

BAB I PENDAHULUAN

1.1 Latar Belakang

Di era digital yang terus berkembang pesat, sektor *e-commerce* telah mengalami transformasi signifikan dalam cara konsumen berbelanja. Platform *e-commerce* seperti *Amazon* menjadi salah satu tempat utama bagi konsumen untuk membeli berbagai produk, termasuk dalam kategori fashion. Kategori fashion merupakan salah satu kategori yang paling dinamis dan kompetitif di platform *e-commerce* [1]. Ulasan pengguna memainkan peran yang sangat penting dalam membantu konsumen lain membuat keputusan pembelian. Ulasan produk tidak hanya memberikan informasi mengenai kualitas dan fitur produk, tetapi juga mempengaruhi kepercayaan serta kredibilitas merek di mata konsumen. Selain itu, Ulasan juga berfungsi sebagai umpan balik yang berharga bagi penjual untuk meningkatkan kualitas produk dan layanan, sehingga meningkatkan reputasi dan penjualan[2]

Women dresses adalah produk fashion populer di *Amazon* yang tidak hanya digunakan untuk kebutuhan sehari-hari, tetapi juga untuk gaya hidup dan acara khusus. Dengan beragam pilihan model, merek, dan desain, produk ini sering mendapat ulasan mendalam dari konsumen mengenai kualitas, kenyamanan, dan daya tahan. Namun, dengan volume ulasan yang sangat besar, proses analisis manual menjadi tidak efisien dan memakan waktu. Hal ini menyulitkan untuk mendapatkan wawasan yang cepat dan akurat mengenai sentimen konsumen terhadap produk, terutama dalam kategori fashion di platform *Amazon*. Oleh karena itu, diperlukan sistem otomatis yang dapat mempercepat dan mengotomatiskan proses analisis sentimen.

Beberapa algoritma dapat digunakan dalam analisis sentimen untuk memproses ulasan produk, di antaranya adalah *Naive Bayes*, *Logistic Regression*, *Random Forest*, *K-Nearest Neighbor (KNN)*, dan *Convolutional Neural Network (CNN)*. Masing-masing algoritma memiliki kelebihan dan keterbatasan dalam konteks analisis sentimen. *Naive Bayes* dikenal karena kemampuannya dalam mengelola

data besar dengan akurasi yang tinggi, tetapi kadang memiliki keterbatasan dalam menangani kompleksitas data[3]. Algoritma ini mengasumsikan independensi antar fitur, yang jarang terjadi dalam teks ulasan produk yang sering kali mengandung konteks dan ketergantungan antar kata. *Logistic Regression* menawarkan keakuratan yang baik dalam mengklasifikasikan data biner, tetapi bisa kurang efektif dalam menangani data yang tidak seimbang [4]. Situasi yang sering terjadi dalam analisis sentimen, di mana sebagian kategori sentimen (misalnya, negatif atau netral) jauh lebih sedikit dibandingkan kategori lainnya. *Random Forest* menawarkan keakuratan yang baik dengan menggunakan beberapa pohon keputusan, tetapi bisa kurang efektif dalam menangani data yang sangat besar[5], dan bisa memperlambat proses pelatihan dan prediksi. Karena proses pelatihan yang memerlukan banyak memori dan waktu, serta kurang mampu memberikan interpretasi terhadap hasil model. *KNN* memiliki kelebihan dalam mengklasifikasikan data berdasarkan kesamaan, tetapi bisa lambat untuk dataset besar[6], karena algoritma ini memerlukan perhitungan jarak untuk setiap data uji terhadap seluruh data latih. *CNN* menawarkan akurasi yang tinggi dalam analisis sentimen, tetapi memerlukan sumber daya komputasi yang besar[7], dan dapat menjadi tidak efisien jika digunakan untuk dataset yang lebih kecil atau dalam kondisi dengan keterbatasan perangkat keras.

Untuk mengatasi permasalahan ini, penelitian ini mengusulkan penggunaan *Support Vector Machine (SVM)* sebagai solusi untuk menganalisis sentimen ulasan produk secara otomatis dan efisien. *SVM* dipilih karena kemampuannya dalam mencari *hyperplane* optimal yang memisahkan kelas-kelas sentimen dengan margin maksimal, bahkan dalam ruang fitur yang kompleks. *SVM* telah terbukti memiliki akurasi yang tinggi dalam analisis sentimen, terutama dalam mengklasifikasikan ulasan positif dan negatif[5]. Salah satu kelebihan utama dari *SVM* adalah kemampuannya untuk menangani data yang tidak seimbang serta mengurangi risiko *overfitting*, yang sering menjadi masalah dalam analisis sentimen. Oleh karena itu, *SVM* menjadi pilihan yang sangat tepat untuk menganalisis sentimen ulasan produk di *Amazon*, khususnya dalam kategori *fashion Women dresses*.

Dengan menggunakan SVM, penelitian ini bertujuan untuk mengklasifikasikan ulasan produk *women dresses* di Amazon menjadi tiga kategori sentimen: positif, negatif, dan netral. Selain itu, hasil penelitian ini juga diharapkan dapat meningkatkan kualitas pengalaman belanja online bagi konsumen, serta membantu perusahaan dalam merancang produk dan layanan yang lebih sesuai dengan kebutuhan pasar.

1.2 Rumusan Masalah

Berdasar latar belakang masalah didapatkan rumusan masalah yaitu:

1. Bagaimana proses analisis sentimen menggunakan algoritma *Support Vector Machine (SVM)* untuk mengklasifikasikan ulasan produk menjadi sentimen positif, negatif, atau netral?
2. Apa saja faktor yang memengaruhi tingkat akurasi algoritma *SVM* dalam menganalisis sentimen ulasan produk di platform *e-commerce*?

1.3 Batasan Masalah

batasan penelitian antara lain:

1. Penelitian ini hanya akan menganalisis ulasan produk dengan kategori fashion , khususnya *Women dresses* yang diambil dari platform *e-commerce Amazon*.
2. Analisis sentimen yang dilakukan akan dibatasi pada tiga kategori sentimen, yaitu positif, negatif, dan netral.
3. Algoritma yang digunakan dalam penelitian ini adalah *Support Vector Machine (SVM)*, tanpa membandingkan dengan algoritma lainnya.
4. Bahasa pemrograman yang digunakan untuk membuat sistem adalah *python*

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah :

1. Untuk menganalisis sentimen ulasan produk di platform *e-commerce* menggunakan algoritma *Support Vector Machine (SVM)*.
2. Untuk mengevaluasi tingkat akurasi algoritma *SVM* dalam

mengklasifikasikan sentimen ulasan ke dalam kategori positif, negatif, atau netral.

1.5 Manfaat Penelitian

1.5.1 Manfaat Bagi Objek Penelitian

1. Membantu platform e-commerce mengidentifikasi kepuasan pelanggan: Analisis sentimen memungkinkan platform e-commerce seperti Amazon untuk mengetahui sentimen pelanggan terhadap produk, sehingga dapat melakukan perbaikan pada produk atau layanan yang mendapat ulasan negatif.
2. Mengoptimalkan pengambilan keputusan dalam pengembangan produk: Hasil analisis membantu perusahaan memperbaiki atau mengembangkan produk sesuai dengan kebutuhan dan preferensi pelanggan.
3. Meningkatkan efisiensi dalam merespons feedback pelanggan: Platform dapat dengan cepat merespons masalah yang dihadapi pelanggan, meningkatkan kepuasan dan mempertahankan reputasi.

1.5.2 Manfaat Bagi Peneliti Selanjutnya

1. Memberikan referensi metodologis untuk penelitian serupa: Penelitian ini memberikan dasar dalam menggunakan SVM untuk analisis sentimen, yang bisa diterapkan pada jenis produk atau platform lain.
2. Memberikan wawasan dalam pengolahan teks berbahasa Inggris: Penelitian ini memberikan solusi terhadap tantangan pengolahan teks ulasan dalam bahasa Inggris, berguna untuk penelitian lebih lanjut.
3. Membuka peluang pengembangan penelitian lanjutan: Penelitian ini memungkinkan pengembangan lanjutan dengan metode atau dataset berbeda untuk analisis sentimen yang lebih mendalam.

1.6 Sistematika Penulisan

Berisi sistematika penulisan skripsi yang memuat uraian secara garis besar isi skripsi untuk tiap-tiap bab.

BAB I PENDAHULUAN : Bab ini menguraikan tentang latar belakang

dilakukannya penelitian analisis sentimen menggunakan metode SVM, perumusan masalah yang akan diselesaikan, batasan-batasan masalah agar penelitian lebih fokus, tujuan yang ingin dicapai, manfaat yang diharapkan dari penelitian, serta sistematika penulisan skripsi

BAB II TINJAUAN PUSTAKA : Bab ini memaparkan tentang penelitian-penelitian terdahulu yang relevan dengan topik penelitian ini. Selain itu, bab ini juga menjelaskan berbagai teori yang mendukung penelitian, meliputi konsep analisis sentimen, teknik preprocessing data, penjelasan tentang algoritma Support Vector Machine (SVM), serta metode evaluasi yang digunakan dalam pengujian model.

BAB III METODE PENELITIAN : Bab ini menjelaskan tahapan-tahapan yang dilakukan dalam penelitian, mulai dari proses pengumpulan data ulasan, teknik preprocessing yang diterapkan pada data, implementasi metode SVM untuk klasifikasi sentimen, hingga metode pengujian dan evaluasi yang digunakan untuk mengukur performa model.

BAB IV HASIL DAN PEMBAHASAN : Bab ini menyajikan hasil implementasi sistem yang telah dibuat, termasuk hasil preprocessing data, hasil klasifikasi menggunakan SVM, serta analisis performa model berdasarkan berbagai metrik evaluasi. Pembahasan mencakup interpretasi hasil pengujian dan perbandingan dengan penelitian sejenis.

BAB V PENUTUP : Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian dan pembahasan yang telah dilakukan. Selain itu, bab ini juga memberikan saran-saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.