

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi telah mengubah cara konsumen dalam berinteraksi dengan produk dan layanan. Perkembangan ini membawa berbagai keuntungan dan perubahan gaya hidup masyarakat dalam hal cara berjualan dan berbelanja [1]. Dengan kemudahan saat ini ditunjukkan dengan semakin banyaknya aplikasi belanja online yang menyediakan fitur ulasan produk yang dimana memiliki manfaat untuk memudahkan dalam membantu pengguna menemukan produk yang dicari dan dibutuhkan.[2].

Di Indonesia, industri perawatan diri dan kecantikan mengalami pertumbuhan yang sangat besar. Melansir data dari Katadata yang dikutip oleh CISAS, sektor perawatan tubuh (*body care*) memiliki volume pasar tertinggi sebesar 3,18 miliar USD pada tahun 2022, dengan proyeksi kenaikan tahunan sebesar 5,81% (CAGR) hingga tahun 2027. Skincare memiliki volume pasar sebesar 2,05 miliar USD dan makeup memiliki volume pasar sebesar 1,61 miliar USD, masing-masing.[3] Angka-angka tersebut menunjukkan bahwa minat masyarakat terhadap produk kecantikan sangat tinggi dan terus meningkat. Hal ini tidak hanya membuka peluang besar bagi pelaku industri, tetapi juga mendorong munculnya kebutuhan akan pemahaman yang lebih mendalam terhadap perilaku dan preferensi konsumen, terutama melalui ulasan-ulasan online yang mereka berikan di berbagai platform kecantikan.

Ulasan pelanggan online sangat membantu untuk memperoleh informasi bagi calon pembeli. Data menunjukkan 93% konsumen yang membaca ulasan online untuk membuat keputusan pembelian dan 85% dari mereka mempercayai ulasan tersebut[4]. Hal ini membuktikan bahwa ulasan pelanggan online sebagai salah satu faktor utama yang dapat memberikan dampak signifikan dalam membentuk persepsi dan keyakinan konsumen sebelum melakukan pembelian[5].

Dalam industri kecantikan, platform seperti *Female Daily* menjadi tempat utama bagi konsumen untuk mencari referensi sebelum membeli produk. *Female Daily* merupakan platform khusus berbasis web maupun *mobile phone* yang berisi konten produk kecantikan tentang perawatan wajah hingga tubuh untuk siap diberi

ulasan oleh konsumen yang telah mencoba atau sedang menggunakan produknya[2]. Setiap produk kecantikan umumnya memiliki kelebihan dan kekurangan. Pengalaman pengguna ini yang kemudian dituangkan dalam sebuah *review* produk, baik itu kritik atau penyampaian kepuasan terhadap penggunaan produk tersebut. Sentimen positif atau negatif pengguna terhadap suatu produk merujuk ke beberapa aspek produk. Misalkan pengguna suka terhadap *packaging* suatu produk tetapi harganya terlalu mahal.

Maka dari permasalahan tersebut dibangun sebuah sistem yang dapat dilakukan yaitu *Aspect Based Sentiment Analysis (ABSA)*. Meskipun analisis sentimen terhadap ulasan produk kecantikan telah banyak dilakukan, tetapi sebagian besar penelitian tersebut masih bersifat umum dan hanya mengkategorikan sentimen sebagai positif atau negatif tanpa memperhatikan aspek spesifik yang menjadi perhatian konsumen. Padahal, ulasan konsumen sering kali mencakup penilaian berbagai bagian produk yang berbeda-beda. Misalnya, seorang konsumen mungkin puas dengan kualitas produk, tetapi dia mungkin tidak menyukai kemasannya atau harganya. Hal ini menunjukkan bahwa metode analisis yang lebih mendalam dan terarah seperti Analisis Sentimen Berbasis Aspek (ABSA) diperlukan.

Analisis sentimen berbasis aspek adalah proses mengekstraksi sentimen dari setiap aspek yang terkandung dalam sebuah kalimat [6]. Analisis sentimen berbasis aspek atau ABSA merupakan bentuk lanjutan dari analisis sentimen yang berfokus pada kalimat. ABSA melibatkan dua tugas utama, yaitu ekstraksi aspek untuk mengidentifikasi aspek yang dibahas dan klasifikasi aspek untuk menentukan sentimen sebagai positif atau negatif. Dengan pendekatan ini, ABSA memungkinkan pengungkapan aspek atau atribut produk yang diekspresikan oleh pengguna[7]. ABSA lebih baik daripada klasifikasi sentimen berbasis kalimat karena mempertimbangkan istilah aspek yang ada dalam ulasan untuk menentukan sentimen daripada mempertimbangkan kalimat individual[8].

ABSA terdiri dari 2 proses yaitu *aspect extraction* dan *sentiment analysis*[9]. Analisis sentimen berbasis aspek memerlukan langkah *Aspect Term Extraction* untuk mengetahui aspek yang dibahas oleh suatu sentimen[10]. *Latent Dirichlet Allocation (LDA)* merupakan metode yang populer pada *aspect extraction*[11]. *Latent Dirichlet Allocation (LDA)* adalah model probabilistik yang digunakan

untuk identifikasi topik dalam kumpulan dokumen. LDA mengasumsikan bahwa setiap dokumen adalah campuran dari beberapa topik, dan setiap topik adalah distribusi probabilitas dari kata-kata. Dengan menggunakan LDA, kita dapat menemukan pola tersembunyi dalam data teks dan mengelompokkan kata-kata yang sering muncul bersama ke dalam topik yang sama[12].

Penelitian oleh Hilal et al. (2024) menunjukkan efektivitas metode *Latent Dirichlet Allocation* (LDA) efektif dalam proses ekstraksi aspek pada ulasan pengguna aplikasi PLN Mobile yang tersedia di *Google PlayStore*. Dua komponen penting *Application Service* dan *Customer Service* diidentifikasi dengan sukses oleh LDA menggunakan pendekatan *topic modeling*. Dengan nilai *coherence score* sebesar 0,52, evaluasi model menunjukkan bahwa kata-kata dalam setiap topik memiliki tingkat keselarasan semantik yang cukup baik. Hasil ini menunjukkan bahwa LDA dapat digunakan untuk analisis sentimen berbasis aspek karena dapat mengekstraksi aspek secara otomatis dan bermakna dari data teks tidak terstruktur. [10].

Dalam analisis sentimen, *Support Vector Machine* (SVM) adalah salah satu metode yang sering digunakan. Kelebihan menggunakan SVM pada analisis sentimen adalah SVM akan bekerja baik jika terdapat batas margin yang jelas antara kelas, selain itu algoritma SVM memberi hasil yang lebih akurat jika dibandingkan dengan algoritma lain yang lebih sederhana seperti naïve bayes[13].

Pada penelitian sebelumnya oleh Kustyaningsih dan Permana (2024) yang menggunakan pemodelan topik dengan *Latent Dirichlet Allocation* (LDA) untuk menentukan aspek utama, serta klasifikasi sentimen menggunakan *Support Vector Machine* (SVM). Dari hasil pemodelan topik, ditemukan tiga aspek utama, yaitu *Usability*, *Reliability*, dan *Performance Efficiency*, dengan nilai koherensi tertinggi sebesar 0,487. Sedangkan klasifikasi menggunakan SVM dengan pembagian data 90:10 menghasilkan akurasi 85,45%, yang meningkat menjadi 90% setelah penerapan metode resampling SMOTE[14].

Selanjutnya, penelitian yang dilakukan oleh Arthamevia et al. (2021) memanfaatkan algoritma *Support Vector Machine* (SVM) dengan fitur yang diekstraksi menggunakan metode TF-IDF untuk melakukan klasifikasi sentimen. Penelitian ini membandingkan berbagai jenis kernel dan menemukan bahwa kernel

linear dengan parameter $C = 1$ memberikan hasil terbaik. Hasil pengujian menunjukkan bahwa model SVM memiliki akurasi tertinggi sebesar 88,35% dengan konfigurasi ini. Kernel lain, seperti RBF, polynomial, dan sigmoid, memiliki akurasi yang lebih rendah. Hasil ini menegaskan bahwa SVM dapat digunakan dengan baik untuk mengklasifikasikan sentimen dalam data teks yang besar[1].

Berdasarkan penelitian sebelumnya, metode usulan yang akan digunakan dalam analisis sentimen berbasis aspek pada *review* produk di website *Female Daily* adalah kombinasi *Latent Dirichlet Allocation (LDA)* dan *Support Vector Machine (SVM)*. LDA akan digunakan sebagai teknik ekstraksi aspek untuk mengidentifikasi topik utama dalam ulasan produk, seperti harga, kemasan, dan aroma. Selanjutnya, hasil ekstraksi aspek akan digunakan sebagai fitur dalam klasifikasi sentimen menggunakan SVM, yang terbukti memiliki akurasi tinggi dalam membedakan sentimen positif dan negatif. Kombinasi metode ini diharapkan dapat menghasilkan analisis sentimen berbasis aspek yang lebih akurat dan mendalam, sehingga memberikan wawasan yang lebih jelas bagi pengembang produk dan pemasar dalam memahami persepsi konsumen terhadap berbagai aspek produk kecantikan.

1.2 Rumusan Masalah

1.2.1 Permasalahan

1. Ulasan produk kecantikan di *Female Daily* bersifat bebas, tidak terstruktur, panjangnya bervariasi, dan banyak menggunakan bahasa informal, sehingga menyulitkan dalam mengidentifikasi aspek-aspek utama yang sering dibahas pengguna.
2. Klasifikasi sentimen pada ulasan produk menghadapi tantangan karena data teks memiliki dimensi fitur yang tinggi dan hubungan antar kata yang kompleks.

1.2.2 Metode Usulan

Untuk menjawab permasalahan yang telah diidentifikasi, penelitian ini mengusulkan kombinasi dua metode utama yaitu :

1. Untuk mengatasi kesulitan dalam mengenali aspek-aspek utama dari ulasan produk kecantikan di *Female Daily* yang bersifat tidak terstruktur dan

menggunakan bahasa informal, digunakan metode *Latent Dirichlet Allocation* (LDA). LDA dapat memetakan kumpulan kata yang sering muncul bersamaan menjadi beberapa kelompok topik yang mewakili aspek-aspek produk, sehingga aspek dominan dapat ditemukan secara otomatis tanpa memerlukan data berlabel.

2. Untuk mengatasi tantangan klasifikasi sentimen pada data teks berdimensi tinggi dan kompleks, digunakan algoritma *Support Vector Machine* (SVM). SVM bekerja dengan mencari *hyperplane* atau garis pemisah terbaik untuk memisahkan data ke dalam dua kelas berbeda, sehingga setiap ulasan yang telah dipetakan ke aspek tertentu dapat diklasifikasikan polaritas sentimennya menjadi positif atau negatif.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang dan serta identifikasi masalah yang telah dipaparkan, penelitian ini dirumuskan dalam beberapa pertanyaan sebagai berikut :

1. Aspek-aspek utama apa saja yang dapat diidentifikasi dari ulasan pengguna *Female Daily* terkait produk kecantikan menggunakan metode *Latent Dirichlet Allocation* (LDA)?
2. Berapa tingkat akurasi algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen terhadap masing-masing aspek yang dihasilkan dari *Latent Dirichlet Allocation* (LDA)?

1.3 Batasan Masalah

Untuk menjaga penelitian ini tetap terarah dan fokus, ditetapkan beberapa batasan sebagai berikut :

1. Ulasan yang dianalisis terbatas pada yang menggunakan Bahasa Indonesia, sementara ulasan dalam bahasa lain tidak akan dipertimbangkan.
2. Penelitian ini tidak melakukan perbandingan metode, baik untuk ekstraksi aspek maupun klasifikasi sentimen.
3. Sentimen yang diklasifikasikan dibatasi pada dua kelas, yaitu positif dan negatif, tanpa mempertimbangkan sentimen netral atau ambiguitas opini.
4. Hasil klasifikasi sentimen tidak dianalisis berdasarkan kategori produk tertentu.

5. Data ulasan yang digunakan berjumlah 3.216 ulasan dengan rentang waktu tahun 2020 hingga 2025.
6. Penelitian tidak menganalisis pengaruh *rating* bintang terhadap hasil *sentiment analysis*.
7. Aspek atau topik yang dianalisis telah ditetapkan secara khusus menjadi empat aspek

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan

Penelitian ini memiliki tujuan yang ingin dicapai sebagai berikut :

1. Untuk mengidentifikasi aspek-aspek utama yang terkandung dalam ulasan pengguna *Female Daily* terkait produk kecantikan dengan menggunakan metode pemodelan topik *Latent Dirichlet Allocation* (LDA).
2. Untuk mengukur tingkat akurasi algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen pengguna terhadap masing-masing aspek yang diperoleh dari hasil pemodelan topik LDA.

1.4.2 Manfaat

Manfaat yang diharapkan dengan dilakukannya penelitian ini adalah :

1. Secara teoritis, dalam penelitian ini diharapkan menjadi referensi dan menambah literatur dalam pengembangan *text mining* dan *machine learning* untuk analisis sentimen berbasis aspek dalam *review* produk.
2. Secara praktis, menyediakan wawasan bagi pengelola produk kecantikan dalam memahami persepsi konsumen terhadap berbagai aspek produk.

1.5 Sistematika Penulisan

Dalam penyusunan laporan penelitian ini pembahasan terbagi menjadi tiga bab yang secara singkat akan diuraikan sebagai berikut :

1. BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang penelitian yang mendasari pentingnya analisis sentimen berbasis aspek pada ulasan produk kecantikan di *Female Daily*. Selain itu, bab ini juga berisi rumusan masalah, batasan masalah, tujuan penelitian, serta manfaat penelitian bagi berbagai pihak.

2. BAB II KAJIAN PUSTAKA

Bab ini menguraikan teori-teori yang mendukung penelitian, termasuk konsep analisis sentimen berbasis aspek, *preprocessing*, serta metode *Latent Dirichlet Allocation* (LDA) dan *Support Vector Machine* (SVM). Selain itu, dibahas pula penelitian terdahulu yang relevan untuk memberikan landasan bagi metode yang digunakan dalam penelitian ini.

3. BAB III METODE PENELITIAN

Bab ini menjelaskan secara rinci metode yang digunakan dalam penelitian, mencakup sumber dan teknik pengumpulan data, proses *preprocessing* teks, ekstraksi aspek menggunakan LDA, serta klasifikasi sentimen menggunakan SVM. Selain itu, dijelaskan pula teknik evaluasi model menggunakan *accuracy*, *precision*, *recall*, *F1-score*.

4. BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian yang meliputi *preprocessing* data, ekstraksi aspek menggunakan LDA, dan klasifikasi sentimen menggunakan SVM. Bab ini juga menjelaskan skenario pengujian, seperti pembagian data latih dan data uji, jumlah topik LDA, serta konfigurasi SVM yang digunakan. Selain itu, ditampilkan hasil evaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*, beserta pembahasannya untuk melihat efektivitas metode yang diterapkan.

5. BAB V KESIMPULAN DAN SARAN

Bab ini berisi ringkasan dari keseluruhan hasil penelitian yang telah diperoleh, termasuk kesimpulan mengenai penerapan LDA dalam ekstraksi aspek dan SVM dalam klasifikasi sentimen pada ulasan *Female Daily*. Selain itu, bab ini juga memuat saran yang dapat dijadikan acuan untuk penelitian selanjutnya, baik dalam pengembangan metode, peningkatan kualitas data, maupun penerapan teknologi yang lebih mutakhir untuk analisis sentimen berbasis aspek.