

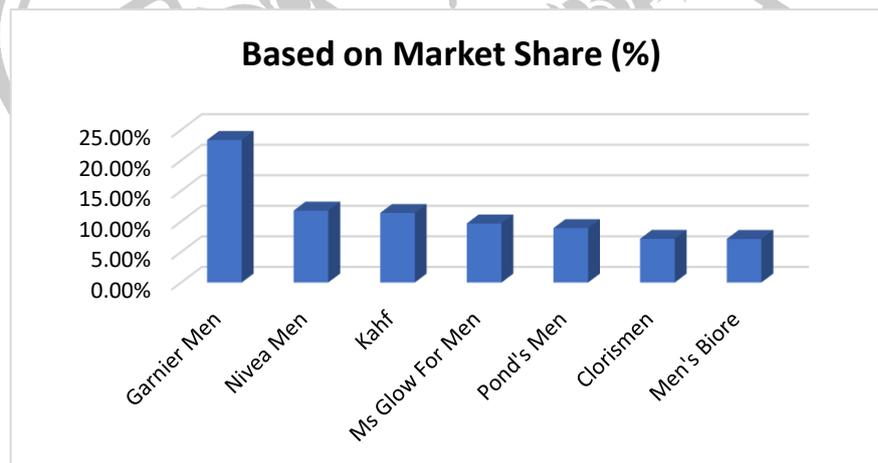
# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Facial wash*, merupakan salah satu produk perawatan wajah yang banyak digunakan Masyarakat. Fungsinya adalah untuk membersihkan kotoran, minyak berlebih, dan sisa make-up yang menempel pada wajah. Perawatan wajah adalah aspek penting yang perlu diperhatikan oleh baik pria maupun wanita. Tujuan perawatan bukan sekedar untuk penampilan, tetapi juga untuk menjaga kulit tetap optimal (Setyaningsih dkk, 2023). Penggunaan *facial wash* secara rutin tidak hanya membantu mencegah timbulnya masalah kulit seperti jerawat dan komedo, tetapi juga menjaga keseimbangan pH kulit agar tetap sehat. Selain itu, semakin banyaknya varian *facial wash* dengan kandungan yang disesuaikan untuk berbagai jenis kulit menjadikan produk ini semakin diminati oleh masyarakat (Prastika dkk, 2021). Pengguna produk perawatan wajah menganggap *facial wash* sebagai produk yang paling penting dalam rangkaian *skincare* mereka. Hal ini menunjukkan bahwa *facial wash* bukan hanya sekedar pelengkap, tetapi menjadi bagian esensial dalam rutinitas perawatan wajah sehari-hari.

Beragam merk *facial wash* seperti Garnier Men, Nivea Men, Kahf dan lain-lain bersaing di pasar dan menjadi merk yang populer di pasar Indonesia.



Gambar 1.1 Grafik berdasarkan data riset internal tim compas

Berdasarkan grafik di atas, yang diperoleh melalui *crawling* data *e-commerce* Shopee dan Tokopedia pada kategori pembersih wajah pria, dapat dilihat bahwa

Garnier Men mendominasi dengan pangsa pasar tertinggi, diikuti oleh Nivea Men, Kahf, serta beberapa merek lainnya seperti Ms Glow for Men dan Pond's Men. Kahf menempati posisi ketiga dengan persentase yang hampir setara dengan Nivea Men, menunjukkan bahwa produk ini mampu bersaing secara kompetitif dengan merek internasional. Jika dilihat dari sisi brand lokal, maka Kahf menempati posisi teratas sebagai produk dalam negeri dengan performa terbaik. Hal ini mencerminkan adanya dukungan dan kepercayaan konsumen terhadap produk lokal serta memperlihatkan potensi besar Kahf untuk terus berkembang di industri perawatan wajah pria.

*Facial wash* Kahf dikenal sebagai produk yang halal, aman, dan berbahan alami, cocok untuk berbagai jenis kulit. Produk ini memiliki beberapa varian untuk mengatasi masalah seperti kulit berminyak, jerawat, dan kusam (Nabilla dan Hafiar, 2024). Kombinasi kemasan yang praktis dan harga yang sesuai dengan daya beli konsumen menengah membuat produk ini semakin diminati. Sebagian besar ulasan menunjukkan kepuasan konsumen, meskipun ada juga kritik terkait efek samping atau ketidakcocokan pada kulit tertentu. Karena itu, analisis ulasan konsumen penting untuk memahami persepsi mereka, apalagi di era digital yang memengaruhi cara orang berbelanja.

Perkembangan teknologi telah mempermudah masyarakat dalam memilih dan membeli produk. Konsumen kini cenderung memanfaatkan internet dan *e-commerce* untuk mencari informasi dan membaca ulasan sebelum membeli (Alamin dkk, 2023). Ulasan pengguna dianggap lebih objektif dibanding iklan. Penelitian oleh Akmal dkk. (2024) menunjukkan bahwa jumlah pengguna *e-commerce* di Indonesia mengalami tren kenaikan signifikan dari tahun 2019 hingga 2023, dan diprediksi akan mencapai lebih dari 245 juta pengguna pada tahun 2025 berdasarkan model *regresi linear time series*. Dalam konteks perubahan perilaku belanja yang semakin bergantung pada ulasan online, *e-commerce* seperti Shopee memainkan peran penting dalam memfasilitasi proses tersebut.

Shopee merupakan platform *e-commerce* paling populer di Indonesia dengan beragam produk (Alamin dkk, 2023). Sejak diluncurkan pada 2015, Shopee menarik banyak pengguna melalui fitur seperti gratis ongkir, *flash sale*, dan *shopee live* (Adha dan Hayani, 2024). Shopee juga menyediakan sistem ulasan dan rating

yang transparan, membantu pembeli menilai produk berdasarkan pengalaman pengguna lain. Pada awal 2024, Shopee mencatat 228 juta kunjungan, menjadikannya sumber data potensial untuk analisis sentimen. Dengan banyaknya ulasan yang tersedia di Shopee, diperlukan metode analisis yang mampu mengelola data teks secara efektif untuk mengetahui persepsi konsumen salah satunya *Multinomial Naïve Bayes*.

*Multinomial Naïve Bayes* merupakan algoritma *machine learning* yang banyak digunakan dalam klasifikasi teks, terutama untuk analisis sentimen. Algoritma ini didasarkan pada teorema Bayes dan mengasumsikan bahwa fitur-fitur dalam dataset bersifat independen satu sama lain. Dalam konteks analisis sentimen, *Multinomial Naïve Bayes* bekerja dengan menghitung probabilitas kemunculan kata-kata tertentu dalam sebuah ulasan untuk menentukan apakah sentimen yang terkandung bersifat positif, negatif, atau netral. Algoritma ini dikenal memiliki kinerja yang baik pada data teks yang bersifat multinomial, seperti frekuensi kata dalam dokumen. Namun, pada dataset yang tidak seimbang, performa model dapat menurun karena kecenderungan bias terhadap kelas mayoritas. Untuk mengatasi hal tersebut, penelitian ini menerapkan *SMOTE (Synthetic Minority Over-sampling Technique)*, yaitu teknik oversampling yang menghasilkan data sintesis pada kelas minoritas untuk menyeimbangkan distribusi kelas pada data latih (Moulaei dkk., 2023).

Menurut Setiawan dan Rachmawati (2022), algoritma *Multinomial Naïve Bayes* unggul dalam hal kecepatan, kemudahan implementasi, dan cocok digunakan untuk data berukuran besar. Hal ini membuatnya ideal untuk klasifikasi teks secara cepat. Namun, kelemahan utamanya adalah asumsi bahwa setiap fitur bersifat independen, padahal dalam teks alami fitur sering saling terkait, sehingga bisa menurunkan akurasi. Jika dibandingkan dengan *Support Vector Machine (SVM)*, *Naïve Bayes* lebih cepat saat training, tetapi *SVM* biasanya lebih akurat karena mampu menangani data kompleks. Dibandingkan dengan *Decision Tree*, *Naïve Bayes* lebih tahan terhadap overfitting, meski kurang fleksibel dalam mengelola fitur yang saling bergantung. Putra (2023) juga menekankan bahwa performa *Naïve Bayes* bisa ditingkatkan dengan teknik *preprocessing* seperti *stopword removal*, *stemming*, dan penggunaan *TF-IDF*.

Beberapa studi menunjukkan bahwa algoritma ini bisa mencapai akurasi tinggi jika digunakan pada data teks yang sesuai, sehingga tetap menjadi pilihan populer untuk analisis sentimen yang mengutamakan efisiensi. Salah satunya penelitian Setyaningsih dkk. (2023) yang menggunakan algoritma *Naïve Bayes* untuk menganalisis opini masyarakat tentang *skincare* di Twitter, dengan capaian akurasi 86%, presisi 84%, dan recall 80%. Berbeda dengan penelitian tersebut, penelitian ini menerapkan algoritma *Multinomial Naïve Bayes* dengan *TF-IDF*, *SMOTE*, dan *K-Fold Cross Validation* untuk mengklasifikasikan *ulasan facial wash* Kahf di Shopee. Perbandingan hasil kedua penelitian diharapkan dapat memberikan gambaran perbedaan performa model pada jenis data dan platform yang berbeda.

Oleh karena itu, penerapan *Multinomial Naïve Bayes* dalam menganalisis sentimen pengguna terhadap produk *facial wash* Kahf di Shopee diharapkan dapat memberikan hasil yang akurat dan bermanfaat bagi perusahaan dalam memahami persepsi konsumen.

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, dapat dirumuskan permasalahan sebagai berikut yaitu:

1. Bagaimana hasil analisis sentimen ulasan pengguna shopee terhadap produk *facial wash* kahf yang diklasifikasikan menggunakan algoritma *Multinomial Naïve Bayes*?
2. Bagaimana hasil evaluasi performa model klasifikasi sentimen berdasarkan akurasi, presisi, dan recall?

## 1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini yaitu:

1. Mengetahui hasil analisis sentimen ulasan pengguna shopee terhadap produk *facial wash* kahf yang diklasifikasikan menggunakan algoritma *Multinomial Naïve Bayes*.

2. Mengetahui hasil evaluasi performa model klasifikasi sentimen berdasarkan metrik akurasi, presisi, dan recall untuk mengetahui efektivitas dan akurasi model yang dibangun.

#### 1.4 Manfaat Penelitian

Manfaat dari tercapainya penelitian ini yaitu:

1. Penelitian ini diharapkan mampu memberikan pemahaman yang mendalam tentang kepuasan pelanggan terhadap produk pembersih wajah kahf melalui analisis sentimen mereka masing-masing. Hasil dari studi ini dapat digunakan untuk menentukan kecenderungan sentimen positif dan negatif terhadap ulasan yang ditinggalkan oleh pengguna shopee serta menilai aspek produk mana yang perlu diperbaiki untuk meningkatkan kualitas dan kepuasan pelanggan.
2. Penelitian ini bermanfaat dalam menyelesaikan proyek akhir serta memberikan wawasan tentang penerapan analisis sentimen yang didukung oleh pembelajaran mesin, khususnya pada penggunaan algoritma *Multinomial Naïve Bayes* pada data ulasan *e-commerce*. Dengan model klasifikasi sentimen yang akurat, penelitian ini dapat digunakan untuk membantu dalam pengambilan keputusan mengenai pengembangan produk berdasarkan data ulasan pengguna.

#### 1.5 Batasan Penelitian

Adapun beberapa batasan masalah yang diangkat dalam melakukan penelitian ini, antara lain:

1. Data yang digunakan berasal dari platform *e-commerce* Shopee, berupa 600 komentar teks terkait produk *facial wash* kahf yang dikumpulkan dalam rentang waktu 2024–2025. Dataset yang digunakan hanya mencakup data teks, tanpa melibatkan rating, elemen video dan gambar.
2. Metode yang diterapkan adalah *Multinomial Naïve Bayes* tanpa perbandingan metode yang lain.

3. Data awal mencakup berbagai jenis komentar, namun melalui pra-pemrosesan hanya ulasan terkait produk yang dipertahankan, sedangkan komentar tentang pengiriman dan layanan logistik dihilangkan.
4. Analisis sentimen yang dilakukan mengklasifikasikan ulasan menjadi 2 kelas, yaitu positif dan negatif.
5. Validasi model dilakukan menggunakan *K-Fold Cross Validation* dengan nilai *K* sebesar 4, 5, 6, dan 10.

