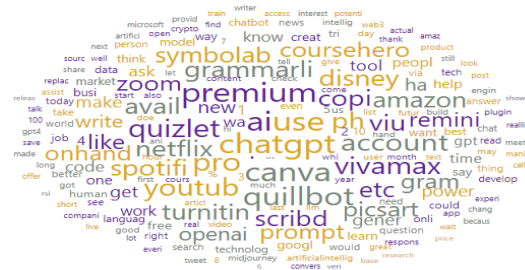
Gambar 3. Alur prapemrosesan

Word Cloud merupakan suatu penggambaran grafis dari kata-kata untuk bagian tertentu dalam konten tertulis, kemudian diatur menurut

frekuensinya (Jayashankar dan Sridaran, 2016). Seperti yang terlihat pada gambar 3, beberapa widget telah digunakan untuk proses word cloud, proses ini menggunakan widget Corpus, preprocess text, dan word cloud. Widget corpus merupakan salah satu widget yang memungkinkan pengguna untuk memanipulasi dan menganalisis data teks. Widget ini dapat digunakan untuk tugas-tugas seperti pemrosesan teks, ekstraksi fitur, dan analisis teks. Widget Preprocess Text dapat memungkinkan pengguna untuk dapat membersihkan, mempersiapkan, dan mengoptimalkan data teks sebelum melakukan analisis lebih lanjut, seperti pemodelan topik, klasifikasi, atau pengelompokan. Widget word cloud berfungsi sebagai visualisasi frekuensi kemunculan kata dalam sebuah data teks sehingga dapat mudah dipahami. Gambar 4 dan 5 memberikan representasi grafis dari data teks (*word cloud*) dari chatGPT. Gambar 4 memberikan gambaran *word cloud* sebelum *preprocessing*, dan gambar 5 setelah *word cloud* melalui tahap *preprocessing*.



Gambar. 4 Sebelum *preprocessing*

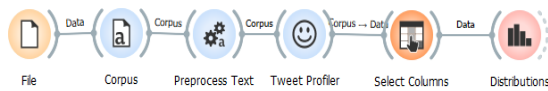


Gambar. 5 Setelah *preprocessing*

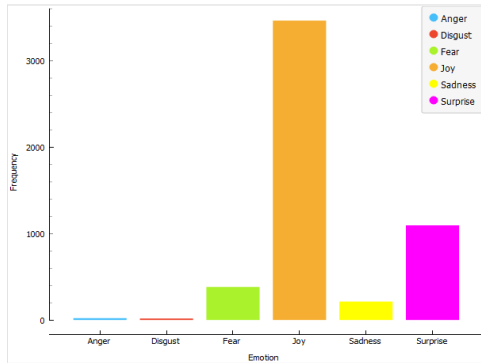
Dapat dilihat pada gambar 4 dan 5 memiliki perbedaan. Terlihat pada gambar 5 sudah tidak terdapat tanda baca, *mention* (@), *hashtag* (#), *retweet*, dan kata atau tanda baca yang tidak dibutuhkan.

4.4 Analisis Emosional

Analisis emosional dilakukan untuk mengetahui tingkat emosional terkait dengan chatGPT. Analisis ini dapat dilakukan menggunakan widget *tweet profile* yang menggunakan algoritma Ekman (*multiclass*). Alur perhitungan menggunakan orange data mining dapat dilihat pada gambar 6.



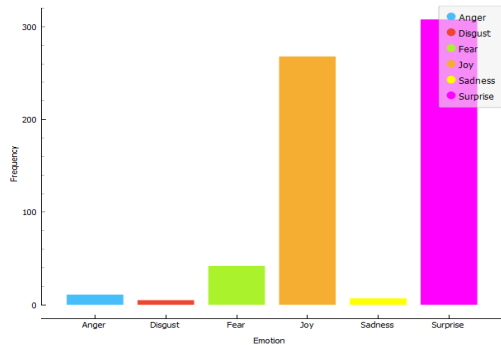
Gambar 6. Alur analisis emosional



Gambar 7. Hasil analisis emosional bahasa Inggris

Gambar 7 menunjukkan hasil perhitungan yang dilakukan. Marah = 21, penghinaan = 17, ketakutan = 382, kebahagiaan = 3464, kesedihan = 213, dan terkejut = 1095. Dari hasil tersebut, dapat diambil kesimpulan bahwa mayoritas masyarakat yang menggunakan twitter dengan bahasa Inggris merasa bahagia dengan dirilisnya chatGPT sebagai *chatbot artificial intelligence*.

Analisis juga dilakukan pada tweet berbahasa Indonesia. Dengan tweet sebanyak 641, menghasilkan emosional pengguna dengan nilai tertinggi yaitu terkejut. Hasil dapat dilihat pada gambar 8.



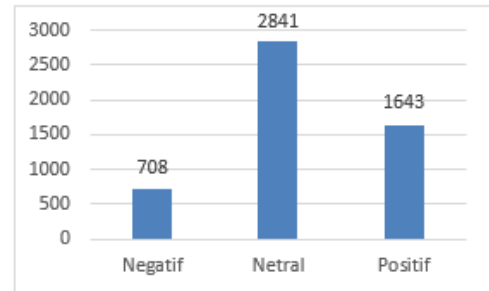
Gambar 8. Hasil analisis emosional bahasa Indonesia

Hasil analisis tweet berbahasa Indonesia menunjukkan ekspresi marah = 11, penghinaan = 5, ketakutan = 42, kebahagiaan = 268, kesedihan = 7, dan terkejut = 308. Hal ini menunjukkan bahwa, mayoritas masyarakat dengan Bahasa Indonesia merasa terkejut dengan dirilisnya chatGPT.

4.5 Klasifikasi polaritas

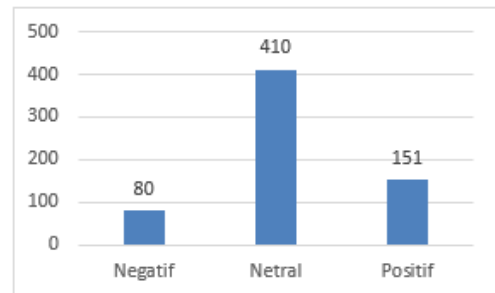
Klasifikasi polaritas sentimen dilakukan menggunakan teknik *Liu-Hu* dan *Multilingual sentiment*. Orange data mining menyediakan alat tersebut pada widget *sentiment analysis*. Alur yang digunakan pada tahap ini mencakup beberapa widget yang digunakan, mulai dari *widget file*, *corpus*,

preprocess text, *sentiment analysis*, *data table*, *select columns*, dan *save data*.



Gambar 9. Hasil analisis sentimen bahasa Inggris

Gambar 9 memberikan penjelasan skor sentimen 5.192 tweet terkait dengan chatGPT dengan tweet berbahasa Inggris. Dapat diamati melalui grafik bahwa sentimen netral memiliki angka maksimum sebesar 2.841, sentimen positif sebesar 1.643, dan sentimen negatif sebesar 708. Jika dihitung persentase analisis polaritas, didapatkan hasil sebesar, sentimen netral = 54,72%, positif = 31,64%, dan negatif = 13,64%. Tidak jauh berbeda dengan hasil analisis sentimen chatGPT berbahasa Indonesia dengan menganalisis tweet sebanyak 641. Dapat dilihat pada gambar 10.



Gambar 10. Hasil analisis sentimen bahasa Indonesia

Grafik pada gambar 10 menjelaskan bahwa sentimen netral memiliki angka maksimum sebesar 410, sentimen positif sebesar 151, dan sentimen negatif sebesar 80. Jika dihitung persentase analisis polaritas, didapatkan hasil sebesar, sentimen netral = 63,96%, positif = 23,56%, dan negatif = 12,48%. Sehingga dapat diambil kesimpulan bahwa, dirilisnya *chatbot artificial intelligence* chatGPT, mendapatkan respon netral dan tidak terdapat penolakan oleh masyarakat.

5. KESIMPULAN DAN SARAN

ChatGPT menjadi salah satu model yang dirilis oleh OpenAI. ChatGPT adalah model chatbot yang dilatih untuk dapat memahami dan menghasilkan teks dalam berbagai bahasa. Kecerdasan chatGPT tersebut menimbulkan kekhawatiran di antara beberapa pendidik. Publik saat ini cenderung menggunakan media sosial dalam menyampaikan opini terhadap sesuatu. Twitter menjadi salah satu media sosial yang sering digunakan dalam menyampaikan opini

tersebut. Kumpulan tweet dari publik tersebut menunjukkan perasaan orang terhadap suatu peristiwa. Hal ini dapat menjadi kumpulan data yang dapat digunakan untuk mengukur sentimen publik terhadap rilisnya chatGPT.

Metode analisis sentimen disajikan dan dijelaskan secara rinci dengan menggunakan perangkat orange data mining. Hasilnya menunjukkan bahwa sebagian besar masyarakat mengungkapkan kebahagiaan terhadap rilisnya chatGPT, menempatkan emosional terkejut pada kategori tertinggi kedua berdasarkan publik dengan bahasa Inggris. Berbeda hasil dengan analisis sentimen berbahasa Indonesia, sebagian besar terkejut atas rilisnya chatGPT oleh OpenAI. Hasil analisis sentimen menunjukkan bahwa chatGPT mendapatkan reaksi netral oleh publik, artinya perilis chatGPT tidak terdapat penolakan oleh publik.

Penelitian ini masih terbatas pada bahasa Indonesia dan bahasa Inggris, masih banyak tweet tentang chatGPT dalam bahasa lain. Kedepan analisis dapat dilakukan dengan menggunakan banyak bahasa. Selain itu penelitian ini masih dibatasi dengan tweet tekstual yang ada di twitter. Di masa mendatang, data baru seperti suara, artikel, dan bentuk multimedia lainnya dapat diikutsertakan dalam proses pengumpulan data.

DAFTAR PUSTAKA

- ABAYOMI-ALLI, A., ABAYOMI-ALLI, O., MISRA, S. and FERNANDEZ-SANZ, L., 2022. Study of the Yahoo-Yahoo Hash-Tag Tweets Using Sentiment Analysis and Opinion Mining Algorithms. *Information (Switzerland)*, 13(3). <https://doi.org/10.3390/info13030152>.
- AGHEMO, A., FORNER, A. and VALENTI, L., 2023. Should Artificial Intelligence-based language models be allowed in developing scientific manuscripts? A debate between ChatGPT and the editors of *Liver International*. *Liver International*, <https://doi.org/10.1111/liv.15580>.
- AGUSTINA, N. and IHSAN, C.N., 2023. Pendekatan Ensemble Untuk Analisis Sentimen Covid19 Menggunakan Pengklasifikasi Soft Voting An Ensemble Approach For Covid19 Sentiment Analysis Using Soft Voting Classifier. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 10(2), pp.263–270. <https://doi.org/10.25126/jtiik.2023106215>.
- ALAMMARY, A.S., 2022. BERT Models for Arabic Text Classification: A Systematic Review. *Applied Sciences (Switzerland)*, <https://doi.org/10.3390/app12115720>.
- ALSAEEDI, A. and KHAN, M.Z., 2019. A Study on Sentiment Analysis Techniques of Twitter Data. *IJACSA) International Journal of Advanced Computer Science and Applications*, [online] 10(2). Available at: <www.ijacsa.thesai.org>.
- AMINIMOTLAGH, M., SHAHHOSEINI, H.S. and FATEHI, N., 2023. A reliable sentiment analysis for classification of tweets in social networks. *Social Network Analysis and Mining*, 13(1). <https://doi.org/10.1007/s13278-022-00998-2>.
- BASHIR, S., BANO, S., SHUEB, S., GUL, S., MIR, A.A., ASHRAF, R., SHAKEELA and NOOR, N., 2021. Twitter chirps for Syrian people: Sentiment analysis of tweets related to Syria Chemical Attack. *International Journal of Disaster Risk Reduction*, 62. <https://doi.org/10.1016/j.ijdr.2021.102397>.
- BOUKES, M., 2019. Social network sites and acquiring current affairs knowledge: The impact of Twitter and Facebook usage on learning about the news. *Journal of Information Technology and Politics*, 16(1), pp.36–51. <https://doi.org/10.1080/19331681.2019.1572568>.
- CARRILLO-DE-ALBORNOZ, J., VIDAL, J.R. and PLAZA, L., 2018. Feature engineering for sentiment analysis in e-health forums. *PLoS ONE*, 13(11). <https://doi.org/10.1371/journal.pone.0207996>.
- COOPER, G., 2023. Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-023-10039-y>.
- GHOUS, H. and KOVACS, L., 2020. Efficiency comparison of Python and RapidMiner. *Multidisziplináris Tudományok*, 10(3), pp.212–220. <https://doi.org/10.35925/j.multi.2020.3.26>.
- HOLZINGER, A., KEIBLINGER, K., HOLUB, P., ZATLOUKAL, K. and MULLER, H., 2023. AI for life: Trends in artificial intelligence for biotechnology. *New Biotechnology*, 74, pp.16–24. <https://doi.org/10.1016/j.nbt.2023.02.001>.
- HU, M. and LIU, B., 2004. Mining Opinion Features in Customer Reviews. In *Proceedings of the 19th national conference on Artificial Intelligence (AAAI'04)*. [online] Available at: <www.aaai.org>.
- IMRAN, A.S., DAUDPOTA, S.M., KASTRATI, Z. and BATRA, R., 2020. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets. *IEEE Access*, 8, pp.181074–181090. <https://doi.org/10.1109/ACCESS.2020.3027350>.
- JAYASHANKAR, S. and R. SRIDARAN, 2016. Moving word cloud from visual towards text analysis to endow eLearning. *Proceedings of the 10th INDIACom; 2016 3rd International Conference on Computing for Sustainable Global Development*, (3493–3498).
- KARAALI, G., 2023. Artificial Intelligence, Basic Skills, and Quantitative Literacy. *Numeracy*,

- 16(1). <https://doi.org/10.5038/1936-4660.16.1.1438>.
- LABILLE, K., GAUCH, S. and ALFARHOOD, S., 2017. Creating Domain-Specific Sentiment Lexicons via Text Mining. In: *In Proceedings of the Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM'17)*. IEEE Computer Society. pp.1–8. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>.
- LACATUS, C. and MEIBAUER, G., 2023. Crisis, Rhetoric and Right-Wing Populist Incumbency: An Analysis of Donald Trump's Tweets and Press Briefings. *Government and Opposition*, 58(2), pp.249–267. <https://doi.org/10.1017/gov.2021.34>
- LI, W., GUO, K., SHI, Y., ZHU, L. and ZHENG, Y., 2018. DWWP: Domain-specific new words detection and word propagation system for sentiment analysis in the tourism domain. *Knowledge-Based Systems*, 146, pp.203–214. <https://doi.org/10.1016/j.knosys.2018.02.004>.
- MUMTAZ, D. and AHUJA, B., 2018. A Lexical and Machine Learning-Based Hybrid System for Sentiment Analysis. *Innovations in Computational Intelligence*, 713, pp.165–175.
- NAWANGSARI, R.P., KUSUMANINRUM, R. and WIBOWO, A., 2019. Word2vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study. In: *Procedia Computer Science*. Elsevier B.V. pp.360–366. <https://doi.org/10.1016/j.procs.2019.08.178>.
- ÖZTURK, N. and AYVAZ, S., 2018. Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telematics and Informatics*, 35(1), pp.136–147. <https://doi.org/10.1016/j.tele.2017.10.006>.
- PANG, B. and LEE, L., 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), pp.1–135.
- PATTERWAR, T. and JAIN, D., 2022. Stock prediction analysis by customers opinion in Twitter data using an optimized intelligent model. *Social Network Analysis and Mining*, 12(1). <https://doi.org/10.1007/s13278-022-00979-5>
- SAJWAN, V., AWASTHI, M., GOEL, A. and SHARMA, P., 2023. Sentiment analysis of Twitter data regarding the agnipath scheme of the defense forces. *Indonesian Journal of Electrical Engineering and Computer Science*, 30(3), pp.1643–1650. <https://doi.org/10.11591/ijeecs.v30.i3.pp1643-1650>.
- SCHART, M., 2022. The ChatGPT chatbot is blowing people away with its writing skills. The University of Sydney [online] Tersedia di: <<https://www.sydney.edu.au/news-opinion/news/2022/12/08/the-chatgpt-chatbot-is-blowing-people-away-with-its-writing-skil.html>> [Diakses 23 November 2023].
- SEKIOKA, S., HATANO, R. and NIAHIYAMA, H., 2023. Market prediction using machine learning based on social media specific features. *Artificial Life and Robotics*. <https://doi.org/10.1007/s10015-023-00857-z>.
- T. ELSHOUSH, H. and A. DINAR, E., 2019. Using Adaboost and Stochastic gradient descent (SGD) Algorithms with R and Orange Software for Filtering E-mail Spam. 2019 11th Computer Science and Electronic Engineering (CEECE), pp.41–46. <https://doi.org/10.1109/CEECE47804.2019.8974319>
- TONG, Y. and ZHANG, L., 2023. Discovering the next decade's synthetic biology research trends with ChatGPT. *Synthetic and Systems Biotechnology*, <https://doi.org/10.1016/j.synbio.2023.02.004>.
- TRI SAPUTRA, F., NURHADRYANI, Y., HARTONO WIJAYA, S. and DEFINA, 2021. Analisis Sentimen Bahasa Indonesia Pada Twitter Menggunakan Struktur Tree Berbasis Leksikon. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 8(1), pp.135–146. <https://doi.org/10.25126/jtiik.202184133>.
- WADHWANI, G.K., VARSHNEY, P.K., GUPTA, A. and KUMAR, S., 2023. Sentiment Analysis and Comprehensive Evaluation of Supervised Machine Learning Models Using Twitter Data on Russia–Ukraine War. *SN Computer Science*, 4(4). <https://doi.org/10.1007/s42979-023-01790-5>
- WONGKAR, M. and ANGDRESEY, A., 2019. Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler : Twitter. 2019 Fourth International Conference on Informatics and Computing (ICIC). <https://doi.org/10.1109/ICIC47613.2019.8985884>
- YANG, W., 2022. Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100061>.