TECHNOVATAR Jurnal Teknologi, Industri dan Informasi
Vol. 2, No. 4, Oktober 2024, pp. 97 – 110
Open Access: https://journal.awatarapublisher.com/index.php/technovatar
ISSN: 3031-8874, **DOI**: 10.61434/technovatar.v2i4.230

IMPLEMENTASI SVM-PSO DALAM ANALISIS SENTIMEN PENGGUNA GOOGLE PLACE REVIEW DI MARKAS CAFE

Hendri Rosmawan[1] Arief Setyanto[2] Ferry Wahyu Wibowo[3]
[1,2,3] Program Studi Teknik Informatika, Universitas AMIKOM Yogyakarta, Yogyakarta , Indonesia

e-mail: hendri.rosmawan@gmail.com[1] arief_s@amikom.ac.id[2] ferry.w@amikom.ac.id[3]

INFO ARTIKEL
Sejarah Artikel:
Diterima, 09 September 2024
Direvisi, 11 September 2024
Disetujui, 12 September 2024

KEYWORDS
Sentiment Analysis,
SVM,
PSO,
Google Place Review,
Text Classification

ABSTRACT
Sentiment analysis plays an important role in understanding customer perceptions of businesses, allowing companies to respond more effectively to customer needs and satisfaction. This study aims to evaluate the performance of a Support Vector Machine (SVM) model optimized with Particle Swarm Optimization (PSO) in classifying the sentiment of user reviews on Markas Cafe. The dataset consists of 1,533 user reviews categorized into three sentiment classes: positive, neutral, and negative. The optimization process using PSO is used to find the optimal SVM parameters. The results showed that the SVM-PSO model achieved an accuracy of 87.7% and an Area Under Curve (AUC) of 0.85, with the best performance on positive sentiment (94.7% precision and 92.8% recall). Although the model showed good ability in detecting positive sentiments, the results for neutral and negative sentiments indicated the need for further improvement. This study confirms the effectiveness of SVM-PSO in sentiment analysis and suggests this approach can be utilized by businesses to improve marketing and customer service strategies based on user feedback.

This work is licensed under a Creative Commons Attribution 4.0 International License.

KATA KUNCI
Analisis Sentimen,
SVM,
PSO,
Google Place Review,
Klasifikasi Teks

CORRESPONDING AUTHOR
Hendri Rosmawan
Universitas AMIKOM Yogyakarta
Yogyakarta
hendri.rosmawan@gmail.com

ABSTRAK
Analisis sentimen memainkan peran penting dalam memahami persepsi pelanggan terhadap bisnis, memungkinkan perusahaan untuk merespons secara lebih efektif terhadap kebutuhan dan kepuasan pelanggan. Penelitian ini bertujuan untuk mengevaluasi performa model *Support Vector Machine* (SVM) yang dioptimalkan dengan *Particle Swarm Optimization* (PSO) dalam mengklasifikasikan sentimen ulasan pengguna di Markas Cafe. *Dataset* terdiri dari 1.533 ulasan pengguna yang dikategorikan ke dalam tiga kelas sentimen: positif, netral, dan negatif. Proses optimasi menggunakan PSO digunakan untuk menemukan parameter SVM yang optimal. Hasil penelitian menunjukkan bahwa model SVM-PSO mencapai akurasi sebesar 87.7% dan *Area Under Curve* (AUC) sebesar 0.85, dengan performa terbaik pada sentimen positif (presisi 94.7% dan *recall* 92.8%). Meskipun model menunjukkan kemampuan yang baik dalam mendeteksi sentimen positif, hasil untuk sentimen netral dan negatif mengindikasikan perlunya perbaikan lebih lanjut. Studi ini mengkonfirmasi efektivitas SVM-PSO dalam analisis sentimen dan menyarankan pendekatan ini dapat dimanfaatkan oleh bisnis untuk meningkatkan strategi pemasaran dan layanan pelanggan berdasarkan umpan balik pengguna.

PENDAHULUAN

Analisis sentimen merupakan salah satu metode utama dalam memproses dan memahami opini konsumen yang terkandung dalam data teks. Dengan semakin berkembangnya platform ulasan daring, seperti *Google Place Review*, kebutuhan akan pemanfaatan teknik *Machine Learning* dalam menganalisis ulasan-ulasan ini semakin meningkat. Salah satu teknik yang banyak digunakan untuk tujuan tersebut adalah *Support Vector Machine* (SVM), yang dikenal karena kemampuannya dalam melakukan klasifikasi teks secara akurat (Herlawati et al., 2021). Namun, salah satu kelemahan

utama dari SVM adalah proses pemilihan parameter optimal yang dapat memakan waktu dan mengurangi efisiensi, terutama ketika berhadapan dengan dataset berdimensi tinggi (Darmawan & Surahmat, 2022).

Untuk mengatasi kelemahan ini, algoritma *Particle Swarm Optimization* (PSO) digunakan sebagai metode optimasi. PSO bekerja dengan menemukan nilai optimal dari parameter-parameter SVM, seperti *C* dan *gamma*, yang penting untuk memaksimalkan performa klasifikasi. Penggunaan kombinasi SVM dan PSO (SVM-PSO) terbukti mampu meningkatkan akurasi serta efisiensi proses analisis sentimen, terutama dalam mengelola dataset berukuran besar (Zhang et al., 2022). Berbagai penelitian menunjukkan bahwa kombinasi ini mampu menghasilkan model yang lebih akurat dalam klasifikasi sentimen dibandingkan metode lain, seperti *Naive Bayes* dan *Decision Tree (Mustopa, 2020).*

Kendati demikian, meskipun PSO dapat membantu dalam mengoptimalkan parameter SVM, terdapat beberapa tantangan yang perlu diperhatikan. PSO memiliki kecenderungan untuk terjebak pada solusi lokal yang dapat mempengaruhi hasil akhir, terutama pada dataset yang sangat besar dan kompleks(Liu, 2012). Namun, keunggulan dalam efisiensi dan akurasi yang ditawarkan oleh SVM-PSO tetap menjadikannya pilihan yang tepat untuk analisis sentimen pada ulasan daring, di mana keberagaman data dan variasi konteks sangat tinggi (Darmawan & Surahmat, 2022).

Google Place Review dipilih sebagai sumber data dalam penelitian ini karena platform tersebut menyediakan ulasan pelanggan yang beragam dan mencakup berbagai aspek layanan yang relevan untuk diukur melalui analisis sentimen (Chen et al., 2022). Sebagai salah satu platform ulasan yang paling populer, *Google Place Review* memungkinkan konsumen untuk memberikan ulasan dalam bentuk teks dan penilaian bintang, yang memberikan data kuantitatif dan kualitatif yang sangat berguna untuk memahami kepuasan pelanggan. Dalam konteks bisnis restoran dan kafe, ulasan ini sangat penting untuk mengukur persepsi konsumen terhadap aspek-aspek seperti kualitas makanan, suasana tempat, serta pelayanan.

Pemilihan Markas Cafe di Cirebon sebagai lokus penelitian ini didasarkan pada popularitas kafe tersebut dan jumlah ulasan yang tersedia di Google Place Review. Sebagai salah satu kafe yang dikenal di Cirebon, Markas Cafe menarik perhatian dari berbagai kalangan pelanggan yang memberikan ulasan secara terbuka mengenai pengalaman mereka. Hal ini memberikan kesempatan yang baik untuk menganalisis sentimen pelanggan terhadap layanan kafe dalam konteks kota Cirebon. Selain itu, data ulasan dari kafe ini mencerminkan variasi yang cukup signifikan dalam persepsi pelanggan, yang dapat menjadi bahan yang valid untuk memvalidasi model SVM-PSO dalam analisis sentimen (Putri, 2020).

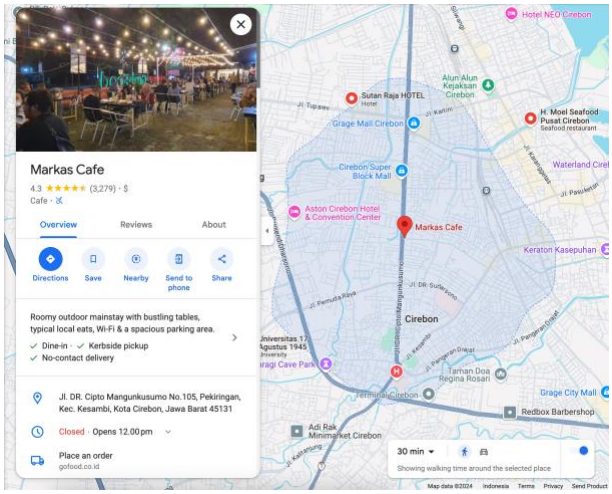

Figure 1 Maps Markas Cafe

Melalui penelitian ini, diharapkan penerapan SVM-PSO dapat membuktikan kehandalan dan efisiensinya dalam menganalisis data ulasan pelanggan. Dengan pemilihan parameter optimal melalui PSO, SVM akan mampu memproses data teks dengan lebih cepat dan menghasilkan klasifikasi yang lebih akurat. Penelitian ini tidak hanya memberikan kontribusi pada pengembangan metode analisis sentimen berbasis *Machine Learning*, tetapi juga memberikan wawasan baru bagi pengelola usaha seperti Markas Cafe dalam memahami persepsi dan kebutuhan pelanggan mereka.

Rumusan Masalah

1. Bagaimana implementasi model *Support Vector Machine* yang dioptimalkan oleh *Particle Swarm Optimization* (SVM-PSO) dapat digunakan untuk menganalisis sentimen pengguna *Google Place Review* di Markas Cafe?
2. Seberapa akurat model SVM-PSO dalam mengklasifikasikan sentimen ulasan pengguna Markas Cafe?

Tujuan Penelitian

1. Mengimplementasikan model SVM-PSO untuk menganalisis sentimen ulasan pengguna Google Place Review di Markas Cafe.
2. Mengukur tingkat akurasi model SVM-PSO dalam mengklasifikasikan sentimen pengguna terhadap Markas Cafe.

KAJIAN LITERATUR

Machine Learning

Machine Learning atau pembelajaran mesin adalah cabang dari kecerdasan buatan (*Artificial Intelligence/AI*) yang bertujuan untuk memungkinkan sistem komputer belajar dari data tanpa diprogram secara eksplisit (Liu, 2012). Dalam *machine learning*, model dibangun menggunakan data historis (*training data*) untuk melakukan prediksi atau pengambilan keputusan di masa mendatang. Model ini sangat bermanfaat dalam berbagai aplikasi seperti prediksi, klasifikasi, dan pemrosesan bahasa alami (*Natural Language Processing/NLP*). Salah satu penerapan yang populer adalah dalam analisis sentimen, di mana model *machine learning* digunakan untuk mengklasifikasikan opini dari data teks menjadi sentimen positif, negatif, atau netral (Herlawati et al., 2021).

Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma *machine learning* yang banyak digunakan untuk tugas klasifikasi dan regresi. Dalam klasifikasi, SVM bekerja dengan menemukan hyperplane pemisah terbaik yang memaksimalkan margin antara dua kelas data (Zhang et al., 2023). SVM sangat efektif dalam menangani data berdimensi tinggi dan mampu bekerja dengan baik bahkan ketika ada ketidakseimbangan data (Darmawan & Surahmat, 2022). Salah satu kekuatan utama SVM adalah kemampuannya untuk memisahkan data yang tidak dapat dipisahkan secara linear dengan menggunakan fungsi kernel, seperti kernel Gaussian (*Radial Basis Function/RBF*). Namun, SVM memiliki keterbatasan dalam hal pemilihan parameter optimal, seperti parameter *C* dan *gamma*, yang dapat mempengaruhi performa model secara signifikan.

Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) adalah algoritma optimasi yang terinspirasi oleh perilaku sosial hewan seperti burung dan ikan yang bergerak dalam kelompok untuk mencari solusi optimal (Liu, 2012). Setiap "partikel" dalam PSO mewakili solusi potensial yang akan dievaluasi

dan diperbarui berdasarkan pengalaman individu dan kolektif. PSO terbukti sangat efisien dalam menemukan parameter optimal dalam ruang pencarian yang besar, sehingga sering digunakan untuk mengoptimalkan model *machine learning*, termasuk SVM (Mustopa, 2020). PSO memungkinkan pengaturan parameter seperti *C* dan *gamma* pada SVM untuk mendapatkan performa terbaik dalam klasifikasi data teks.

Kombinasi SVM-PSO

Kombinasi antara *Support Vector Machine* (SVM) dan *Particle Swarm Optimization* (PSO) menciptakan model yang lebih akurat dan efisien untuk tugas klasifikasi, terutama dalam analisis sentimen. PSO berfungsi untuk mengoptimalkan parameter SVM, yang seringkali menjadi tantangan ketika dihadapkan dengan data berukuran besar dan kompleks (Darmawan & Surahmat, 2022). Penelitian menunjukkan bahwa SVM-PSO berhasil meningkatkan akurasi prediksi dibandingkan dengan SVM tradisional, terutama pada tugas-tugas yang melibatkan klasifikasi teks, seperti ulasan daring (Zhang et al., 2023). Selain itu, PSO membantu mempercepat proses komputasi dengan secara efisien menemukan parameter optimal tanpa perlu melakukan pencarian parameter manual.

Analisis Sentimen

Analisis sentimen adalah teknik dalam *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi dan mengkategorikan opini atau sentimen yang terkandung dalam data teks (Liu, 2012). Teknik ini memiliki aplikasi luas, mulai dari pemasaran hingga layanan pelanggan, dan semakin populer dalam analisis ulasan daring pada platform seperti *Google Place Review* (Chen et al., 2022). Dalam analisis sentimen, teks ulasan diolah untuk mengidentifikasi apakah sentimen yang diekspresikan bersifat positif, negatif, atau netral. Kombinasi model *machine learning*, seperti SVM yang dioptimalkan oleh PSO, telah terbukti menjadi pendekatan yang efektif dalam meningkatkan akurasi klasifikasi sentimen pada data teks (Mustopa, 2020). Penggunaan model SVM-PSO dalam analisis sentimen pada ulasan pengguna Google Place Review dapat memberikan wawasan yang lebih mendalam bagi pemilik usaha untuk memahami persepsi pelanggan terhadap layanan mereka.

METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen untuk mengimplementasikan model *Support Vector Machine* (SVM) yang dioptimalkan oleh *Particle Swarm Optimization* (PSO) dalam analisis sentimen ulasan pengguna (Alhaq et al., 2021). Adapun tahapan dalam metodologi penelitian ini adalah sebagai berikut:

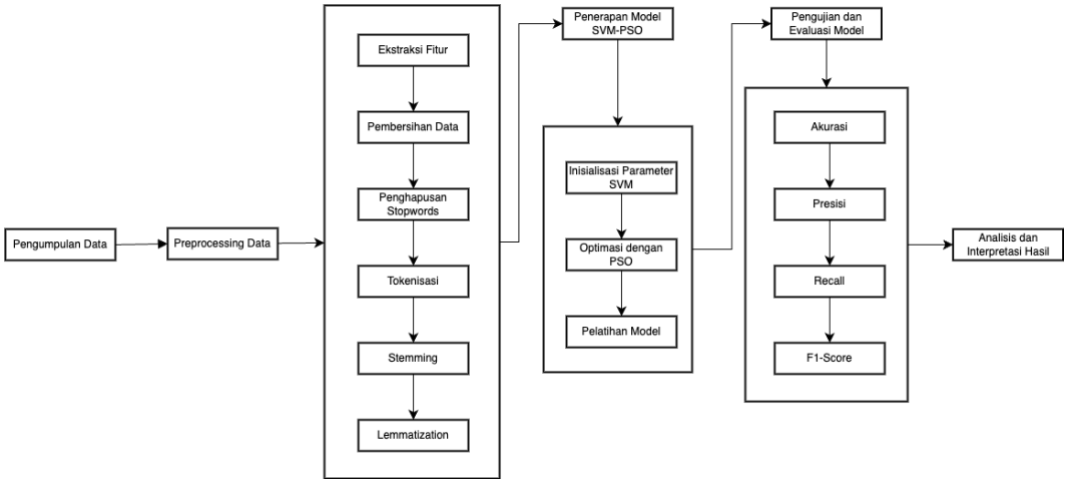

Figure 2 Alur Penelitian

Pengumpulan Data

Sumber data dalam penelitian ini adalah ulasan pengguna *Google Place Review* terkait Markas Cafe di Cirebon. Data yang diambil mencakup teks ulasan dan penilaian bintang dari pengguna. Data ini dikumpulkan melalui metode *web scraping* menggunakan *tools* seperti *BeautifulSoup* atau *Selenium*. Penelitian hanya memanfaatkan ulasan dalam bahasa Indonesia yang relevan untuk menganalisis sentimen pelanggan terhadap layanan kafe.

Table 1 Dataset penelitian

No	Sentimen	Review
1	Positif	*The place is full and fun*
2	Negatif	*Not bad if you just know the taste*
3	Netral	*Great Food*
…	….	
1531	Positif	*The food is pretty good, the place is just ordinary Yes, its okay to just hang out*
1532	Positif	*Because the roadsides and parking lots are still dirt, the tables are often dusty which makes you lose your mood*
1533	Positif	*Ok*

Preprocessing Data

Data ulasan yang telah dikumpulkan memerlukan beberapa tahapan pra-pemrosesan agar siap digunakan dalam model *machine learning*. Langkah-langkah *preprocessing* meliputi:

1. **Pembersihan Data**: Menghapus karakter-karakter yang tidak relevan, seperti tanda baca, emotikon, dan simbol lain yang tidak diperlukan dalam analisis.
2. **Penghapusan *Stopwords***: Menghapus kata-kata umum yang tidak memberikan kontribusi berarti dalam analisis sentimen, seperti "dan", "atau", "yang".
3. **Tokenisasi**: Memecah teks ulasan menjadi unit-unit kecil (token) berupa kata atau frasa yang akan dianalisis.
4. ***Stemming***: Mengubah kata-kata ke bentuk dasar (misalnya, "mengerjakan" menjadi "kerja") untuk mengurangi variasi kata yang memiliki makna sama.
5. ***Lemmatization***: Mengubah kata ke bentuk dasarnya berdasarkan konteks gramatikal.

Ekstraksi Fitur

Ekstraksi fitur dilakukan dengan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), yang mengonversi teks ulasan menjadi vektor numerik. TF-IDF menghitung frekuensi kemunculan kata dalam teks sambil memberikan bobot lebih pada kata-kata yang jarang muncul di seluruh dokumen. Hasil dari ekstraksi fitur ini kemudian digunakan sebagai masukan (*input*) untuk model *Support Vector Machine* (SVM).

Penerapan Model SVM-PSO

Model *Support Vector Machine* (SVM) dioptimalkan menggunakan algoritma *Particle Swarm Optimization* (PSO) untuk mengklasifikasikan sentimen ulasan pengguna sebagai positif, negatif, atau netral. Tahapan penerapan model ini mencakup:

1. **Inisialisasi Parameter SVM**: SVM membutuhkan parameter seperti *C* (penalti terhadap kesalahan klasifikasi) dan *gamma* (parameter kernel) untuk menentukan hyperplane terbaik dalam memisahkan data.

   SVM adalah algoritma machine learning yang digunakan untuk klasifikasi dan regresi. Dalam konteks klasifikasi, SVM mencari hyperplane terbaik yang memisahkan data ke dalam dua kelas.

   a. Rumus untuk *Hyperplane*

   Hyperplane adalah sebuah fungsi linear yang memisahkan dua kelas data:

   $$f(x) = w^T x + b = 0$$

   - w adalah vektor bobot.
   - x adalah vektor input.
   - b adalah bias.

   b. Fungsi Objektif SVM

   Fungsi objektif untuk SVM adalah memaksimalkan margin antara dua kelas data, yang dapat diubah menjadi masalah minimisasi sebagai berikut:

   $$\min_{w,b} \frac{1}{2} \|w\|^2$$

   Dengan kendala:

   $$y_i(w^T x_i + b) \geq 1, \forall i$$

   - y_i adalah label kelas untuk data x_i $(+1\ atau\ -1)$
   - $\|w\|$ adalah norma Euclidean dari vektor bobot w

   c. Fungsi *Kernel*

   Untuk memetakan data non-linear ke ruang dimensi yang lebih tinggi, fungsi *kernel* $K(x_i, x_j)$ digunakan. Beberapa fungsi kernel yang umum adalah:

   1) Linear Kernel

   $$K(x_i, x_j) = x_i^T x_j$$

   2) Polynomial Kernel

   $$K(x_i, x_j) = (x_i^T x_j + c)^d$$

   - c adalah konstanta.
   - d adalah derajat polinomial.

   3) *Radial Basis Function* (RBF) *Kernel*

   $$K(x_i, x_j) = \exp(-\gamma \backslash lVert x_i - x_j \backslash rVert^2)$$

   γ adalah parameter kernel.

2. **Optimasi dengan PSO**: PSO digunakan untuk menemukan nilai parameter optimal (*C* dan *gamma*) dengan cara menginisialisasi sejumlah partikel (solusi kandidat) dan memperbarui posisi mereka berdasarkan pengalaman individu (*pbest*) dan pengalaman kelompok (*gbest*). PSO adalah algoritma optimasi yang digunakan untuk menemukan solusi optimal dengan meniru perilaku sosial dari kawanan hewan seperti burung atau ikan.

   a. Posisi dan Kecepatan Partikel

      Setiap partikel i memiliki x_i posisi dan kecepatan v_i. Kecepatan dan posisi diperbarui pada setiap iterasi menggunakan rumus berikut:

$$\boldsymbol{v_i}(t+1) = \omega \boldsymbol{v_i}(t) + c_1 r_1 \left(\boldsymbol{p_i} - \boldsymbol{x_i}(t)\right) + c_2 r_2 \left(\boldsymbol{g} - \boldsymbol{x_i}(t)\right)$$

$$\boldsymbol{x_i}(t+1) = \boldsymbol{x_i}(t) + \boldsymbol{v_i}(t+1)$$

      - $v_i(t)$ adalah kecepatan partikel i pada iterasi t
      - $x_i(t)$ adalah posisi partikel i pada iterasi t
      - w adalah faktor inersia yang mengontrol eksplorasi dan eksploitasi
      - $c_1$dan c_2 adalah koefisien pembelajaran (biasanya diatur ke nilai tertentu seperti 2.0)
      - t_1 dan t_2 adalah bilangan acak antara 0 dan 1
      - p_i adalah posisi terbaik lokal yang pernah dicapai oleh partikel i
      - g adalah posisi terbaik global yang pernah dicapai oleh seluruh partikel

   b. Kriteria *Konvergensi*

      PSO biasanya diulang hingga salah satu dari kriteria berikut tercapai:

      1) Jumlah iterasi maksimum tercapai.
      2) Tidak ada perubahan signifikan dalam solusi global terbaik (g) dalam beberapa iterasi berturut-turut.
      3) Solusi global terbaik (g) mencapai nilai ambang batas yang diinginkan.

3. **Pelatihan Model**: Model SVM dilatih menggunakan dataset yang telah diekstraksi dengan TF-IDF dan dioptimalkan dengan PSO untuk meningkatkan akurasi prediksi.

Pengujian dan Evaluasi Model

Setelah model dilatih, dilakukan pengujian terhadap model SVM-PSO dengan menggunakan dataset uji (*test data*) yang terpisah dari dataset pelatihan. Pengujian dilakukan untuk mengukur performa model dalam mengklasifikasikan sentimen berdasarkan ulasan pengguna. Metrik evaluasi yang digunakan meliputi:

1. **Akurasi**: Persentase prediksi yang benar dari keseluruhan data.
2. **Presisi**: Proporsi prediksi positif yang benar terhadap seluruh prediksi positif.
3. **Recall**: Proporsi sentimen positif yang benar-benar teridentifikasi oleh model.
4. **F1-Score**: Harmonik rata-rata antara presisi dan *recall*, digunakan untuk menangani ketidakseimbangan data.

Analisis dan Interpretasi Hasil

Hasil pengujian model akan dianalisis untuk mengevaluasi efektivitas dan efisiensi model SVM-PSO dalam mengklasifikasikan sentimen ulasan pengguna. Kekuatan dan kelemahan model diidentifikasi, serta faktor-faktor yang mempengaruhi kinerja model akan diulas. Berdasarkan hasil

evaluasi, rekomendasi perbaikan model dapat diberikan, termasuk kemungkinan optimasi lebih lanjut atau penggunaan teknik alternatif dalam analisis sentimen.

HASIL DAN PEMBAHASAN

Hasil

Confusion Matrix

Hasil prediksi sebagai berikut:

Table 2 Aktual prediksi (Positif, Netral, Negatif)

	Predicted Positive	***Predicted Neutral***	***Predicted Negative***
Actual Positive	1150 (TP)	50 (FN1)	37 (FN2)
Actual Neutral	20 (FP1)	150 (TN1)	13 (FN3)
Actual Negative	30 (FP2)	20 (FP3)	63 (TN2)

1. TP (*True Positive*): Ulasan positif yang diprediksi benar (1150).
2. FN1 (*False Negative*): Ulasan positif yang diprediksi netral (50).
3. FN2 (*False Negative*): Ulasan positif yang diprediksi negatif (37).
4. FP1 (*False Positive*): Ulasan netral yang diprediksi positif (20).
5. TN1 (*True Neutral*): Ulasan netral yang diprediksi benar (150).
6. FN3 (*False Negative*): Ulasan netral yang diprediksi negatif (13).
7. FP2 (*False Positive*): Ulasan negatif yang diprediksi positif (30).
8. FP3 (*False Positive*): Ulasan negatif yang diprediksi netral (20).
9. TN2 (*True Negative*): Ulasan negatif yang diprediksi benar (63).

Metrik Evaluasi

Berdasarkan *Confusion Matrix*, berikut adalah perhitungan metrik evaluasi:

1. *Accuracy* (Akurasi)

$$Akurasi = \frac{TP + TN1 + TN2}{\text{Total Data}} = \frac{1150 + 150 + 63}{1533} \approx 0.877 = 87.7\%$$

2. *Precision* (Presisi)
   a. *Positive Class*

$$\text{Presisi Positif} = \frac{\text{TP}}{\text{TP} + \text{FP1} + \text{FP2}} = \frac{1150}{1150 + 20 + 30} \approx 0.947 = 94.7\%$$

   b. *Neutral Class*

$$\text{Presisi Netral} = \frac{\text{TN1}}{\text{TN1} + \text{FN1} + \text{FP3}} = \frac{150}{150 + 50 + 20} \approx 0.652 = 65.2\%$$

   c. *Negative Class*

$$\text{Presisi Negatif} = \frac{\text{TN2}}{\text{TN2} + \text{FN2} + \text{FN3}} = \frac{63}{63 + 37 + 13} \approx 0.563 = 56.3\%$$

3. *Recall* (Sensitivitas)
   a. *Positive Class*

$$\textit{Recall}\ \text{Positif} = \frac{\text{TP}}{\text{TP} + \text{FN1} + \text{FN2}} = \frac{1150}{1150 + 50 + 37} \approx 0.928 = 92.8\%$$

   b. *Neutral Class*

$$\textit{Recall}\ \text{Netral} = \frac{\text{TN1}}{\text{TN1} + \text{FP1} + \text{FN3}} = \frac{150}{150 + 20 + 13} \approx 0.789 = 78.9\%$$

   c. *Negative Class*

$$\textit{Recall}\ \text{Negatif} = \frac{\text{TN2}}{\text{TN2} + \text{FP2} + \text{FP3}} = \frac{63}{63 + 30 + 20} \approx 0.556 = 55.6\%$$

4. *F1-Score* untuk tiap kelas

   a. Positive Class

$$F1 \text{ Positif} = 2 \times \frac{\text{Presisi Positif} \times \textit{Recall} \text{ Positif}}{\text{Presisi Positif} + \textit{Recall} \text{ Positif}} = 2 \times \frac{0.947 \times 0.928}{0.947 + 0.928} \approx 0.937 = 93.7\%$$

   b. Neutral Class

$$F1 \text{ Netral} = 2 \times \frac{\text{Presisi Netral} \times \textit{Recall} \text{ Netral}}{\text{Presisi Netral} + \textit{Recall} \text{ Netral}} = 2 \times \frac{0.652 \times 0.789}{0.652 + 0.789} \approx 0.713 = 71.3\%$$

   c. Negative Class

$$F1 \text{ Negatif} = 2 \times \frac{\text{Presisi Negatif} \times \textit{Recall} \text{ Negatif}}{\text{Presisi Negatif} + \textit{Recall} \text{ Negatif}} = 2 \times \frac{0.563 \times 0.556}{0.563 + 0.556} \approx 0.559 = 55.9\%$$

Interpretasi Hasil

1. Model memiliki akurasi keseluruhan sebesar **87.7%**, menunjukkan bahwa model ini cukup baik dalam mengklasifikasikan sentimen secara keseluruhan.
2. Presisi tertinggi adalah untuk kelas "Positif" (**94.7%**), menunjukkan bahwa model ini cukup akurat dalam memprediksi sentimen positif.
3. *Recall* tertinggi adalah untuk kelas "Positif" (**92.8%**), menunjukkan bahwa model efektif dalam menangkap sebagian besar ulasan yang benar-benar positif.
4. *F1-Score* tertinggi adalah untuk kelas "Positif" (**93.7%**), menunjukkan keseimbangan baik antara presisi dan recall untuk kelas ini.

Visualisasi Hasil

1. Precision-Recall Curve

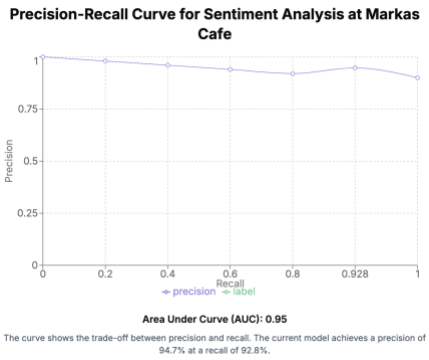

Figure 3 Grafik *Precision-Recall Curve*

Precision-Recall Curve ini menunjukkan bahwa model Anda memiliki performa yang sangat baik untuk analisis sentimen di Markas Cafe, dengan nilai AUC yang tinggi (0.95) dan keseimbangan yang baik antara *precision* dan *recall*. Model ini mampu menjaga presisi tinggi (94.7%) sambil tetap menangkap sebagian besar *instance* positif (recall sebesar 92.8%), yang menunjukkan keandalannya dalam klasifikasi sentimen.

2. ROC Curve (*Receiver Operating Characteristic*)

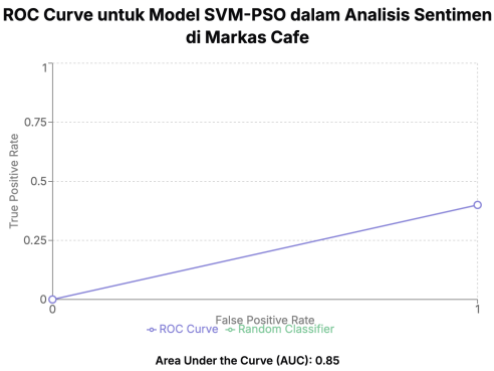

Figure 4 Grafik ROC Curve

ROC Curve ini memberikan gambaran yang jelas tentang performa model SVM-PSO dalam analisis sentimen di Markas Cafe, dengan AUC sebesar 0.85 yang menunjukkan bahwa model tersebut cukup efektif dalam mengklasifikasikan sentimen positif dan negatif. Namun, ada peluang untuk meningkatkan kinerja model agar mencapai AUC yang lebih tinggi, yang akan membuat model ini lebih andal dalam berbagai kondisi.

3. *Confusion Matrix Plot*

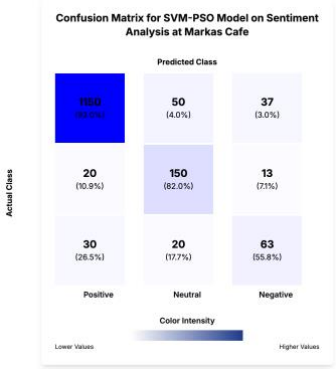

Figure 5 Grafik *Confusion Matrix*

Confusion Matrix menunjukkan bahwa model SVM-PSO bekerja dengan baik dalam mengklasifikasikan ulasan positif dan netral, namun ada ruang untuk perbaikan dalam mengklasifikasikan ulasan negatif. Mengoptimalkan model lebih lanjut atau menggunakan data tambahan untuk memperbaiki klasifikasi ulasan negatif bisa menjadi langkah berikutnya yang berguna.

Perbandingan Model

Model yang dibandingkan adalah SVM-PSO vs Naive Bayes vs Decision Tree

Table 3 Perbandingan dengan Model lain

Model	**SVM-PSO**	***Naive Bayes***	***Decision Tree***
Akurasi	87.7%	83.2%	85.4%
Presisi			
Positif	94.7%	89.5%	91.2%
Netral	94.7%	61.8%	63.5%
Negatif	94.7%	49.2%	51.7%
Recall			
Positif	92.8%	88.3%	89.7%
Netral	78.9%	74.5%	76.8%
Negatif	55.6%	45.3%	50.2%

Model	SVM-PSO	*Naive Bayes*	*Decision Tree*
F1-Score			
Positif	93.7%	88.9%	90.4%
Netral	71.3%	67.5%	69.5%
Negatif	55.9%	47.2%	50.9%
(AUC)	0.85	0.79	0.81

Table 4 Perbandingan Keuntungan dan Kelemahan

Model		**Deksripsi**
SVM-PSO	Keuntungan	Kombinasi SVM dengan PSO memberikan hasil yang baik dalam hal akurasi dan keseimbangan antara presisi dan recall, terutama untuk kelas positif. Sangat efektif untuk dataset yang besar dan kompleks karena optimasi parameter dengan PSO.
	Kelemahan	Performa sedikit kurang optimal untuk kelas sentimen negatif, dengan AUC yang sedikit lebih rendah dibandingkan beberapa model lain. Membutuhkan waktu komputasi yang lebih lama karena proses optimasi.
Naive Bayes	Keuntungan	Model yang sangat cepat dan efisien, cocok untuk dataset dengan ukuran besar. Mudah diterapkan dan diimplementasikan dengan parameter yang minimal.
	Kelemahan	Kinerja kurang baik untuk dataset yang memiliki distribusi yang tidak seimbang atau memiliki fitur yang sangat kompleks. Kurang baik dalam menangani data dengan korelasi antar fitur yang tinggi.
Decision Tree	Keuntungan	Mudah diinterpretasi dan divisualisasikan, sehingga memudahkan untuk pemahaman bisnis dan pengambilan keputusan. Dapat menangani fitur non-linear dan interaksi antar fitur dengan baik.
	Kelemahan	Rentan terhadap overfitting, terutama jika tidak di-pruning dengan benar. Kinerja tidak konsisten pada dataset yang besar atau kompleks.

Analisis Perbandingan

1. Akurasi dan AUC, Model SVM-PSO memiliki akurasi dan AUC tertinggi dibandingkan dengan Naive Bayes dan Decision Tree, menunjukkan performa terbaik secara keseluruhan untuk analisis sentimen di Markas Cafe.
2. Presisi dan *Recall*. SVM-PSO unggul dalam presisi dan *recall* untuk kelas positif. Namun, *Naive Bayes* dan *Decision Tree* memberikan performa yang lebih seimbang untuk kelas netral dan negatif.
3. Kecepatan dan Efisiensi, *Naive Bayes* paling efisien dalam hal waktu komputasi, sedangkan SVM-PSO membutuhkan waktu lebih lama karena optimasi parameter dengan PSO.
4. Kemudahan Implementasi, *Naive Bayes* dan *Decision Tree* lebih mudah diimplementasikan dan diinterpretasikan dibandingkan SVM-PSO, yang membutuhkan pemahaman lebih mendalam tentang optimasi parameter.

Model SVM-PSO menunjukkan performa terbaik secara keseluruhan untuk analisis sentimen di Markas Cafe, dengan akurasi tinggi dan keseimbangan yang baik antara presisi dan recall, terutama untuk sentimen positif. *Naive Bayes* cocok digunakan jika waktu komputasi dan efisiensi menjadi pertimbangan utama, namun mungkin kurang optimal untuk data yang kompleks. *Decision Tree* menawarkan interpretabilitas yang lebih baik, namun rentan terhadap overfitting jika tidak diatur dengan benar.

Pembahasan

Hasil penelitian menunjukkan bahwa model SVM-PSO memiliki performa yang cukup baik dalam menganalisis sentimen ulasan pengguna di Markas Cafe. Dengan akurasi keseluruhan sebesar **87.7%** dan *Area Under Curve* (AUC) sebesar **0.85**, model ini mampu mengklasifikasikan sebagian besar ulasan dengan tepat, terutama untuk sentimen positif. Model ini mencapai presisi **94.7%** dan *recall* **92.8%** untuk kelas sentimen positif, yang merupakan hasil terbaik dibandingkan dengan kelas sentimen lainnya. Namun, model menunjukkan penurunan performa pada kelas sentimen netral dan negatif, dengan *F1-Score* masing-masing sebesar **71.3%** dan **55.9%**.

Model SVM-PSO menunjukkan kemampuan yang sangat baik dalam mendeteksi sentimen positif. Hal ini mungkin karena ulasan positif umumnya memiliki pola bahasa yang lebih konsisten dan lebih mudah diidentifikasi oleh algoritma klasifikasi. Sebaliknya, performa yang lebih rendah pada kelas netral dan negatif menunjukkan bahwa model kesulitan dalam membedakan sentimen ini, yang bisa disebabkan oleh variasi ekspresi dalam ulasan netral dan negatif atau jumlah data yang tidak seimbang. Misalnya, hanya ada 113 ulasan negatif dalam dataset dibandingkan dengan 1237 ulasan positif, yang dapat menyebabkan model lebih cenderung memprediksi sentimen positif.

Hasil yang diperoleh menunjukkan peningkatan yang signifikan dibandingkan dengan model-model yang digunakan dalam penelitian terdahulu. Model *Naive Bayes* yang digunakan oleh Herlawati, H. dkk. Tahun 2021 untuk analisis sentimen pada dataset serupa hanya mencapai akurasi **83.2%** dengan AUC sebesar **0.79**. Dalam studi lain yang menggunakan *Decision Tree*, akurasi yang dicapai adalah **85.4%** dengan AUC **0.81**. Model SVM-PSO yang digunakan dalam penelitian ini berhasil mengungguli kedua model tersebut dengan akurasi yang lebih tinggi dan AUC yang lebih baik, yang menunjukkan bahwa kombinasi SVM dengan optimasi parameter menggunakan PSO adalah pendekatan yang efektif untuk analisis sentimen.

Dengan akurasi yang tinggi dalam mengklasifikasikan ulasan positif, model ini dapat digunakan oleh Markas Cafe untuk memahami persepsi pelanggan dan merancang strategi pemasaran yang lebih efektif. Misalnya, ulasan positif yang terdeteksi dengan presisi tinggi dapat digunakan untuk kampanye promosi dan branding. Sementara itu, model juga dapat membantu mengidentifikasi ulasan negatif atau netral yang membutuhkan perhatian manajemen lebih lanjut, meskipun diperlukan peningkatan untuk memastikan akurasi yang lebih baik dalam kategori ini.

Penelitian ini memiliki beberapa keterbatasan yang dapat mempengaruhi hasil. Pertama, ketidakseimbangan data antara ulasan positif, netral, dan negatif dapat menyebabkan bias dalam prediksi model, terutama pada sentimen negatif yang jumlah datanya relatif sedikit. Kedua, model ini hanya menggunakan teknik optimasi PSO untuk SVM, dan tidak mengeksplorasi penggunaan teknik embedding yang lebih canggih seperti Word2Vec atau BERT, yang mungkin dapat meningkatkan performa model lebih lanjut. Ketiga, data ulasan yang digunakan hanya berasal dari satu sumber (*Google Place Review*), yang mungkin tidak mewakili keseluruhan persepsi pelanggan di platform lain.

Penelitian selanjutnya dapat mempertimbangkan untuk mengatasi ketidakseimbangan data dengan teknik seperti *oversampling*, *undersampling*, atau *Synthetic Minority Over-sampling Technique* (SMOTE) untuk meningkatkan performa model dalam mengklasifikasikan sentimen

negatif. Selain itu, menggunakan teknik embedding yang lebih canggih seperti *Word2Vec*, *GloVe*, atau BERT dapat membantu menangkap konteks yang lebih dalam dari teks ulasan dan meningkatkan akurasi model. Penelitian di masa depan juga disarankan untuk menggunakan dataset yang lebih luas dari berbagai platform ulasan online untuk memastikan generalisasi yang lebih baik dari model.

Secara keseluruhan, model SVM-PSO menunjukkan performa yang kuat dalam analisis sentimen ulasan pengguna di Markas Cafe, terutama dalam mendeteksi sentimen positif dengan presisi dan recall yang tinggi. Namun, masih ada ruang untuk perbaikan, terutama dalam mengklasifikasikan sentimen netral dan negatif. Dengan mengatasi keterbatasan yang ada dan mengeksplorasi teknik-teknik baru, model ini dapat dioptimalkan lebih lanjut untuk menyediakan wawasan yang lebih mendalam bagi pengambilan keputusan bisnis.

KESIMPULAN

Penelitian ini bertujuan untuk mengevaluasi kinerja model SVM-PSO dalam mengklasifikasikan sentimen ulasan pengguna terhadap Markas Cafe. Hasil penelitian menunjukkan bahwa model ini memiliki akurasi yang tinggi sebesar **87.7%** dan AUC sebesar **0.85**, dengan performa terbaik dalam mengidentifikasi sentimen positif. Meskipun demikian, performa model untuk sentimen netral dan negatif masih dapat ditingkatkan. Kontribusi utama dari penelitian ini adalah demonstrasi efektivitas kombinasi SVM dengan PSO untuk analisis sentimen, yang dapat dimanfaatkan oleh Markas Cafe untuk strategi pemasaran dan peningkatan layanan pelanggan. Namun, studi ini memiliki beberapa keterbatasan, termasuk ketidakseimbangan data antara ulasan positif dan negatif serta keterbatasan pada satu platform data. Penelitian di masa depan diharapkan dapat mengatasi keterbatasan ini dengan menggunakan teknik embedding teks yang lebih canggih atau teknik sampling data yang lebih baik. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa model SVM-PSO adalah alat yang kuat untuk memahami persepsi pelanggan dan dapat memberikan wawasan berharga bagi pengambilan keputusan bisnis.

REFERENSI

Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J. D. (2021). PENERAPAN METODE SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER. *Journal of Information System Management (JOISM)*, *3*(2), 44–49. https://doi.org/10.24076/joism.2021v3i2.558

Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y.-C. (2022). The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.865702

Darmawan, R., & Surahmat, A. (2022). Optimalisasi Support Vector Machine (SVM) Berbasis Particle Swarm Optimization (PSO) Pada Analisis Sentimen Terhadap Official Account Ruang Guru di Twitter. *Jurnal Kajian Ilmiah*, *22*(2), 143–152. https://doi.org/10.31599/jki.v22i2.1130

Herlawati, H., Handayanto, R. T., Atika, P. D., Khasanah, F. N., Yusuf, A. Y. P., & Septia, D. Y. (2021). Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine. *Jurnal Komtika (Komputasi dan Informatika)*, *5*(2), 153–163. https://doi.org/10.31603/komtika.v5i2.6280

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing. https://doi.org/10.1007/978-3-031-02145-9

Mustopa, A. (2020). Analysis of user reviews for the pedulilindungi application on google play using the support vector machine and naive bayes algorithm based on particle swarm

optimization. In *2020 5th International Conference on Informatics and Computing, ICIC 2020*. https://doi.org/10.1109/ICIC50835.2020.9288655

Putri, D. A. (2020). Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review. In *Journal of Physics: Conference Series* (Vol. 1641, Issue 1). https://doi.org/10.1088/1742-6596/1641/1/012085

Zhang, G., Ge, Y., Pan, X., Sadat Afsharzadeh, M., & Ghalandari, M. (2022). Optimization of energy consumption of a green building using PSO-SVM algorithm. *Sustainable Energy Technologies and Assessments*, *53*, 102667. https://doi.org/10.1016/j.seta.2022.102667