

Analisis Sentimen Ulasan Pelanggan Tomoro Coffee Menggunakan Algoritma Support Vector Machine (SVM)

Yustitio Caesar¹, Dwijananda Galih Prameswara²

^{1,2}Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

E-mail: ¹siosanmar1234@gmail.com, ²dwigapra011@gmail.com

Abstrak – Tomoro Coffee memiliki banyak ulasan pelanggan di Google Maps sehingga sulit dianalisis secara manual. Penelitian ini bertujuan menganalisis sentimen ulasan pelanggan Tomoro Coffee menggunakan algoritma Support Vector Machine (SVM). Data diperoleh dari Google Maps melalui SerpAPI sebanyak 1.829 ulasan. Tahapan penelitian meliputi preprocessing teks, ekstraksi fitur menggunakan TF-IDF, serta klasifikasi sentimen dengan SVM kernel linear dan penanganan ketidakseimbangan data menggunakan SMOTE. Evaluasi dilakukan menggunakan confusion matrix dengan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan model SVM mencapai akurasi 96,17% dengan performa sangat baik pada kelas sentimen positif, meskipun kinerja pada kelas netral masih rendah. Hasil ini menunjukkan bahwa SVM efektif untuk analisis sentimen ulasan pelanggan dan dapat digunakan sebagai dasar evaluasi peningkatan kualitas layanan.

Kata Kunci — Analisis sentimen, SVM, Tomoro Coffee

1. PENDAHULUAN

Kopi adalah salah satu minuman yang banyak dikonsumsi di dunia [1]. Indonesia dikenal sebagai salah satu pembuat kopi terbesar di dunia, dengan berbagai jenis dan variasi kopi yang tersebar di berbagai wilayah. Kopi juga termasuk komoditas utama di sektor perkebunan di Indonesia, dengan nilai ekonomi yang tinggi, serta konsumsi kopi di Indonesia mencapai 3,4% pertahunnya [2]. Salah satu kedai kopi yang terkenal di Indonesia adalah Tomoro Coffee, dengan konsep inovatif yang memanfaatkan teknologi digital. Seiring meningkatnya persaingan industri kedai kopi, Tomoro Coffee diwajibkan untuk meningkatkan kualitas produk serta pelayanan.

Permasalahan yang dihadapi oleh Tomoro Coffee adalah ulasan pelanggan yang tersebar dari berbagai platform digital khususnya Google Maps. Jumlah ulasan yang banyak menyulitkan perusahaan dalam menganalisis tanggapan pelanggan secara manual. Akibatnya, informasi terkait kepuasan pelanggan, kualitas produk sering terabaikan.

Untuk mengatasi permasalahan tersebut, penelitian ini memanfaatkan teknik analisis sentimen berbasis *text mining*. Analisis sentimen yang dilakukan menggunakan algoritma Support Vector Machine (SVM), yang mampu mengklasifikasi ulasan ke dalam kategori positif, negatif, dan netral secara otomatis.

Penelitian sebelumnya mengenai analisis sentimen ulasan kedai kopi. Pada penelitian oleh azhar [3], analisis sentimen ulasan kedai kopi. Metode yang digunakan dalam penelitian tersebut adalah Naive Bayes dengan menerapkan Algoritma Genetika sebagai seleksi fiturnya. Hasil dari penelitian tersebut menunjukkan bahwa metode Naive Bayes memiliki akurasi 90% sedangkan setelah menggunakan Algoritma Genetika meningkat menjadi 94%. Analisis sentimen pada ulasan pengguna aplikasi Starbucks juga dilakukan pada penelitian oleh Palimbani [4], yang bertujuan untuk mengevaluasi performa aplikasi mobile Starbucks. Metode yang digunakan dalam penelitian ini adalah Support Vector Machine (SVM). Hasil dari penelitian tersebut memiliki akurasi 88%.

Berdasarkan penelitian-penelitian sebelumnya, penelitian ini diharapkan memberikan kontribusi dalam penerapan analisis sentimen terhadap ulasan pelanggan. Selain itu, penelitian ini diharapkan dapat menjadi bahan evaluasi dan pertimbangan untuk meningkatkan kualitas pelayanan.

2. METODE PENELITIAN

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa ulasan pelanggan Tomoro Coffee yang diperoleh dari Google Maps. Pengumpulan data dilakukan secara otomatis menggunakan *Application Programming Interface* (API) SerpAPI, yang berfungsi untuk mengambil data ulasan Google Maps secara tersusun.

Proses pengambilan data dilakukan menggunakan bahasa pemrograman Python, dengan memanfaatkan library *serpapi*, *pandas*, dan *urllib*. Data ulasan diambil dan dikumpulkan menjadi beberapa atribut, yaitu nama, tanggal, rating, dan ulasan.

2.2 Preprocessing Data

Tahap *preprocessing* data dilakukan untuk membersihkan dan menyiapkan teks ulasan pelanggan sebelum digunakan pada proses analisis. Proses ini bertujuan mengurangi noise dan meningkatkan kualitas data sehingga kinerja algoritma klasifikasi sentimen menjadi lebih optimal.

Pada penelitian ini, *preprocessing* data dilakukan melalui beberapa tahapan sebagai berikut:

1. *Case Folding*
Tahap *case folding* dilakukan dengan mengubah seluruh teks ulasan menjadi huruf kecil [5].
2. *Pembersihan Teks (Text Cleaning)*
Pada tahap ini dilakukan penghapusan karakter yang tidak diperlukan seperti simbol, tanda baca, angka, dan karakter khusus lainnya. Selain itu spasi ganda yang terdapat dalam teks juga dihilangkan untuk menghasilkan teks yang lebih rapi [6].
3. *Tokenizing*
Proses ini dilakukan dengan memecah teks ulasan menjadi unit-unit kata atau token [7].
4. *Stopword Removal*
Tahap *stopword removal* dilakukan dengan menghapus kata-kata umum dalam bahasa Indonesia yang tidak memiliki makna signifikan terhadap analisis sentimen, seperti “dan”, “yang”, “di”, dan “ke” [8].
5. *Stemming*
Setiap kata diubah menjadi bentuk dasar, seperti “memuaskan” menjadi “puas”.

2.3 TF-IDF

Pada penelitian ini, proses ekstraksi fitur dilakukan menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengubah teks ulasan pelanggan menjadi numerik berdasarkan kata yang sering muncul dalam dokumen [9]. Pembobotan dilakukan dengan mengalikan frekuensi kemunculan kata dengan invers frekuensi dokumen, seperti ditunjukkan pada Persamaan 1.

2.4 Algoritma Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma *supervised learning* yang dilatih menggunakan data berlabel dan sering digunakan dalam klasifikasi. Algoritma SVM juga bekerja untuk menemukan *hyperplane* yang terbaik [10]. *Hyperplane* merupakan garis yang digunakan untuk membedakan data antar kelas, dengan menentukan jarak terdekat antara *hyperplane* dengan data dari masing-masing kelas. Fungsi keputusan SVM ditunjukkan pada Persamaan (1).

$$f(x) = w \cdot x + b \dots \dots \dots (1)$$

Dimana w adalah vektor bobot normal terhadap *hyperplane* b adalah bias. Proses tersebut bertujuan untuk meminimalkan nilai sehingga jarak antara kelas menjadi optimal.

2.5 Evaluasi Algoritma

Evaluasi algoritma dilakukan untuk mengukur kinerja model SVM dalam klasifikasi sentimen ulasan pelanggan. Evaluasi menggunakan *confusion matrix* sebagai perhitungan metrik akurasi, presisi, *recall*, dan *F1-score*.

2.5.1. Akurasi

Akurasi digunakan untuk mengukur persentase data yang berhasil diklasifikasi dengan benar oleh model. Akurasi dihitung menggunakan Persamaan (2) sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \dots \dots \dots (2)$$

2.5.2. Precision

Precision digunakan untuk mengukur ketepatan model dalam memprediksi suatu kelas. Perhitungan *precision* ditunjukkan pada Persamaan (3) sebagai berikut:

$$\text{Precision} = \frac{TP}{TP+FP} \dots \dots \dots (3)$$

2.5.3. Recall

Recall digunakan untuk mengukur model dalam mengenali seluruh data. Perhitungan *recall* ditunjukkan pada Persamaan (4) sebagai berikut:

$$\text{Recall} = \frac{TP}{TP+FN} \dots \dots \dots (4)$$

2.5.4. F1-Score

F1-Score digunakan untuk mengukur antara *precision* dan *recall* sehingga memberikan gambaran model secara menyeluruh. Perhitungan *F1-Score* ditunjukkan pada Persamaan (5) sebagai berikut:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \dots (5)$$

Keterangan:

TP (True Positive): Data positif yang diklasifikasikan dengan benar.

TN (True Negative): Data negatif yang diklasifikasikan dengan benar.

FP (False Positive): Data negatif yang diklasifikasikan sebagai positif.

FN (False Negative): Data positif yang diklasifikasikan sebagai negatif.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Berdasarkan proses pengumpulan data yang telah dilakukan, diperoleh sebanyak 1.829 data ulasan pelanggan Tomoro Coffee yang diambil dari Google Maps pada dua cabang, yaitu Tomoro Coffee Joyoboyo dan Tomoro Coffee Sekartaji.

Data yang diperoleh terdiri dari beberapa atribut, yaitu nama pengguna (nama), tanggal ulasan (at), rating, dan teks ulasan (ulasan). Atribut rating digunakan sebagai acuan awal dalam penentuan sentimen, sedangkan teks ulasan digunakan sebagai data utama dalam proses analisis sentimen.

Tabel 1. Data Ulasan Pelanggan

Nama	Tanggal	Rating	Ulasan
Inot	3 minggu lalu	1.0	PELAYANAN NYA BURUK SEKALI
Gunawan	sebulan lalu	5.0	Tempat sangat nyaman, sangat cocok untuk berkumpul dengan teman

3.2 Hasil *Preprocessing* Data

Preprocessing data dilakukan untuk meningkatkan kualitas data teks ulasan sebelum dilakukan proses analisis sentimen. Hasil *preprocessing* berupa teks ulasan yang telah bersih dan terstruktur. Contoh perbandingan antara teks ulasan sebelum dan sesudah *preprocessing* ditampilkan pada Tabel 2.

Tabel 2. Data sebelum dan setelah *preprocessing*

No	Ulasan asli	Setelah <i>Preprocessing</i>
1	pelayanan bagus, dan juga enak buat bikin tugas	layan bagus enak bikin tugas
2	tempat nyaman dan minumannya enak	nyaman minum enak

3.3 Hasil TF-IDF

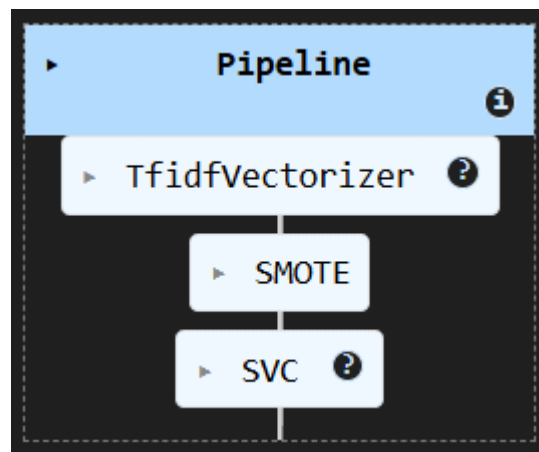
Hasil ekstraksi fitur TF-IDF menghasilkan matriks fitur berukuran 1.829×1924 , di mana 1.829 merupakan jumlah data ulasan pelanggan dan 1924 merupakan jumlah fitur kata unik yang dihasilkan dari proses pelatihan. Setiap baris pada matriks merepresentasikan satu ulasan, sedangkan setiap kolom merepresentasikan bobot kepentingan suatu kata.

3.4 Implementasi Algoritma SVM

Dataset dibagi menjadi dua bagian, yaitu data latih dan data uji dengan perbandingan 80% data latih dan 20% data uji. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur performa model.

Algoritma SVM yang digunakan pada penelitian ini adalah SVM dengan kernel linear. Pemilihan kernel linear didasarkan pada karakteristik data teks yang memiliki dimensi tinggi, sehingga kernel ini dinilai efektif dan efisien dalam proses klasifikasi.

Proses ekstraksi fitur TF-IDF dan klasifikasi SVM diintegrasikan dalam sebuah *pipeline*. Selain itu, untuk mengatasi ketidakseimbangan jumlah data pada masing-masing kelas sentimen, diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE) pada data latih.

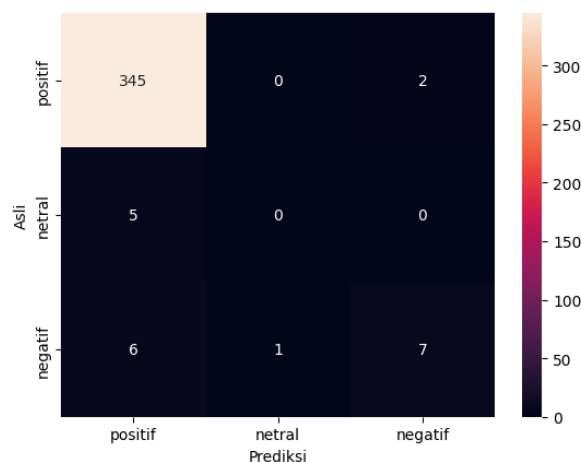


Gambar 1. Alur Implementasi Algoritma SVM

3.5 Hasil Evaluasi

3.5.1. Hasil *Confusion Matrix*

Dalam penelitian ini, evaluasi performa model klasifikasi sentimen terhadap ulasan pelanggan Tomoro Coffee, yang dibentuk menggunakan algoritma Support Vector Machine (SVM), dilakukan dengan menggunakan *Confusion Matrix*. *Confusion Matrix* digunakan untuk menggambarkan perbandingan antara label sebenarnya dengan prediksi model, sehingga dapat diketahui jumlah data yang benar maupun salah pada masing-masing kelas sentimen.



Gambar 2. Visualisasi *Confusion Matrix*

Gambar 2 menunjukkan visualisasi *confusion matrix* hasil klasifikasi sentimen ulasan pelanggan Tomoro Coffee menggunakan SVM. Berdasarkan *confusion matrix* tersebut, kelas positif memiliki jumlah prediksi benar sebanyak 345 data yang berhasil diklasifikasi dengan benar. Pada kelas netral, data sebanyak 5 ulasan tidak berhasil diklasifikasi dengan benar. Hal ini menunjukkan bahwa model masih sulit membedakan kelas netral dengan positif dan negatif. Sedangkan kelas negatif terdapat 7 data yang berhasil diklasifikasi dengan benar, sedangkan 6 data diklasifikasi sebagai kelas positif, dan 1 data sebagai kelas

netral. Kesalahan klasifikasi ini menunjukkan bahwa adanya kemiripan teks antara sentimen negatif dengan kelas yang lain.

3.5.2. Hasil Kinerja SVM

Tabel 3. Hasil Kinerja SVM

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Positif	0.97	0.99	0.98
Negatif	0.78	0.50	0.61
Netral	0.00	0.00	0.00
Akurasi	96.17%		

Hasil kinerja model SVM dalam klasifikasi sentimen ulasan pelanggan Tomoro Coffee dianalisis menggunakan metrik akurasi, *precision*, *recall*, dan *F1-Score*. Berdasarkan hasil pengujian terhadap data uji, model SVM memperoleh nilai akurasi sebesar 96.17% yang menunjukkan bahwa model mampu mengklasifikasi data ulasan dengan baik.

Kinerja terhadap kelas sentimen positif menunjukkan nilai *precision* sebesar 0.97, *recall* sebesar 0.99, dan *F1-Score* sebesar 0.98. Berdasarkan nilai tersebut menunjukkan bahwa model SVM sangat efektif dalam mengenali ulasan dengan sentimen positif.

Pada kelas sentimen negatif, model SVM menunjukkan nilai *precision* sebesar 0.78, *recall* sebesar 0.50, dan *F1-Score* sebesar 0.61. Hasil ini menunjukkan bahwa sebagian dari data negatif masih terdapat kesalahan klasifikasi, terutama pada prediksi sebagai sentimen positif.

Sementara itu, kinerja model pada kelas netral menunjukkan hasil yang kurang ideal dengan nilai *precision*, *recall*, dan *F1-Score* sebesar 0.00. Hal ini disebabkan karena jumlah data pada sentimen netral sangat sedikit, serta terdapat kemiripan karakteristik teks dengan kelas sentimen yang lain terutama pada kelas positif, sehingga model mengalami kesulitan dalam membedakan secara akurat.

Secara keseluruhan, hasil evaluasi ini menunjukkan algoritma SVM memiliki performa yang sangat baik dalam mengklasifikasi sentimen ulasan pelanggan. Jika membandingkan dengan penelitian sebelumnya [11] yang menganalisis ulasan Kopi Kenangan menggunakan algoritma Naïve Bayes dan memperoleh nilai akurasi sebesar 72%, hasil dari penelitian ini menunjukkan bahwa penggunaan algoritma SVM mampu memberikan performa klasifikasi yang lebih baik.

4 SIMPULAN

Berdasarkan hasil pengujian dan analisis yang dilakukan terhadap sentimen ulasan pelanggan Tomoro Coffee, dapat disimpulkan sebagai berikut:

1. Penerapan algoritma SVM dengan fitur TF-IDF efektif digunakan untuk klasifikasi sentimen ulasan pelanggan, yang memiliki akurasi sebesar 96.17%. Hal ini menunjukkan metode mampu menggambarkan kepuasan pelanggan mayoritas dengan baik.
2. Kelebihan model terletak pada presisi dengan nilai 0.97 dan *recall* 0.99. Namun model masih memiliki kelemahan pada kelas netral yang memiliki nilai *F1-Score* 0.00, dan kelas negatif hanya 0.50.
3. Ketidakseimbangan jumlah data antar kelas sentimen, terutama pada kelas netral dan negatif. Sehingga menyebabkan kesalahan klasifikasi pada kelas yang memiliki data sedikit.

5 SARAN

Berdasarkan hasil penelitian dan kendala yang dihadapi, berikut saran untuk pengembangan penelitian selanjutnya:

1. Penelitian selanjutnya disarankan menggunakan *Adaptive Synthetic Sampling* (ADASYN), guna mengatasi ketidakseimbangan jumlah data antar kelas sentimen.
2. Pengembangan penelitian dapat membandingkan algoritma Support Vector Machine (SVM) dengan algoritma yang lain, atau menggunakan metode deep learning.
3. Penelitian selanjutnya dapat memperluas sumber data dari platform digital yang lain, agar data lebih beragam.

DAFTAR PUSTAKA

- [1] R. F. Siahaan and R. Muliono, “Coffee Quality Classification Based on Customer Reviews Using C4.5 Algorithm,” *JITE (Journal of Informatics and Telecommunication Engineering) Algorithm*, vol. 8, no. 3, pp. 101–109, 2025, doi: 10.31289/jite.v8i3Spc.14427.
- [2] M. P. Muhammad Rizwan, *Budidaya Kopi*. Cv. Azka Pustaka, 2022.
- [3] N. Azhar, P. P. Adikara, S. Adinugroho, and P. Korespondensi, “Analisis Sentimen Ulasan Kedai Kopi Menggunakan Metode Naïve Bayes Dengan Seleksi Fitur Algoritme Genetika Sentiment Analysis For Coffee Shop Reviews Using Naïve Bayes Method With Genetic Algorithm Feature Selection,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 8, no. 3, pp. 609–618, 2021, doi: 10.25126/jtiik.202184436.
- [4] M. A. Palimbani, R. P. Hasuti, and R. A. Rajagede, “Analisis Sentimen Berbasis Aspek pada Ulasan Pengguna Aplikasi Starbucks Menggunakan Algoritma Support Vector Machine,” *Journal of Internet and Software Engineering*, vol. 5, no. 1, 2024.
- [5] A. M. Putri, W. K. Nofa, and D. A. P. Hapsari, “Penerapan Metode Bert Untuk Analisis Sentimen Ulasan Pengguna Aplikasi Segari Di Google Play Store,” *Jurnal Ilmiah Teknik*, vol. 4, no. 1, pp. 89–104, 2025.
- [6] L. A. Fitriana, M. F. Julianto, R. Dahlia, M. R. Firdaus, and A. Fazriansyah, “Analisis Ulasan Konsumen sebagai Data Non-Keuangan dalam Sistem Informasi Akuntansi,” *Profitabilitas*, vol. 5, no. 1, pp. 64–74, 2025.
- [7] A. F. Setyawan, Ariyanto, Fari Katul Fikriah, and Rozaq Isnaini Nugraha, “Analisis Sentimen Ulasan iPhone di Amazon Menggunakan Model Deep Learning BERT Berbasis Transformer,” *Elkom: Jurnal Elektronika dan Komputer*, vol. 17, no. 2, pp. 447–452, Dec. 2024, doi: 10.51903/elkom.v17i2.2150.
- [8] R. Rahmadani, A. Rahim, and R. Rudiman, “Analisis Sentimen Ulasan ‘Ojol the Game’ Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024.
- [9] T. I. Alfawas, A. Rahim, and R. Rudiman, “Penerapan Fitur Ekstraksi TF-IDF untuk Analisis Sentimen Ulasan Game Bus Simulator Indonesia dengan Algoritma Naive Bayes,” *Innovative: Journal Of Social Science Research*, vol. 4, no. 5, pp. 3177–3193, 2024.
- [10] B. Salungweni, W. Weku, and E. Ketaren, “Analisis pengaruh film ‘Ice Cold’ kasus kopi sianida terhadap sentimen pengguna YouTube dengan SVM dan Random Forest,” *Jurnal TIMES*, vol. 13, no. 2, pp. 31–37, 2024.
- [11] H. Sakdiyah, “Analisis sentimen customer review brand Kopi Kenangan menggunakan metode Naive Bayes,” *Fakultas Sains dan Teknologi*, 2023.