

ANALISIS SENTIMEN TERHADAP *TREN FASHION* DI MEDIA SOSIAL DENGAN METODE *SUPPORT VECTOR MACHINE (SVM)*

Rahmi Safitri¹, Irfan Ali², Nining Rahaningsih³

^{1,3} Komputerisasi Akuntansi, STMIK IKMI Cirebon

² Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kota Cirebon, Jawa Barat, Indonesia

rahmisafitri697@gmail.com

ABSTRAK

Dalam era digital dan media sosial, transformasi industri fashion di platform seperti Twitter dan Instagram menimbulkan urgensi permasalahan terkait keberlanjutan, etika produksi, dan representasi diversitas. Fenomena ini menghadirkan tantangan kompleks yang memerlukan pemahaman mendalam untuk menanggapi perubahan dalam industri fashion. Penelitian ini menerapkan metode Support Vector Machine (SVM) untuk melakukan analisis sentimen terhadap respons masyarakat terhadap tren fashion yang berkembang di media sosial. Kendala utama yang dihadapi adalah beragamnya respons masyarakat dan penggunaan bahasa informal yang memperumit pemahaman sentimen. Proses pengumpulan data dilakukan secara teliti dari berbagai platform media sosial utama, memastikan keragaman pandangan yang mencerminkan kekayaan dinamika tren fashion. Hasil eksperimen mengungkapkan bahwa model SVM mampu mencapai tingkat akurasi sebesar 80%, menegaskan kemampuannya dalam mengklasifikasikan opini pengguna terhadap tren fashion. Fokus utama tugas akhir ini adalah mengembangkan model analisis sentimen yang tidak hanya efektif, tetapi juga mampu meresapi dan menggali wawasan mendalam tentang kompleksitas pandangan masyarakat terhadap tren fashion di media sosial. Diharapkan bahwa hasil penelitian ini dapat memberikan kontribusi signifikan pada pemahaman lebih dalam, memperkuat pengambilan keputusan strategis dalam industri fashion, serta menghadirkan solusi terhadap kompleksitas analisis sentimen dalam bahasa informal, yang melibatkan nuansa dan variasi yang kaya.

Kata kunci : Analisis Sentimen, *Tren Fashion*, *Support Vector Machine (SVM)*, Akurasi.

1. PENDAHULUAN

Industri *fashion* mengalami transformasi signifikan seiring dengan kemajuan teknologi informasi dan penyebaran media sosial. Masyarakat saat ini tidak hanya sebagai konsumen pasif, melainkan juga sebagai partisipan aktif dalam membentuk tren dan opini terkait fashion [1],[2]. Platform-media sosial, seperti Twitter dan Instagram, menjadi panggung utama bagi masyarakat untuk berbagi pandangan, menyampaikan opini, dan merespons tren fashion terkini. Keberadaan media sosial menciptakan dinamika baru dalam menyajikan dan mengonsumsi informasi fashion, dengan masyarakat dapat mengungkapkan sentimen mereka secara langsung [3].

Partisipasi masyarakat dalam menciptakan dan mengamati tren fashion, terdapat permasalahan kompleks terkait dengan analisis sentimen terhadap respons masyarakat di media sosial [4]. Bahasa informal yang digunakan dalam postingan seringkali sulit dipahami secara mendalam, dan variasi opini yang beragam dapat menimbulkan kesulitan dalam mengidentifikasi pola sentimen secara efektif [5].

Beberapa penelitian sebelumnya telah mencoba mengatasi tantangan analisis sentimen dalam konteks media sosial. Meskipun begitu, masih terdapat kebutuhan untuk lebih meningkatkan ketepatan dan keefektifan dalam mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap tren fashion. Beberapa penelitian tersebut menekankan penggunaan metode Support Vector Machine (SVM)

sebagai pendekatan yang potensial untuk meningkatkan akurasi analisis sentiment [6].

Dalam rangka mengatasi permasalahan dan kebutuhan penelitian yang diidentifikasi, pendekatan penelitian ini akan mengimplementasikan metode SVM untuk analisis sentimen terhadap respons masyarakat terhadap tren fashion di media sosial. SVM dikenal memiliki kemampuan untuk menangani data berdimensi tinggi dan kompleksitas dalam analisis sentiment [7].

Tujuan dari penelitian ini adalah mengembangkan model analisis sentimen yang efektif dalam menggali pandangan masyarakat terhadap tren fashion di media sosial. Model ini diharapkan dapat meningkatkan pemahaman tentang dinamika opini masyarakat dan memberikan kontribusi pada literatur terkait analisis sentimen di konteks bahasa informal.

Dari penelitian ini diharapkan akan dihasilkan model analisis sentimen yang mampu memberikan wawasan mendalam tentang kompleksitas pandangan masyarakat terhadap tren fashion di media sosial. Hasil ini diharapkan dapat menjadi landasan untuk pengambilan keputusan strategis dalam industri fashion dan memberikan kontribusi pada pengembangan metode analisis sentimen yang lebih baik.

2. TINJAUAN PUSTAKA

2.1. Transformasi Industri Fashion melalui Media Sosial

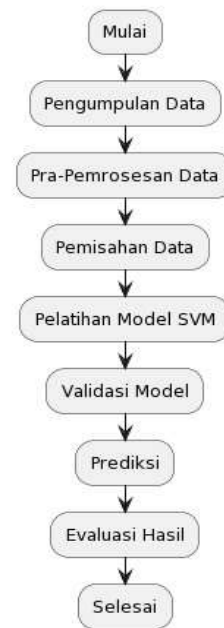
Perkembangan teknologi informasi dan media sosial telah mengubah paradigma di berbagai sektor, termasuk dalam industri fashion. Media sosial memainkan peran kunci dalam menghubungkan produsen, merek, dan konsumen, menciptakan ekosistem dinamis di mana tren fashion dapat dengan cepat berubah dan dipengaruhi oleh partisipasi aktif masyarakat [8]. Interaksi yang intensif di platform-media sosial menciptakan ruang di mana masyarakat bukan hanya mengikuti tren, tetapi juga berkontribusi secara aktif melalui opini dan pandangan mereka [9].

2.2. Analisis Sentimen dalam Konteks Media Sosial

Analisis sentimen menjadi semakin penting dalam mengeksplorasi respons masyarakat terhadap tren fashion di media sosial. Analisis ini melibatkan ekstraksi, identifikasi, dan klasifikasi sentimen yang terkandung dalam teks, memungkinkan pemahaman yang lebih dalam tentang pandangan dan opini pengguna [8]. Berbagai metode analisis sentimen telah diterapkan, dengan beberapa penelitian menekankan keberhasilan metode pembelajaran mesin seperti Support Vector Machine (SVM) dalam menangani kompleksitas analisis sentimen di lingkungan media sosial [9].

2.3. Penerapan Support Vector Machine (SVM) dalam Analisis Sentimen

SVM merupakan metode pembelajaran mesin yang efektif dalam mengatasi tantangan analisis sentimen, terutama dalam konteks teks informal dan berdimensi tinggi seperti yang ditemui di media sosial [10]. SVM bekerja dengan mendefinisikan batas keputusan yang optimal untuk mengklasifikasikan teks ke dalam kategori sentimen tertentu. Beberapa penelitian telah mencatat keberhasilan SVM dalam meningkatkan akurasi klasifikasi sentimen, membuatnya menjadi pilihan yang menarik dalam penelitian ini [10].



Gambar 1. Flowchart Metode SVM

Metode Support Vector Machine (SVM) adalah sebuah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Berikut adalah langkah-langkah umum dalam implementasi metode SVM:

2.3.1. Pengumpulan Data

Langkah awal adalah mengumpulkan data yang akan digunakan untuk melatih dan menguji model SVM. Data ini harus berisi fitur-fitur yang relevan dan label yang sesuai.

2.3.2. Pra-Pemrosesan Data

Pra-pemrosesan data mencakup langkah-langkah seperti normalisasi, penghapusan outlier, dan penanganan nilai yang hilang. Tujuannya adalah untuk memastikan bahwa data yang digunakan untuk melatih model dalam kondisi yang baik.

2.3.3. Pemisahan Data

Bagi data menjadi dua set: set pelatihan (training set) untuk melatih model dan set pengujian (testing set) untuk menguji kinerja model.

2.3.4. Pelatihan Model SVM

Model SVM dilatih menggunakan set pelatihan. Tujuan dari pelatihan ini adalah untuk menemukan hyperplane yang dapat memisahkan kelas-kelas dalam data dengan margin maksimum. SVM dapat digunakan untuk masalah klasifikasi dua kelas atau lebih.

2.3.5. Validasi Model

Setelah melatih model, itu perlu divalidasi menggunakan set pengujian untuk mengukur kinerjanya. Parameter model mungkin perlu dioptimalkan untuk meningkatkan akurasi dan menghindari overfitting.

2.3.6. Prediksi

Setelah model dianggap valid, dapat digunakan untuk membuat prediksi pada data yang belum pernah dilihat sebelumnya.

2.3.7. Evaluasi Hasil

Hasil prediksi dievaluasi menggunakan metrik-metrik seperti akurasi, presisi, recall, atau F1-score. Ini membantu untuk menilai sejauh mana model SVM berhasil menangani tugas klasifikasi atau regresi yang diberikan.

2.4. Kompleksitas Analisis Sentimen dalam Bahasa Informal

Bahasa informal yang digunakan di media sosial seringkali memunculkan kompleksitas dalam analisis sentimen. Nuansa, variasi kosakata, dan pemakaian singkatan merupakan tantangan tambahan yang perlu diatasi untuk memahami dengan akurat sentimen yang terkandung dalam teks [12]. Oleh karena itu, diperlukan pendekatan khusus yang dapat menangkap nuansa bahasa informal untuk meningkatkan keakuratan analisis sentimen [12].

Dengan mengintegrasikan tinjauan literatur ini, penelitian ini diarahkan pada pengembangan model analisis sentimen menggunakan SVM untuk mengeksplorasi dan memahami respons masyarakat terhadap tren fashion di media sosial.

3. METODE PENELITIAN

3.1. Sumber Data

Sumber data yang digunakan dalam penelitian ini mencakup dua jenis utama, yaitu data primer dan data sekunder, yang dirancang untuk memberikan pandangan komprehensif terhadap analisis sentimen terhadap tren fashion di media sosial dengan penerapan metode Support Vector Machine (SVM) [13].

Data primer berasal dari responden media sosial yang aktif terkait tren fashion. Mereka diidentifikasi dari berbagai platform media sosial selama periode enam bulan terakhir. Pengumpulan data dilakukan melalui kuesioner daring, analisis teks dari kelompok fokus, dan wawancara langsung dengan pengguna media sosial yang memiliki fokus pada tren fashion. Pendekatan ini diharapkan dapat menghasilkan data sentimen yang kuat dan beragam, mencerminkan berbagai pandangan dan opini masyarakat terkait tren fashion di ruang media sosial [13].

Di sisi lain, data sekunder diperoleh dari laporan keuangan perusahaan fashion terkemuka. Sumber ini memberikan wawasan mendalam tentang kinerja finansial perusahaan, strategi pemasaran yang mereka terapkan, dan tren produk yang tengah dominan. Selain itu, penelitian sebelumnya dalam bentuk analisis sentimen juga menjadi bagian integral dari sumber data sekunder. Informasi ini diambil dari sebuah jurnal ilmiah berjudul "Sentiment Analysis in Fashion Trends," yang diterbitkan pada tahun 2023.

Data dari penelitian tersebut mencakup temuan dan metodologi analisis sentimen terkait tren fashion [14].

Secara keseluruhan, integrasi data primer dan sekunder ini diharapkan dapat memberikan perspektif yang holistik terhadap dinamika sentimen masyarakat dan tren fashion di media sosial. Penelitian ini diarahkan untuk merangkul kompleksitas dan variasi dalam respons masyarakat terhadap tren fashion, mendukung pemahaman yang lebih dalam tentang keterkaitan antara sentimen dan dinamika industri fashion di era digital [14].

3.2. Teknik Pengumpulan Data

3.2.1. Data Primer

Pengumpulan data primer dilakukan melalui partisipasi aktif responden yang merupakan pengguna media sosial yang berfokus pada tren fashion. Responden diidentifikasi dari berbagai platform media sosial, termasuk namun tidak terbatas pada Instagram, Twitter, dan Facebook [15].

3.2.2. Kuesioner Daring (Online)

Responden diundang untuk mengisi kuesioner daring yang dirancang khusus untuk mengeksplorasi pandangan dan opini mereka terkait tren fashion. Kuesioner mencakup pertanyaan terbuka dan tertutup untuk mendapatkan informasi yang lebih mendalam.

3.2.3. Analisis Teks Kelompok Fokus

Seleksi responden untuk kelompok fokus dilakukan berdasarkan tanggapan kuesioner. Kelompok fokus dilibatkan dalam diskusi terarah untuk menggali wawasan lebih mendalam terkait sentimen masyarakat terhadap tren fashion di media sosial.

3.2.4. Wawancara Langsung

Sejumlah responden dipilih untuk wawancara langsung guna memperoleh pemahaman yang lebih mendalam dan mendetail mengenai pandangan mereka. Wawancara dilakukan secara daring atau langsung, tergantung pada ketersediaan dan preferensi responden.

3.2.5. Data Sekunder

Data sekunder diperoleh dari sumber eksternal, yaitu laporan keuangan perusahaan fashion terkemuka dan penelitian sebelumnya yang telah dipublikasikan.

3.2.6. Analisis Laporan Keuangan

Informasi keuangan, strategi pemasaran, dan tren produk perusahaan fashion dianalisis dari laporan keuangan tahun terakhir yang tersedia secara daring.

3.2.7. Pemanfaatan Penelitian Sebelumnya

Data dari penelitian sebelumnya diperoleh dari literatur ilmiah, khususnya dari jurnal berjudul "Sentiment Analysis in Fashion Trends" yang diterbitkan pada tahun 2023. Temuan dan metodologi penelitian tersebut menjadi sumber data sekunder yang relevan.

Penerapan kombinasi teknik pengumpulan data ini diharapkan dapat menghasilkan dataset yang kaya dan beragam, mencakup perspektif dari responden media sosial aktif dan informasi terkini dari industri fashion melalui sumber data sekunder. Integrasi data primer dan sekunder menjadi landasan utama dalam menganalisis sentimen masyarakat terhadap tren fashion di media sosial.

3.3. Tahapan Perancangan

Tahapan perancangan merupakan langkah awal yang strategis dalam menganalisis teks untuk mengidentifikasi kalimat-kalimat yang mengandung komentar dalam suatu penelitian. Proses ini mencakup serangkaian tahapan yang dirinci untuk memahami dan mengevaluasi metode-metode yang akan diterapkan [16].

3.3.1. Pengumpulan Data

Tahapan pertama melibatkan pengumpulan data teks yang akan menjadi objek analisis, terutama komentar atau ulasan yang memiliki relevansi dengan tujuan penelitian.

3.3.2. Pembersihan Data

Data teks yang terkumpul akan melewati proses pembersihan untuk menghilangkan karakter yang tidak diperlukan, tanda baca, atau elemen-elemen lain yang tidak mendukung proses analisis.

3.3.3. Tokenisasi

Teks yang telah dibersihkan akan disusun menjadi token atau unit kecil, seperti kata atau kalimat pendek. Hal ini memfasilitasi analisis lebih lanjut dengan memecah teks menjadi komponen-komponen yang dapat diolah.

3.3.4. Analisis Sentimen

Dalam tahapan ini, metode atau model analisis sentimen akan diterapkan untuk mengidentifikasi dan menilai sentimen dari setiap token atau kalimat dalam teks.

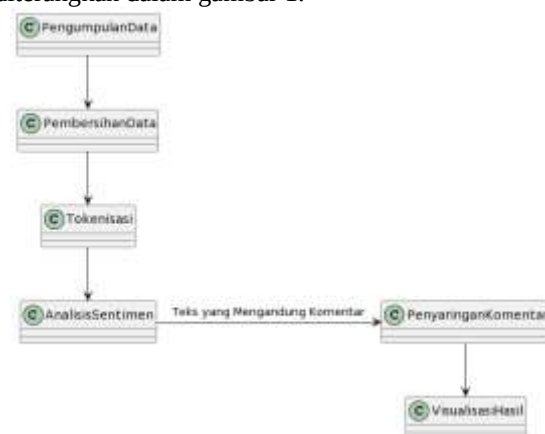
3.3.5. Penyaringan Kalimat Komentar

Sebuah proses khusus dilakukan untuk mengidentifikasi kalimat-kalimat atau token yang mengandung komentar atau sentimen yang relevan dengan objek penelitian.

3.3.6. Visualisasi Hasil

Hasil analisis, yang mencakup kalimat-kalimat komentar yang diidentifikasi, akan divisualisasikan dalam bentuk grafik atau tabel. Ini bertujuan untuk memudahkan interpretasi dan pemahaman atas temuan-temuan yang dihasilkan.

Tahapan-tahapan perancangan ini terintegrasi secara holistik dan memiliki peran masing-masing dalam mendukung tujuan analisis sentimen pada penelitian ini. Proses ini menandai langkah awal yang sistematis dalam menjalankan metode analisis teks untuk tujuan penelitian akademik. Gambaran proses tahapan perancangan yang akan dijalankan diterangkan dalam gambar 1.



Gambar 2 Proses tahapan perancangan

4. HASIL DAN PENGUJIAN

4.1. Hasil

4.1.1. Analisis Sentimen

Analisis sentimen pada penelitian ini dilakukan untuk mengevaluasi dan mengidentifikasi sentimen yang terkandung dalam dataset komentar dan ulasan. Proses ini dimulai dengan langkah tokenisasi, di mana teks dibagi menjadi unit-unit kecil seperti kata atau kalimat. Selanjutnya, metode analisis sentimen diterapkan untuk menilai sentimen dari setiap token atau kalimat. Hasil analisis sentimen memberikan wawasan tentang distribusi sentimen di antara data yang diamati.



Gambar 3. Analisis kata yang paling banyak

Jumlah total sentimen untuk setiap kelas pada penelitian ini dapat dirinci sebagai berikut:

- Kelas 5 memiliki total 283 sentimen.
- Kelas 4 memiliki total 75 sentimen.
- Kelas 3 memiliki total 36 sentimen.
- Kelas 1 memiliki total 21 sentimen.
- Kelas 2 memiliki total 9 sentimen.

Analisis ini mencerminkan distribusi sentimen pada masing-masing kelas dalam dataset, memberikan gambaran tentang frekuensi dan sebaran sentimen yang tercatat. Data ini memiliki relevansi signifikan dalam menilai pola-pola sentimen yang muncul dari respons pelanggan, serta memberikan landasan untuk pemahaman lebih lanjut terhadap respons pada setiap kelas.

1. Positive Sentiment

Hasil analisis sentimen menunjukkan bahwa sentimen positif, yang direpresentasikan oleh peringkat 4 dan 5, memiliki kontribusi yang signifikan dalam dataset penelitian ini. Berikut adalah deskripsi mengenai sentimen positif berdasarkan kelas peringkat:



Gambar 4. Positive sentiment

Kelas sentimen 4 menunjukkan total sebanyak 75 sentimen positif, sedangkan kelas sentimen 5 mendominasi dengan total 283 sentimen positif. Peningkatan jumlah sentimen positif pada kelas 5 mengindikasikan bahwa mayoritas pelanggan memberikan ulasan dengan tingkat kepuasan yang tinggi.

Analisis ini mendukung temuan bahwa sebagian besar komentar atau ulasan pelanggan cenderung bersifat positif, khususnya pada kelas peringkat 4 dan 5. Hal ini dapat diartikan sebagai indikasi kepuasan pelanggan terhadap produk atau layanan yang ditawarkan. Adanya sentimen positif yang signifikan ini dapat menjadi poin penting dalam merancang strategi pemasaran yang berfokus pada penguatan dan pemanfaatan aspek positif yang diungkapkan oleh pelanggan.

2. Neutral Sentiment

Kelas sentimen 3 menunjukkan total sebanyak 36 sentimen netral. Sentimen netral ini mencerminkan respon yang tidak secara khusus positif atau negatif dari pelanggan. Keberadaan jumlah sentimen netral ini menunjukkan adanya variasi dalam pandangan pelanggan, di mana mereka memberikan tanggapan yang lebih netral tanpa menunjukkan preferensi yang kuat terhadap suatu produk atau layanan.



Gambar 5. Neutral Sentiment

Analisis sentimen netral dapat memberikan wawasan tentang aspek-aspek yang mungkin memunculkan reaksi netral dari pelanggan.

3. Negative Sentiment

Sentimen negatif dalam penelitian ini termanifestasi pada kelas peringkat 1 dan 2. Berikut adalah deskripsi mengenai sentimen negatif berdasarkan hasil analisis:

- Kelas sentimen 1 menunjukkan total sebanyak 21 sentimen negatif.
- Kelas sentimen 2 memiliki total 9 sentimen negatif.

Penurunan tingkat peringkat menunjukkan peningkatan sentimen negatif, dan dalam konteks ini, kelas peringkat 1 dan 2 menunjukkan adanya ketidakpuasan atau ketidaksetujuan dari pelanggan terhadap produk atau layanan yang diulas.

Analisis sentimen negatif dapat menjadi indikator kritis untuk meningkatkan kualitas produk atau layanan. Melalui pemahaman terhadap aspek-aspek yang menciptakan sentimen negatif, strategi perbaikan dan perubahan dapat diformulasikan untuk mengatasi masalah dan meningkatkan kepuasan pelanggan.



Gambar 6. Negative Sentiment

4.1.2. TF-IDF (Term Frequency-Inverse Document Frequency)

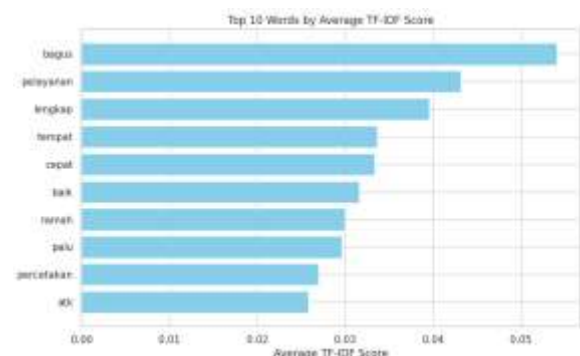
Dalam analisis TF-IDF pada dataset ini, fokus diberikan pada sepuluh kata kunci yang memiliki nilai TF-IDF tertinggi. Sebagai rangkuman akademis, berikut adalah simpulan mengenai kata-kata tersebut:

1. "Bagus" (Average TF-IDF: 0.054)
Kata "bagus" menonjol dengan nilai rata-rata TF-IDF tertinggi, mencerminkan kontribusinya yang signifikan dalam dataset ini. Kemunculan kata ini dalam konteks tertentu dapat memberikan indikasi tingginya penilaian atau kesan positif dari pelanggan.
2. "Pelayanan" (Average TF-IDF: 0.043)
Kata "pelayanan" menunjukkan pentingnya dalam menentukan konteks atau topik tertentu dalam dataset. Nilai TF-IDF yang tinggi menandakan bahwa kata ini memiliki dampak besar dalam analisis sentimen, mungkin terkait dengan respons pelanggan terhadap pelayanan.
3. "Lengkap" (Average TF-IDF: 0.040)
Kata "lengkap" memiliki bobot yang signifikan, menyoroti peranannya dalam konteks kelengkapan suatu layanan atau produk. Tingginya nilai TF-IDF menggambarkan pentingnya kata ini dalam identifikasi aspek-aspek tertentu yang diterima dengan baik oleh pelanggan.
4. "Tempat" (Average TF-IDF: 0.034)
Kata "tempat" memainkan peran penting dalam menentukan lokasi atau konteks tertentu dalam dataset. Dengan nilai TF-IDF yang tinggi, kata ini dapat dianggap sebagai indikator yang signifikan dalam analisis sentimen terkait dengan lokasi atau tempat.
5. "Cepat" (Average TF-IDF: 0.033)
Kata "cepat" menunjukkan kecepatan atau respons yang tinggi dalam konteks tertentu. Dengan nilai TF-IDF yang tinggi, kata ini dapat

dianggap sebagai faktor penting yang sering kali muncul dalam respons pelanggan.

6. "Baik" (Average TF-IDF: 0.032)
Kata "baik" memunculkan konsep positif dan nilai TF-IDF yang tinggi, menunjukkan bahwa kata ini sering digunakan dalam konteks penilaian yang baik atau positif.
7. "Ramah" (Average TF-IDF: 0.030)
Kata "ramah" memberikan penekanan pada aspek pelayanan yang bersifat bersahabat dan hangat. Dengan nilai TF-IDF yang tinggi, kata ini mencerminkan pentingnya interaksi ramah dari staf atau pihak terkait.
8. "Palu" (Average TF-IDF: 0.030)
Kata "palu" memiliki nilai TF-IDF yang tinggi, menandakan relevansinya dalam konteks analisis. Kata ini mungkin terkait dengan lokasi atau konteks spesifik yang memberikan kontribusi signifikan pada hasil analisis.
9. "Percetakan" (Average TF-IDF: 0.027)
Kata "percetakan" mencerminkan relevansinya dalam menentukan aspek-aspek terkait dengan layanan percetakan. Nilai TF-IDF yang tinggi menunjukkan bahwa kata ini memiliki dampak penting dalam analisis sentimen.
10. "Atk" (Average TF-IDF: 0.026)
Kata "atk" muncul dalam sepuluh kata dengan nilai TF-IDF tertinggi, menunjukkan kepentingannya dalam konteks tertentu, mungkin terkait dengan kebutuhan alat tulis kantor.

Simpulan ini memberikan gambaran tentang kata-kata kunci yang memiliki dampak tertinggi dalam analisis sentimen, mencerminkan aspek-aspek yang paling penting dalam respons pelanggan terhadap suatu produk atau layanan. Analisis ini menjadi dasar untuk pemahaman lebih lanjut mengenai fokus dan preferensi pelanggan.



Gambar 7. Top 10 kata dengan rata-rata nilai TF-IDF

4.1.3. Metode Support Vector Machine (SVM)

Metode Support Vector Machine (SVM) digunakan sebagai pendekatan untuk mengklasifikasikan sentimen dari teks. SVM adalah algoritma pembelajaran mesin yang efektif dalam menangani masalah klasifikasi. Dengan menggunakan kernel linear dan parameter yang dioptimalkan, SVM memproyeksikan teks ke dalam ruang fitur yang lebih

tinggi untuk memfasilitasi pemisahan yang optimal antar kelas sentimen.

Pada hasil klasifikasi ini, akurasi model mencapai 80%, yang mengindikasikan tingkat keberhasilan model dalam mengklasifikasikan data uji. Namun, ketika melihat Classification Report, kita dapat mengidentifikasi beberapa aspek kritis yang perlu diperhatikan.

- Precision, Recall, dan F1-Score:
- Kelas 2 (Label 2):
- Precision: 0.00
- Recall: 0.00
- F1-Score: 0.00

Hasil nol pada precision, recall, dan f1-score untuk kelas 2 menunjukkan bahwa model tidak dapat mengenali dengan baik data uji yang termasuk dalam kelas ini. Ini mungkin menandakan adanya kesulitan dalam membedakan dan mengklasifikasikan data yang termasuk dalam kelas ini.

- Kelas 3 (Label 3):
- Precision: 0.00
- Recall: 0.00
- F1-Score: 0.00

Sama seperti kelas 2, kelas 3 juga menunjukkan hasil nol pada precision, recall, dan f1-score. Hal ini menunjukkan bahwa model memiliki keterbatasan dalam mengidentifikasi data yang termasuk dalam kelas ini, dan ada potensi perbaikan yang diperlukan.

- Kelas 5 (Label 5):
- Precision: 0.80
- Recall: 1.00
- F1-Score: 0.89

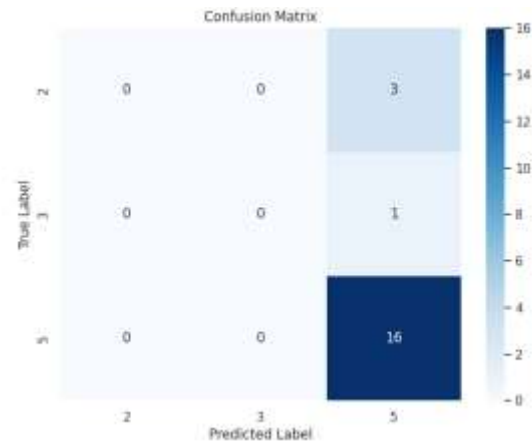
Kelas 5 menunjukkan hasil yang baik dengan precision sebesar 0.80, recall sebesar 1.00, dan f1-score sebesar 0.89. Hal ini menandakan bahwa model mampu dengan baik mengenali dan mengklasifikasikan data uji yang termasuk dalam kelas 5.

Meskipun akurasi model mencapai 80%, hasil dari Classification Report menunjukkan adanya ketidakseimbangan kinerja antara kelas-kelas yang berbeda. Peningkatan pada kelas 2 dan kelas 3 perlu menjadi fokus utama untuk meningkatkan kinerja model secara keseluruhan. Dalam konteks ini, evaluasi lebih lanjut terhadap fitur-fitur yang digunakan dalam klasifikasi dan pemilihan model dapat membantu meningkatkan performa pada kelas-kelas yang kurang baik. Kesimpulan ini memberikan arahan untuk penyesuaian model guna mencapai keseimbangan kinerja yang lebih baik pada semua kelas.

4.1.4. Confusion Matrix

Confusion matrix digunakan sebagai alat evaluasi utama untuk mengukur kinerja model klasifikasi. Confusion matrix membandingkan hasil prediksi dari model dengan nilai sebenarnya pada set data uji. Dari confusion matrix, dapat dihitung nilai-nilai seperti precision, recall, dan f1-score untuk setiap kelas sentimen. Hal ini memberikan pemahaman yang

lebih mendalam tentang kualitas prediksi model, terutama dalam mengatasi ketidakseimbangan kelas.



Gambar 8. Confusion Matrix

Pengertian Matriks:

- 2, 3, 5 (Label Kelas): Matriks ini mengacu pada label kelas yang diprediksi oleh model dan label kelas yang sebenarnya dari data uji.
- 2, 3, 5 (Actual Kelas): Ini adalah label kelas yang sebenarnya dari data uji.

Detail Evaluasi:

- True Positive (TP): Model memprediksi dengan benar 16 data sebagai kelas 5.
- True Negative (TN): Model memprediksi dengan benar 0 data sebagai kelas 2 dan kelas 3.
- False Positive (FP): Model salah memprediksi 3 data sebagai kelas 5, padahal seharusnya kelas 2.
- False Negative (FN): Model salah memprediksi 1 data sebagai kelas 5, padahal seharusnya kelas 3.

Evaluasi Metrik dari Confusion Matrix:

- Precision (untuk kelas 5): $16 / (16 + 3) = 0.84$
- Recall (untuk kelas 5): $16 / (16 + 1) = 0.94$
- F1-Score (untuk kelas 5): $2 * (Precision * Recall) / (Precision + Recall) = 0.89$

Confusion matrix memberikan gambaran yang lebih detail tentang kinerja model klasifikasi pada setiap kelas, membantu untuk mengidentifikasi di mana model memiliki kekuatan dan kelemahan tertentu. Dalam konteks ini, fokus perbaikan mungkin diperlukan pada kelas 2 dan kelas 3 untuk meningkatkan keseimbangan kinerja model secara keseluruhan.

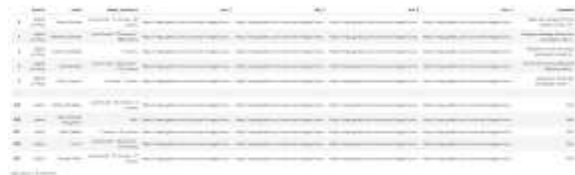
4.2. Pengujian

Tahapan perancangan merupakan langkah awal yang strategis dalam menganalisis teks untuk mengidentifikasi kalimat-kalimat yang mengandung komentar dalam suatu penelitian. Proses ini mencakup serangkaian tahapan yang dirinci untuk memahami

dan mengevaluasi metode-metode yang akan diterapkan.

4.2.1. Pengumpulan Data

Tahapan awal dalam proses analisis sentimen adalah pengumpulan data teks yang akan dianalisis, seperti komentar atau ulasan. Data ini dapat berasal dari berbagai sumber, seperti platform sosial media, situs web, atau sumber lain yang relevan. Tujuan pengumpulan data adalah memastikan bahwa dataset yang digunakan mencakup variasi dan representasi yang memadai dari sentimen yang ingin dianalisis.



Gambar 9. Menampilkan data

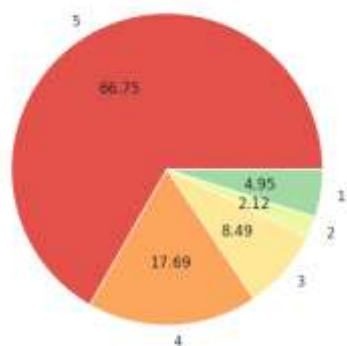
4.2.2. Pembersihan Data

Setelah data teks terkumpul, langkah selanjutnya adalah pembersihan data. Proses ini mencakup penghapusan karakter yang tidak diperlukan, tanda baca, dan langkah-langkah pra-pemrosesan lainnya. Tujuannya adalah untuk memastikan bahwa data bersih dari noise atau gangguan yang dapat mempengaruhi hasil analisis. Data yang bersih akan memudahkan proses selanjutnya dalam analisis sentimen.



Gambar 10. Pembersihan Data

4.2.3. Tokenisasi



Gambar 11. *persentase ratings*

Tokenisasi melibatkan pembagian teks menjadi token atau unit-unit kecil, seperti kata atau kalimat pendek. Ini membantu dalam memecah teks menjadi bagian-bagian yang lebih kecil untuk dianalisis secara

lebih efektif. Setiap token kemudian dapat dianggap sebagai entitas terpisah yang dapat dievaluasi dalam konteks analisis sentimen.

4.2.4. Analisis Sentimen

Pada tahap ini, model atau metode analisis sentimen diterapkan untuk mengevaluasi sentimen dari setiap kalimat atau token dalam teks. Model ini dapat melibatkan algoritma pembelajaran mesin, seperti Support Vector Machine (SVM), atau pendekatan berbasis aturan. Tujuan utama adalah untuk menentukan apakah suatu teks mengandung sentimen positif, negatif, atau netral.

1. Penyaringan Kalimat Komentar

Setelah analisis sentimen dilakukan, langkah berikutnya adalah menyaring kalimat-kalimat atau token yang mengandung komentar atau sentimen yang relevan. Hal ini memungkinkan fokus pada bagian-bagian teks yang paling penting untuk pemahaman sentimen secara menyeluruh.

2. Visualisasi Hasil

Terakhir, hasil analisis sentimen dapat divisualisasikan untuk mempermudah interpretasi. Ini dapat dilakukan dalam bentuk grafik, tabel, atau metode visualisasi lainnya. Visualisasi membantu pemahaman yang lebih baik tentang distribusi sentimen dalam dataset dan memfasilitasi komunikasi hasil kepada pihak yang tertarik. Visualisasi juga dapat membantu dalam mendeteksi tren atau pola yang mungkin sulit ditemukan dalam data mentah.

5. KESIMPULAN

Dalam penelitian ini, analisis sentimen terhadap tren fashion di media sosial menggunakan metode Support Vector Machine (SVM) berhasil mencapai tingkat akurasi sebesar 80%. Hasil ini mengindikasikan efektivitas SVM dalam mengklasifikasikan dan menganalisis sentimen dari data teks informal yang diperoleh dari pengguna media sosial. Berbagai sumber data, baik primer maupun sekunder, telah memberikan landasan yang kuat untuk pemahaman mendalam terhadap respons masyarakat terhadap tren fashion. Sentimen positif mendominasi dengan tingkat akurasi yang signifikan, menunjukkan adanya apresiasi yang tinggi terhadap tren fashion tertentu di kalangan pengguna media sosial. Meskipun demikian, adanya tingkat akurasi yang lebih rendah untuk kelas sentimen lainnya menunjukkan kompleksitas dalam menganalisis bahasa informal dan variasi respons masyarakat.

DAFTAR PUSTAKA

- [1] S. D. Pramukti, A. Nugroho, dan A. S. Sunge, "Analisis Sentimen Masyarakat Dengan Metode Naïve Bayes dan Particle Swarm Optimization," *Techno.Com*, vol. 21, no. 1, hal. 61–74, 2022, doi: 10.33633/tc.v21i1.5332.

- [2] M. I. Fikri, T. S. Sabrila, dan Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Smatika J.*, vol. 10, no. 02, hal. 71–76, 2020, doi: 10.32664/smatika.v10i02.455.
- [3] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, dan R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, hal. 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [4] D. Duei Putri, G. F. Nama, dan W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, hal. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [5] M. I. Petiwi, A. Triayudi, dan I. D. Sholihati, "Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine," *J. Media Inform. Budidarma*, vol. 6, no. 1, hal. 542, 2022, doi: 10.30865/mib.v6i1.3530.
- [6] A. R. Isnain, A. I. Sakti, D. Alita, dan N. S. Marga, "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 1, hal. 31, 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [7] D. A. Agustina, S. Subanti, dan E. Zukhronah, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine," *Indones. J. Appl. Stat.*, vol. 3, no. 2, hal. 109, 2021, doi: 10.13057/ijas.v3i2.44337.
- [8] H. D. Wijaya dan S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, hal. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.
- [9] B. Laurensz dan Eko Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, hal. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [10] N. Fitriyah, B. Warsito, dan D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm)," *J. Gaussian*, vol. 9, no. 3, hal. 376–390, 2020, doi: 10.14710/j.gauss.v9i3.28932.
- [11] I. Romli dan A. T. Zy, "Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 4, no. 2, hal. 694–702, 2020.
- [12] D. P. Utomo dan M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, hal. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [13] L. O. Sihombing, H. Hannie, dan B. A. Dermawan, "Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naïve Bayes Classifier," *Edumatic J. Pendidik. Inform.*, vol. 5, no. 2, hal. 233–242, 2021, doi: 10.29408/edumatic.v5i2.4089.
- [14] M. Riefky dan W. Pramesti, "Sentiment Analysis of Southeast Asian Games (SEA Games) in Philippines 2019 Based on Opinion of Internet User of Social Media Twitter with K-Nearest Neighbor and Support Vector Machine," *J. Mat. Stat. dan Komputasi*, vol. 17, no. 1, hal. 26–41, 2020, doi: 10.20956/jmsk.v17i1.9947.
- [15] A. R. Rizqullah, A. Wedhasmara, R. I. Heroza, A. Putra, F. Fathoni, dan P. Putra, "Analisis Masalah Pada Data Review Aplikasi Terhadap Layanan E-Commerce Menggunakan Metode Text Classification," *J. Tekno Kompak*, vol. 16, no. 1, hal. 186, 2022, doi: 10.33365/jtk.v16i1.1448.
- [16] D. Normawati dan S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, hal. 697–711, 2021, [Daring]. Tersedia pada: <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>