

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pasar *online* di Indonesia berkembang pesat dan kini menjadi salah satu perhatian utama bagi pengguna internet dan platform hiburan *online*. Beragam jenis layanan pasar tersedia dan dapat diakses oleh sebagian besar masyarakat Indonesia. Hal ini menunjukkan perubahan perilaku konsumen yang semakin mengandalkan layanan belanja *online* sebagai opsi utama untuk memenuhi kebutuhan dan keinginan mereka. Salah satu platform belanja *online* yang populer di kalangan masyarakat Indonesia adalah Tokopedia (Aulia & Hermawan, 2023).

Penjualan *online* saat ini berkembang sangat pesat karena transaksi menjadi semakin mudah hanya dengan genggaman tangan, kita sudah bisa mendapatkan produk yang diinginkan. Namun, perkembangan ini juga menghadirkan berbagai tantangan, seperti masalah dalam pelayanan, variasi produk, dan kualitas produk. Akibatnya, tidak sedikit pelanggan yang merasa kecewa dengan kondisi seperti ini (Ernawati *et al.*, 2023).

Ulasan konsumen dalam sistem *e-commerce* umumnya dianggap sebagai sumber daya penting yang mencerminkan pengalaman, perasaan, dan keinginan pengguna dalam membeli suatu produk. Informasi ini dapat mencakup pandangan konsumen yang mengungkapkan minat, sentimen, dan opini mereka. Berbagai penelitian menunjukkan bahwa orang cenderung lebih mempercayai sesama pengguna yang memiliki pandangan serupa terhadap suatu hal. *Data mining* dapat dimanfaatkan untuk menganalisis data pengguna media sosial yang mengunjungi platform *e-commerce*. Penelitian ini menggunakan teknik *data mining* yang bertujuan untuk membandingkan hasil klasifikasi dalam analisis sentimen dari pandangan pelanggan *E-commerce* yang telah dituliskan di Tokopedia. *Opinion Mining* atau yang dikenal juga sebagai analisis sentimen adalah proses yang menganalisis data teks yang dihasilkan dalam bentuk saran, umpan balik, tweet, dan komentar. Portal *E-commerce* menghasilkan sejumlah besar data setiap hari dalam

bentuk ulasan pelanggan. Menganalisis data dari *platform E-commerce* dapat membantu para *reseller online* memahami harapan pelanggan, memberikan pengalaman berbelanja yang lebih baik, dan meningkatkan penjualan (Sapanji *et al.*, 2023).

Algoritma *Support Vector Machine* (SVM) dikenal memiliki performa tinggi dalam klasifikasi teks, termasuk analisis sentimen, terutama ketika bekerja dengan data yang memiliki jumlah fitur besar, seperti teks ulasan pelanggan. Dalam analisis sentimen sering ada ulasan yang memiliki sentimen ganda, dibandingkan dengan beberapa algoritma lain, SVM memberikan hasil yang stabil dalam tugas klasifikasi biner, seperti dalam menentukan apakah ulasan positif atau negatif. Ini sangat penting untuk menjaga kualitas sentimen agar sesuai dengan kenyataan di lapangan (N. K. Putri *et al.*, 2024).

Algoritma *Naive Bayes* merupakan algoritma yang sederhana namun sangat efektif dalam menunjukkan performa dalam analisis sentimen teks. Algoritma ini mampu mengklasifikasikan teks berdasarkan probabilitas kata – kata yang muncul dalam kategori sentimen tertentu, seperti positif, negatif, atau netral. Dalam analisis ulasan, meskipun beberapa kata mungkin terkait, asumsi independensi ini tetap memungkinkan *Naive Bayes* untuk menghasilkan hasil klasifikasi yang baik. Data ulasan sering kali memiliki bahasa atau pola yang beragam. *Naive Bayes* mampu melakukan klasifikasi pada data yang tidak konsisten dengan cukup baik karena pendekatan probabilitiknya (Hanafi & R, 2024).

Permasalahan utama yang mendorong penelitian ini adalah banyaknya ulasan tidak terstruktur di Tokopedia, termasuk bahasa informal, campuran, dan emotikon, yang sulit dianalisis secara manual untuk memahami sentimen positif, negatif, atau netral. Selain itu, terdapat tantangan dalam memilih algoritma yang paling akurat, seperti SVM dan *Naive Bayes*, untuk mengolah data besar dengan variasi bahasa informal.

Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen terhadap ulasan produk di platform *E-commerce* Tokopedia dengan menggunakan algoritma *Support Vector Machine* (SVM) dan algoritma *Naive Bayes*. Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen dari ulasan

pelanggan, apakah bersifat positif, negatif. Dengan menggunakan SVM dan *Naive Bayes* penelitian ini dapat menentukan algoritma yang lebih efektif dalam mengolah data ulasan dan memberikan hasil klasifikasi sentimen yang akurat. Hasil analisis ini diharapkan dapat memberikan wawasan bagi Tokopedia terkait persepsi pelanggan terhadap produk, membantu meningkatkan kualitas layanan, dan memberikan saran untuk pengembangan strategi pemasaran yang lebih baik.

Penggunaan algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* pada penelitian ini memiliki beberapa perbandingan pada aspek - aspek tertentu seperti dalam aspek akurasi dan ketepatan klasifikasi, kecepatan pemrosesan, kemampuan mengatasi dimensi tinggi dan kesesuaian dengan data skala besar.

Penelitian sebelumnya telah membahas penggunaan algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* untuk menganalisis sentimen, penelitian yang ditulis oleh Sugeng Setyabudi dan Enny Aryanny berjudul Analisis Sentimen Penilaian Pengguna *Marketplace* Lazada Dengan Metode *Naive Bayes* Dan *Support Vector Machine*, hasil penelitian menunjukkan bahwa SVM lebih unggul dalam mengklasifikasikan sentimen pengguna Lazada dengan akurasi 75%, dibandingkan *Naive Bayes* yang mencapai 72% (Setyabudi & Aryanny, 2025).

Penelitian penulis juga mengambil metode yang sama seperti penelitian sebelumnya yaitu metode *Support Vector Machine* (SVM) dan *Naive Bayes* dengan permasalahan Analisis Sentimen Ulasan Produk *E - Commerce* Pada Tokopedia Menggunakan Algoritma *Support Vector Machine* (SVM) dan *Naive Bayes*, analisis sentimen ulasan produk *E- Commerce* pada tokopedia dengan membandingkan metode *Naive Bayes* dan *Support Vector Machine* (SVM) untuk menentukan algoritma yang lebih akurat dalam klasifikasi sentimen.

1.2 Rumusan Masalah

Berikut permasalahan yang terkait dengan penelitian ini dirumuskan berdasarkan latar belakang di atas:

1. Bagaimana cara mengklasifikasikan sentimen ulasan produk pada Tokopedia menjadi kategori positif dan negatif menggunakan algoritma *Support Vektor Machine* (SVM) dan *Naive Bayes* ?

2. Seberapa akurat algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* dalam mengidentifikasi sentimen dari ulasan produk di Tokopedia ?
3. Bagaimana analisis sentimen terhadap ulasan produk dapat memberikan wawasan tentang persepsi pelanggan dan mempengaruhi strategi pemasaran di Tokopedia ?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menentukan cara klasifikasi ulasan produk pada Tokopedia menjadi kategori positif dan negatif dengan menggunakan algoritma *Support Vector Machine* (SVM) dan *Naive Bayes*.
2. Mengevaluasi tingkat akurasi algoritma SVM dan *Naive Bayes* dalam mengidentifikasi sentimen ulasan produk di Tokopedia.
3. Menganalisis bagaimana hasil analisis sentimen ulasan produk dapat memberikan wawasan tentang persepsi pelanggan dan bagaimana wawasan tersebut dapat digunakan untuk menyusun strategi pemasaran di Tokopedia.

1.4 Manfaat Penelitian

Adapun beberapa manfaat dari penelitian yang dilakukan yaitu sebagai berikut:

1. Penelitian ini membantu memahami bagaimana sentimen positif dan negatif dari ulasan pengguna terhadap aplikasi Tokopedia. Informasi ini berguna bagi pengelola untuk meningkatkan layanan dan produk.
2. Hasil analisis sentimen dapat memberikan wawasan tentang aspek-aspek yang memerlukan perbaikan, sehingga perusahaan dapat membuat strategi untuk meningkatkan pengalaman pengguna.
3. Penelitian ini menggunakan algoritma *Support Vektor Machine*(SVM) dan *Naive Bayes* sehingga membantu menentukan metode mana yang lebih efektif dalam pengolahan data teks.
4. Informasi yang diperoleh dari hasil analisis dapat digunakan oleh manajemen untuk membuat keputusan yang lebih akurat berdasarkan data

pelanggan. Penelitian ini mendukung pengembangan teknologi dalam pengolahan bahasa alami dan pengaplikasian algoritma machine learning dalam industri *e-commerce*.

5. Hasil penelitian dapat menjadi referensi untuk penelitian-penelitian selanjutnya yang berfokus pada analisis sentimen di platform lain atau pengembangan metode yang lebih canggih.

1.5 Ruang Lingkup Dan Batasan Penelitian

Adapun ruang lingkup dan batasan penelitian adalah sebagai berikut:

1. Hanya mengambil ulasan di aplikasi Tokopedia terkait ulasan produk kecantikan Skintific di toko skintificid .
2. Data ulasan yang diambil 1 tahun di tahun 2024.
3. Fokus hanya data ulasan teks yang diambil.
4. Pada ulasan yang panjang, hanya kalimat yang muncul di halaman yang diambil, sementara sisanya yang tidak muncul tidak diambil.
5. Ulasan yang digunakan adalah ulasan publik yang dapat diakses, sehingga ulasan palsu atau ulasan yang tidak representatif mungkin ikut teranalisis.
6. Analisis hanya mengukur sentimen secara umum (positif dan negatif) tanpa memisahkan aspek tertentu, seperti sentimen terhadap produk, layanan pelanggan, atau pengiriman.
7. Variabel yang digunakan variabel *input* (fitur) dan variabel *output* (target). Variabel *input* seperti teks ulasan pelanggan, panjang teks ulasan, frekuensi kata kunci sentimen, rating produk, jumlah kata negatif atau positif.
8. Karakter data yang digunakan seperti teks tidak terstruktur, kombinasi bahasa formal dan informal, mengandung kata-kata sentimen, data berskala besar, dan data berformat *CSV*.