

Sentiment Analysis Using Support Vector Machine (SVM) of ChatGPT Application Users in Play Store

Muthia Sakhdiah, Admi Salma*, Dony Permana, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: admisalma1@fmipa.unp.ac.id

Submitted : 22 April 2024

Revised : 28 Mei 2024

Accepted : 29 Mei 2024

ABSTRACT

The ChatGPT application is an Artificial Intelligence (AI) technology that responds to conversations in form text and voice messages, and is accessible via smartphones or computers. The ChatGPT provides answers and solutions related to the problems asked, the speed and complexity of the answers are also added values of this application. However, there are negative impacts, one of which is the vulnerability of scientific papers to plagiarism. -Because of this, there are many reviews from the community that assess this application. These reviews can be seen on the Play Store which can be a reference before downloading the ChatGPT application. How the community responds can be seen through sentiment analysis, which will classify positive and negative assessments. Making it easier for companies to evaluate products. Then classification is carried out using Support Vector Machine (SVM), the classification model obtained is used to classify user reviews of the ChatGPT application. The results showed an accuracy of 93.9% with a linear kernel, and the sentiment of people who use the ChatGPT application is more positive.

Keywords: ChatGPT Application, Classification, Sentiment Analysis, Support Vector Machine (SVM)



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Penerapan teknologi pada bidang komputer salah satunya yaitu teknologi *Artificial Intelligence* (AI) atau lebih dikenal sebagai kecerdasan buatan. Menurut Kusmadewi (2003) kecerdasan buatan atau *artificial intelligence* merupakan salah satu bagian ilmu komputer yang membuat mesin (komputer) dapat melakukan pekerjaan seperti dan sebaik yang dilakukan oleh manusia. AI menggunakan kecerdasan buatan yang memungkinkannya bekerja dan menyelesaikan masalah layaknya manusia. Salah satu produk aplikasi yang memanfaatkan teknologi AI adalah aplikasi ChatGPT yang dikembangkan oleh OpenAI, dirancang dengan arsitektur *Generative Pre-Trained* versi GPT-3.5 dan teknologi *deep learning* dengan *Natural Language Processing* (NLP) yang membantu pengguna dalam menemukan informasi, menerjemahkan bahasa, atau menghasilkan teks baru yang mirip dengan ucapan manusia.

Kemampuan aplikasi ChatGPT memberikan manfaat salah satunya menjadi akses pembelajaran mandiri bagi pelajar. Namun, kemudahan dari penerapan teknologi AI pada aplikasi ChatGPT juga menimbulkan dampak negatif, diantaranya menyebabkan semakin berkurangnya minat baca masyarakat, rentannya plagiarisme dalam karya ilmiah, dan masih banyak lagi. Oleh sebab itu, banyak ulasan positif maupun negatif dari masyarakat yang menilai aplikasi ChatGPT. Ulasan pengguna aplikasi ChatGPT tersebut dapat dilihat pada *Google Play Store*. Pengguna dapat menilai kinerja aplikasi dan dapat dijadikan acuan bagi pengguna lain sebelum mengunduh aplikasi yang sama. Untuk melihat bagaimana tanggapan masyarakat dapat dilakukan dengan analisis sentimen. Menurut Zuriel & Fahrurrozi (2021) analisis sentimen dinilai sebagai alat bantu untuk mengetahui bagaimana sentimen atau opini masyarakat terhadap sebuah layanan, kebijakan, dan sebagainya. Analisis sentimen mengolah data teks yang berisi ulasan, pendapat, atau komentar yang nantinya menghasilkan sentimen positif, negatif, atau netral. Analisis sentimen memudahkan perusahaan dalam mengevaluasi suatu produk ataupun layanan dan jasa yang digunakan untuk pengembangan dan peningkatan kualitas produk yang lebih baik lagi.

Algoritma *machine learning* yang banyak digunakan pada analisis sentimen salah satunya *Support Vector Machine* (SVM). Shah dkk (2020) dalam bukunya menjelaskan bahwa SVM dapat diterapkan untuk berbagai macam masalah, termasuk kategorisasi teks, identifikasi gambar, dan prediksi harga pasar. Penelitian sebelumnya oleh Tuhenay & Mailoa (2021) yang membandingkan *Naïve Bayes Classifier* (NBC) dan SVM mendapatkan akurasi SVM yang lebih

tinggi sebesar 96,34%. Kemudian penelitian oleh Iskandar dan Nataliani (2021) yang membandingkan klasifikasi NBC, SVM, dan *K-Nearest Neighbor* (KNN) terhadap analisis sentimen *gadget* berdasarkan aspek harga, spesifikasi, desain, dan citra *merk* mendapatkan hasil klasifikasi terbaik dilakukan oleh SVM dengan tingkat akurasi sebesar 96,43%. Penelitian yang dilakukan oleh Syafia dkk (2023) tentang analisis sentimen komentar *youtube* penggemar BTS dengan membandingkan algoritma SVM dan *Random Forest* mendapatkan akurasi SVM yang lebih tinggi sebesar 96%. Berdasarkan penelitian yang telah dilakukan sebelumnya, diperoleh kesimpulan bahwa SVM baik dalam mengklasifikasikan data teks pada analisis sentimen. Perbedaan penelitian sebelumnya dengan penelitian yang akan dilakukan yaitu data yang digunakan adalah data ulasan pengguna aplikasi ChatGPT pada kolom komentar di *Google Play Store* serta menampilkan visualisasi kata dalam bentuk *wordcloud*. Berdasarkan penjelasan tersebut, tujuan penelitian yang dilakukan adalah untuk mengklasifikasikan sentimen pengguna aplikasi ChatGPT menjadi dua kelas yaitu kelas positif dan kelas negatif, serta mengetahui performa metode SVM dalam mengklasifikasikan data sentimen.

II. METODE PENELITIAN

A. Jenis Penelitian dan Sumber Data

Penelitian ini merupakan penelitian dengan jenis terapan (*Applied Research*). Data yang digunakan adalah data sekunder berupa kumpulan ulasan pengguna aplikasi ChatGPT di Indonesia yang terdapat pada kolom komentar di *Google Play Store* yang diperoleh dengan menggunakan teknik *scrapping* data. Data yang digunakan adalah data ulasan tanggal 27 Juli 2023 pukul 15:21 WIB hingga data ulasan tanggal 6 Januari 2024 pukul 22:10 WIB. Variabel penelitian yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Keterangan	Variabel
<i>Term ke-i</i> yang terdapat pada ulasan	X
Kelas sentimen positif dan kelas sentimen negatif	Y

B. Teknik Analisis Data

Adapun tahapan analisis data dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan Data

Data yang dikumpulkan adalah data ulasan pengguna aplikasi ChatGPT di Indonesia dari *Google Play Store* menggunakan teknik *scrapping* data. *Scrapping* dilakukan dengan memanfaatkan *library Python Scrapper* dan memasukkan *link (Uniform Resource Locator) URL* aplikasi ChatGPT di *Google Play Store*.

2. Melakukan Text Pre-Processing

Adapun tahapan *text pre-processing* menggunakan *software Phyton* sebagai berikut.

- Cleaning* data, menghapus komponen-komponen asing yang tidak dibutuhkan dalam data, seperti *link URL*, *hashtag*, emoji, tanda baca, dan lainnya.
- Case folding*, yaitu mengubah setiap huruf pada ulasan menjadi huruf kecil (*lower case*) untuk menghindari salah ejaan dalam dokumen.
- Stopword Removeral*, menghapus *stopword* pada data ulasan yaitu kata-kata yang tidak bermakna, kata penghubung, dan lainnya dengan memanfaatkan *library nltk.corpus* serta menambahkan *stopword* yang akan dihapus.
- Tokenizing*, memisah kalimat menjadi kata per kata yang disebut *token* atau kata yang berdiri sendiri dengan memanfaatkan *library nltk.tokenize*.
- Normalisasi, yaitu perbaikan kata singkatan dan kata yang tidak baku atau *slangword* menjadi kata baku sesuai dengan Kamus Besar Bahasa Indonesia (KBBI) dengan menambahkan kamus normalisasi ke dalam *software*. Kamus normalisasi diperoleh dari situs Github.
- Stemming*, menghapus imbuhan pada kata dan menjadikannya ke kata dasar yang benar sesuai dengan KBBI dengan memanfaatkan *library Sastrawi* pada *Phyton*.

3. Pelabelan Kelas Sentimen

Klasifikasi dengan *lexicon* sangat mengandalkan kebutuhan kamus *lexicon* bahasa Indonesia, karena kata (*term*) pada data dapat diberi *score* berdasarkan kamus apakah kata positif atau negatif (Fatihin, 2022). Namun saat melakukan pelabelan akan didapati data ulasan yang masuk ke dalam kelas sentimen netral, yaitu ketika sebuah ulasan tidak mengandung *term* positif dan *term* negatif. Hakim (2021) berpendapat kelas sentimen netral dinilai kurang memberi masukan dan manfaat bagi pihak perusahaan. Oleh sebab itu kelas sentimen

netral tidak digunakan dan akan direduksi sehingga sentimen yang diperlukan hanya sentimen positif dan sentimen negatif.

4. Pembobotan Kata

Pembobotan dilakukan dengan metode pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) yang menghitung frekuensi kemunculan *term* atau kata dalam dokumen atau ulasan. TF (*Term Frequency*) merupakan konsep perhitungan bobot dengan menghitung frekuensi kemunculan *term* dalam suatu dokumen. Nilai TF didapatkan dari frekuensi kemunculan *term* (t) dalam dokumen (d). Sedangkan DF (*Document Frequency*) merupakan banyaknya dokumen yang berisi kemunculan *term* yang sama. IDF (*Inverse Document Frequency*) dapat dihitung dengan logaritma dari jumlah seluruh dokumen (N) dibagi dengan banyaknya dokumen yang berisi kemunculan *term* yang sama. Menurut Arifin dan Sulistiyowati (2021) persamaan yang digunakan untuk menghitung TF-IDF adalah sebagai berikut.

$$TF_{t,d} = F_{(t,d)} \quad (1)$$

$$IDF_t = \log \left(\frac{N}{DF_t} \right) \quad (2)$$

$$W_{t,d} = TF_{t,d} \times IDF_t \quad (3)$$

dimana $F_{(t,d)}$ merupakan frekuensi kemunculan *term* ke- t dalam dokumen ke- d , DF_t merupakan Jumlah dokumen yang berisi kemunculan *term* yang sama, $W_{t,d}$ merupakan nilai pembobotan TF-IDF, dan N adalah jumlah dokumen.

5. Membagi Data *Training* dan Data *Testing*

Pratiwi dan Nugroho (2016) data *training* adalah data yang digunakan untuk perhitungan probabilitas dari data berdasarkan data pembelajaran yang dilakukan, sedangkan data *testing* merupakan data yang akan atau sedang terjadi dan dipergunakan sebagai bahan uji yang sebelumnya sudah didapatkan pada data *training*. Penelitian ini menggunakan pembagian data *training* dan data *testing* sebanyak 80%:20% yang dibagi secara acak.

6. Klasifikasi dengan *Support Vector Machine* (SVM)

Konsep dasar SVM adalah memisahkan dua kelas data sebagai proses klasifikasi dengan menentukan garis pemisah yang disebut juga *hyperplane*. *Hyperplane* yaitu sebuah garis atau bidang yang memisahkan data menjadi dua kelas klasifikasi. *Hyperplane* terbaik yang digunakan yaitu *hyperplane* yang memiliki *margin* maksimum. *Margin* adalah jarak antara titik data terluar dengan *hyperplane*. *Margin* dikatakan maksimum yaitu ketika jarak titik data terluar dari masing-masing kelas data yang menuju ke *hyperplane* memiliki jarak yang sama. Istilah *support vector* merupakan titik data terluar dari masing-masing kelas data yang terdekat dengan *hyperplane* sebagai daerah kritis. Persamaan model SVM yang umum digunakan adalah sebagai berikut (Campbell & Ying, 2011: 2).

$$f(x) = \text{sign}(w \cdot x + b)$$

Sehingga data yang berada pada kelas +1 memenuhi:

$$w \cdot x_i + b > 0 \text{ for } y_i = +1$$

dan yang berada pada kelas -1 diperoleh:

$$w \cdot x_i + b < 0 \text{ for } y_i = -1$$

dimana w merupakan parameter bobot, x_i merupakan variabel bebas ke- i , dan b merupakan bias atau *error*. Pastore dkk (2020) dalam buku *Text Mining Praktik Klasifikasi dan Pemodelan Topik dengan Phyton* menjelaskan bahwa mesin vektor dukungan (SVM) dapat diterapkan pada data *linear* dan *non-linear*, dan menggunakan teknik kernel untuk mengubah data menjadi dimensi yang lebih tinggi sehingga dapat dipisahkan oleh *hyperplane*. Teknik kernel dapat digunakan untuk memetakan data ke ruang berdimensi lebih tinggi yang disebut sebagai ruang kernel, dimana akan menjadikan data terpisah secara *linear*. Dimana terdapat beberapa

macam jenis kernel yaitu kernel *linear*, kernel *Radial Bias Function* (RBF), dan kernel *polynomial*. Sehingga fungsi yang digunakan untuk proses klasifikasi data kelas positif dan negatif adalah sebagai berikut.

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \tag{4}$$

dimana $K(x_i, x)$ adalah fungsi kernel yang digunakan, α_i, y_i adalah kelas data ke- i . dan N adalah jumlah data.

7. Pengujian Ketepatan Klasifikasi

Hidayatullah dkk (2019) menjelaskan bahwa *confusion matrix* merupakan matriks yang menyimpan informasi untuk mengetahui performa model dan sebagai acuan performa klasifikasi dari algoritma yang digunakan pada tahap evaluasi. Matriks tersebut dapat dituliskan sebagai berikut (Saputro & Sari, 2019).

Tabel 2. Confusion Matrix

Aktual	Prediksi	
	Negatif	Positif
Negatif	TN (<i>True Negative</i>)	FP (<i>False Positive</i>)
Positif	FN (<i>False Negative</i>)	TP (<i>True Positive</i>)

Setelah mendapatkan nilai dari *confusion matrix*, dapat dihitung nilai akurasi, presisi, dan *recall*. Menurut Saputo & Sari (2019) performa model dapat diukur dengan menghitung nilai *accuracy, precision, dan recall*. Berikut persamaan yang digunakan.

$$Akurasi = \frac{TP + TN}{TN + FN + FP + TP} \tag{5}$$

$$Presisi = \frac{TP}{FP + TP} \tag{6}$$

$$Recall = \frac{TP}{FN + TP} \tag{7}$$

8. Penarikan Kesimpulan dan Visualisasi

Visualisasi dilakukan dengan membuat *wordcloud* yang menggambarkan frekuensi *term* yang paling banyak muncul, *term* dengan ukuran yang besar merupakan informasi penting dalam analisis sentimen.

III. HASIL DAN PEMBAHASAN

Analisis sentimen menggunakan SVM terhadap ulasan aplikasi ChatGPT di *Play Store* dimulai dengan pengumpulan data menggunakan *software Python* yang dijalankan di *Google Collab*. Hasil data yang diperoleh berjumlah 2.500 data ulasan. Data ulasan tersebut menjadi data awal yang akan dilakukan tahap *text pre-processing*. Data yang telah melalui *text pre-processing* dapat dilakukan untuk analisis lebih lanjut. Hasil dari tahapan *text pre-processing* disajikan pada Tabel 3.

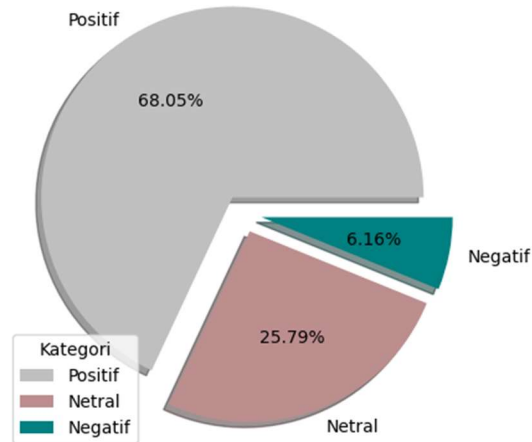
Tabel 3. Hasil Text Pre-Processing

Tahapan	Data Ulasan
Input	Sangat membantu 🙌🙌, semua pertanyaan bisa di jawabnya, bisa juga untuk nanya saran dan pendapat pilihan yang lebih bagus. Semua pekerjaan jadi termudahkan
Output	['bantu', 'tanya', 'jawab', 'tanya', 'saran', 'pendapat', 'pilih', 'bagus', 'kerja', 'mudah']

Setelah melalui tahapan *text pre-processing*, tidak terdapat lagi komponen asing dalam data ulasan seperti *emoticon, tanda baca, hashtag, dan lainnya*. Data ulasan sudah berubah menjadi huruf kecil serta *stopword* pada data

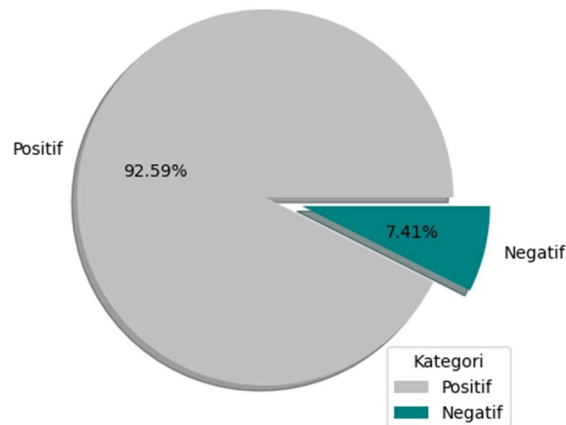
sudah dihapus. Data ulasan sudah terpisah menjadi beberapa *token* atau kata yang berdiri sendiri. Kata-kata yang tidak baku sudah dirubah sesuai dengan KBBI menggunakan kamus normalisasi, dan juga setiap kata yang memiliki imbuhan telah dihapus sehingga menjadi kata-kata dasar.

Hasil dari *text pre-processing* menyisakan 2.307 data ulasan. Kemudian dilanjutkan dengan tahapan pelabelan kelas sentimen dengan metode *lexicon based*. Setelah melakukan pelabelan, didapatkan distribusi kelas sentimen yang disajikan pada Gambar 1.



Gambar 1. Distribusi Kelas Sentimen Pengguna Aplikasi ChatGPT

Terlihat bahwa masyarakat yang memberikan ulasan di *Google Play Store* memiliki sentimen yang positif terhadap penggunaan aplikasi ChatGPT. Hal ini dapat dilihat dari persentase sentimen positif lebih besar dibandingkan persentase sentimen netral dan sentimen negatif. Kelas sentimen netral dapat direduksi sehingga akan menghasilkan kelas sentimen positif dan kelas sentimen negatif saja. Hal ini dikarenakan kelas sentimen netral tidak memberikan banyak keuntungan pada perusahaan dalam melakukan evaluasi produk, layanan atau jasa. Gumilang (2018) menjelaskan apabila kelas sentimen netral tidak teridentifikasi kata sentimen positif dan negatif maka akan dimasukkan ke dalam kelas sentimen positif, dan apabila sentimen netral terdapat kata sentimen positif dan negatif yang seimbang maka akan dimasukkan ke dalam kelas sentimen negatif. Berikut hasil distribusi kelas sentimen setelah dilakukan reduksi yang disajikan pada Gambar 2.



Gambar 2. Distribusi Kelas Sentimen setelah Reduksi

Berdasarkan hasil penelitian, contoh dari sentimen positif adalah “ringan cepat”, “alat pintar”, “bagus bantu”, dan lainnya. Contoh dari sentimen negatif yaitu “susah login”, “error”, “kecewa edit”, dan lainnya. Sentimen masyarakat pengguna aplikasi ChatGPT di *Google Play Store* cenderung bermakna positif, yang artinya meskipun terdapat beberapa ulasan yang bermakna negatif, namun secara keseluruhan tanggapan masyarakat terhadap aplikasi ChatGPT bisa

disimpulkan baik dan mendukung penggunaan aplikasi ChatGPT. Berdasarkan penelitian Rifaldi dkk (2023) menghasilkan sentimen positif pengguna aplikasi ChatGPT lebih banyak dari sentimen negatif, yaitu sentimen positif sebanyak 74,09% dan sentimen negatif sebesar 25,91%.

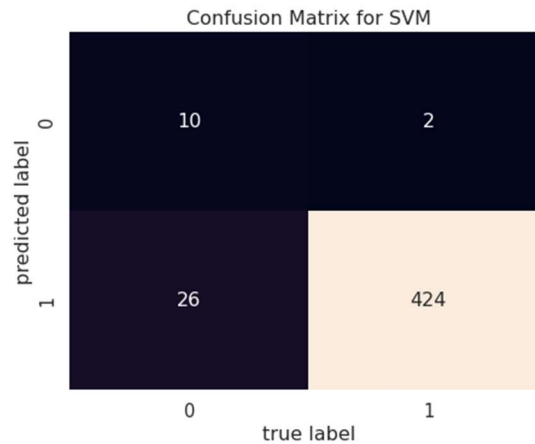
Kelas sentimen positif dan negatif yang telah diperoleh sebelumnya akan diklasifikasikan menggunakan analisis SVM. Sebelum masuk ke proses analisis dilakukan pembagian data, yaitu data *training* dan data *testing* dengan proporsi 80%:20%. Proporsi pembagian data adalah opsional ditentukan oleh penulis jadi tidak ada ketentuan khusus. Setelah dibagi data *training* berjumlah 1.845 dan data *testing* berjumlah 462.

Klasifikasi dengan SVM dilakukan penerapan beberapa fungsi kernel untuk melihat jenis kernel yang tepat untuk digunakan dengan tingkat akurasi tertinggi. Adapun beberapa macam jenis kernel yang digunakan yaitu kernel *linear*, kernel RBF, dan kernel *polynomial*. Hasil akurasi dari penggunaan kernel tersebut disajikan pada Tabel 5.

Tabel 5. Hasil Akurasi Menggunakan Kernel

Jenis Kernel	Hasil Akurasi
<i>Linear</i>	93,9%
RBF	93,5%
<i>Polynomial</i>	93.1%

Klasifikasi dengan metode SVM menghasilkan nilai akurasi yang tinggi, dimana penerapan kernel *linear* mendapatkan nilai akurasi yang lebih tinggi di antara kernel lainnya. Selanjutnya untuk melihat kinerja klasifikasi dengan SVM dapat digunakan *confusion matrix*. Berikut *output confusion matrix* dengan penerapan kernel *linear* yang disajikan pada Gambar 3.

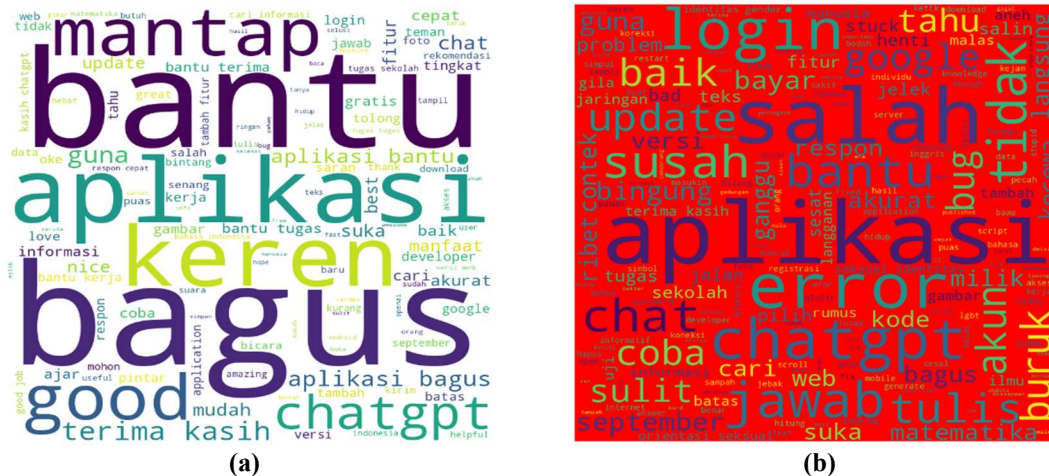


Gambar 3. Output Confusion Matrix

Setelah dilakukan klasifikasi menggunakan SVM dengan kernel *linear*, model mengklasifikasikan 424 data yang awalnya berada pada kelas sentimen positif dan tetap berada di kelas sentimen positif. Sebanyak 10 data yang awalnya berada di kelas sentimen negatif juga diklasifikasikan tetap berada di kelas sentimen negatif. Namun, 26 data yang awalnya berada di kelas sentimen positif, diklasifikasikan ke dalam kelas sentimen negatif, dan terdapat 2 data yang awalnya berada di kelas sentimen negatif, diklasifikasikan ke dalam kelas sentimen positif. Kemudian dari *confusion matrix* pada Gambar 3 dapat dihitung nilai akurasi, presisi, dan *recall* untuk mengukur performa model menggunakan sebagai berikut.

$$\begin{aligned}
 \text{Akurasi} &= \frac{424 + 10}{10 + 26 + 2 + 424} = \frac{434}{462} = 0,94 \\
 \text{Presisi} &= \frac{424}{2 + 424} = \frac{424}{426} = 0,99 \\
 \text{Recall} &= \frac{424}{26 + 424} = \frac{424}{450} = 0,94
 \end{aligned}$$

Berdasarkan hasil dari perhitungan akurasi, presisi, dan *recall* di atas didapatkan bahwa akurasi SVM dengan kernel *linear* dalam mengklasifikasikan ulasan pengguna aplikasi ChatGPT di *Play Store* sebesar 94%. Nilai presisi yang artinya kemampuan model memprediksi data di kelas positif dan tepat berada di kelas positif sebesar 99%. Nilai *recall* yang menunjukkan kemampuan model dalam mengklasifikasikan data positif dari semua data sebesar 94%. Selanjutnya untuk memvisualisasikan *term* yang paling sering muncul dalam sentimen pengguna aplikasi ChatGPT di *Play Store* dapat disajikan dalam bentuk *wordcloud* sebagai berikut.



Gambar 4. (a) “Wordcloud Ulasan Positif”, (b) “Wordcloud Ulasan Negatif”

Gambar 4 (a) merupakan *wordcloud* dari keseluruhan ulasan positif pengguna aplikasi ChatGPT dan Gambar 4 (b) merupakan *wordcloud* dari keseluruhan ulasan negatif pengguna aplikasi ChatGPT di *Google Play Store* yang diperoleh dari tanggal 27 Juli 2023 hingga 6 Januari 2024. Berdasarkan *wordcloud* tersebut dapat diketahui bahwa kata “bagus”, “keren”, “bantu”, “mantap”, dan “good” banyak dikaitkan dengan ulasan positif yang diberikan oleh pengguna aplikasi ChatGPT. Hal ini menunjukkan bahwa kinerja dari aplikasi ChatGPT cukup baik dan membantu para penggunanya. Sedangkan kata “salah”, “error”, “susah”, “bingung”, dan “sakit” banyak dikaitkan dengan ulasan negatif yang diberikan oleh pengguna aplikasi ChatGPT. Hal ini menunjukkan bahwa pengguna aplikasi tersebut masih banyak mengalami kesulitan dalam memahami dan menggunakan aplikasi ChatGPT.

IV. KESIMPULAN

Berdasarkan analisis data yang dikumpulkan sebanyak 2.500 ulasan mengenai sentimen (pandangan) masyarakat pengguna aplikasi ChatGPT dengan pengujian menggunakan algoritma SVM, diperoleh hasil akurasi sebesar 93,9% dengan kernel yang paling baik digunakan yaitu kernel *linear*. Pengujian ketepatan klasifikasi menggunakan *confusion matrix* diperoleh nilai *accuracy*, presisi, dan *recall* berturut-turut sebesar 94%, 99%, dan 94%. Artinya algoritma SVM cocok digunakan pada analisis sentimen pengguna aplikasi ChatGPT di *Google Play Store*. Secara keseluruhan ulasan masyarakat yang menggunakan aplikasi ChatGPT cukup baik. Dilihat dari *wordcloud* yang dihasilkan, masyarakat menilai aplikasi ChatGPT baik dan dapat membantu pengguna dalam melakukan aktifitas sehari-hari baik itu bekerja, menyelesaikan tugas sekolah, ataupun sebagai teman bicara. Namun penilaian negatif dari masyarakat cenderung disebabkan oleh jawaban yang diberikan tidak *update*, susah dalam mengakses aplikasi ChatGPT, *error* pada aplikasi, dan lainnya. Saran untuk penelitian selanjutnya, perlu menambahkan referensi khusus kamus bahasa Indonesia untuk mengidentifikasi kata-kata yang tidak tepat dalam data ulasan, terutama dalam perubahan kata yang tidak baku menjadi kata baku sesuai dengan KBBI.

DAFTAR PUSTAKA

- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan Algoritma Support Vector Machine (Svm) Dengan TF-IDF N-Gram Untuk Text Classification. *Jurnal STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 6(2), 129-136
- Campbell, C., & Ying, Y. (2011). *Learning with Support Vector Machines*. Bristol, Britania Raya: Morgan & Claypool Publishers.
- Fatihin, A. (2022). *Analisis Sentimen terhadap Ulasan Aplikasi Mobile Menggunakan Metode Support Vector Machine (SVM) dan Pendekatan Lexicon Based*. Skripsi. Jakarta: Universitas Islam Negeri Syarif Hidayatullah.
- Gumilang, Z. (2018). *Implementasi Naive-Bayes Classifier dan Asosiasi Untuk Analisis Sentimen Data Ulasan Aplikasi E-commerce Shopee pada Situs Google Play: Skripsi*. Yogyakarta: Universitas Islam Indonesia.
- Hakim, S. N. (2021). *Analisis Sentimen Persepsi Pengguna MyIndihome Menggunakan Metode Support Vector Machine (SVM) dan Naive Bayes Classifier (NBC)*. Skripsi. Yogyakarta: Universitas Islam Indonesia.
- Hidayatullah, A. F., dkk. (2019). Identifikasi Konten Kasar pada Tweet Bahasa Indonesia. *Jurnal Linguistik Komputasional*, 2(1), 1-5.
- Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naive Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal ReSTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(6), 1120-1126.
- Kusmadewi, S. (2003). *Artificial Intelligence*. Yogyakarta: Graha Ilmu.
- Pratiwi, R. W., & Nugroho, Y. S. (2016). Prediksi Rating Film Menggunakan Metode Naive Bayes. *Jurnal Teknik Elektro*, 8(2), 60-63.
- Rifaldi., dkk. (2023). Analisis Sentimen Terhadap Aplikasi Chatgpt Pada Twitter Menggunakan Algoritma Naive Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 7(2), 802-814.
- Saputro, I. W., & Sari, B. W. (2019). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Citec Journal*, 6(1), 1-11.
- Sari, I., & Indarti, D. (2022). *Text Mining Praktik Klasifikasi dan Pemodelan Topik dengan Phyton*. Jawa Timur: Uwais Inspirasi Indonesia.
- Syafia, A. N., Hidayattullah, M. F., & Suteddy, W. (2023). Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS. *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, 8(3), 207-212.
- Tuhenay, D., & Mailoa, E. (2021). Perbandingan Klasifikasi Bahasa Menggunakan Metode Naive Bayes Classifier (NBC) dan Support Vector Machine (SVM). *JIKO (Jurnal Informatika dan Komputer)*, 4(2), 105-111.
- Zuriel, H. P., & Fahrurrozi, A. (2021). Implementasi Algoritma Klasifikasi Support Vector Machine Untuk Analisa Sentimen Pengguna Twitter Terhadap Kebijakan Psbb. *Jurnal Ilmiah dan Informatika Komputer*, 26(2), 149-162.