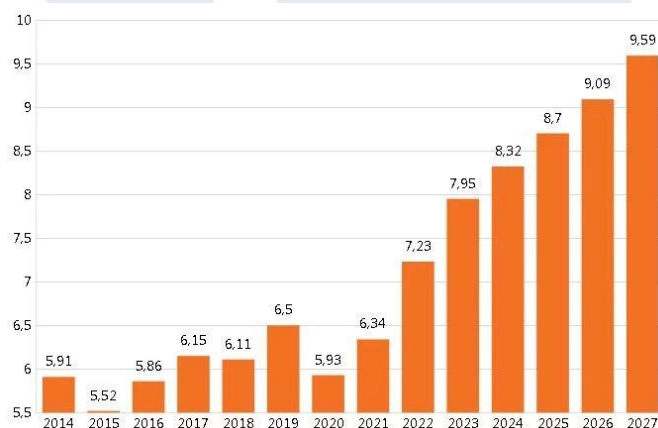


BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini industri *skincare* telah mengalami peningkatan yang signifikan di seluruh dunia, termasuk di Indonesia. Kesadaran masyarakat mengenai pentingnya merawat kulit tidak lagi terbatas pada wanita tapi juga ke berbagai kelompok usia dan gender. Pertumbuhan pasar kosmetik di Indonesia terus meningkat dalam beberapa tahun terakhir, berdasarkan dari laporan Statista pendapatan dari sektor kecantikan dan perawatan diri mencapai 111,83 triliun pada tahun 2022 lalu, sehingga diperkirakan dapat mengalami kenaikan tahunan sebesar 5,81% selama periode 2022 hingga 2027 [1].



Gambar 1.1 Pendapatan Produk *Skincare* di Indonesia

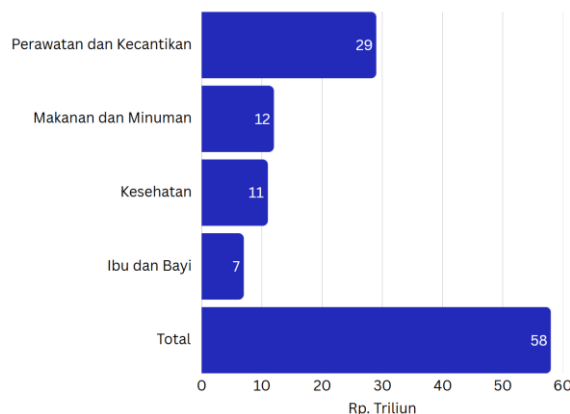
(Sumber: databoks.katadata.co.id) [1]

Peningkatan yang ditunjukkan pada Gambar 1.1 sejalan dengan hasil survei yang dilakukan oleh Katadata Insight Center (KIC) bekerja sama dengan Sirclo pada tahun 2021. Survei tersebut menunjukkan bahwa selama masa pandemi Covid-19, terjadi lonjakan signifikan dalam jumlah transaksi produk-produk kesehatan dan kecantikan. Persentasenya meningkat tajam menjadi 40,1% dibandingkan kondisi sebelum pandemi, yaitu pada tahun 2019 yang hanya sebesar 29,1%. [2].

Jumlah Transaksi Produk Saat Pandemi		
Nama Data	2019	2020-2021
Kesehatan dan Kecantikan	29%	40%
FCMG	31%	31%
Ibu dan Anak	16%	13%
Fashion dan Hobi	16%	10%
Elektronik	7%	3%
Peralatan Rumah Tangga	1%	3%
Lainnya	1%	1%

Gambar 1.2 Jumlah Transaksi Produk Saat Pandemi

Peningkatan saat pandemi melampaui pembelian konsumen terhadap FMCG (Fast Moving Consumer Goods atau produk kebutuhan sehari-hari) yang hanya sebesar 31,2% di tahun 2020-2021 [3].

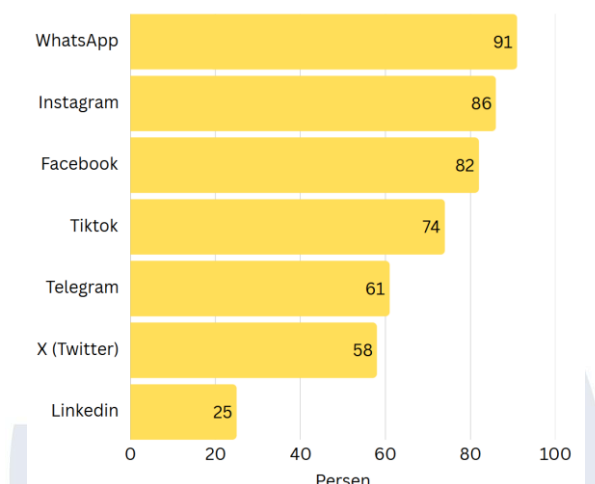


Gambar 1.3 Penjualan Sektor FMCG di E-Commerce Indonesia 2023

Berdasarkan empat kategori utama yang dianalisis, kategori perawatan dan kecantikan menempati posisi teratas dengan nilai penjualan sebanyak 28,2 triliun atau 49% dari total penjualan FMCG di Indonesia tahun 2023 [4].

Hal ini semakin diperkuat dengan hadirnya berbagai merek *skincare* internasional berbahan alami, salah satunya adalah D'Alba yang merupakan *skincare* berbahan asal Italia yang sangat populer di Korea Selatan karena penggunaan *white truffle* sebagai bahan utama dalam produknya [5]. Produk D'Alba salah satunya adalah White Truffle First Spray Serum telah viral di berbagai

platform media sosial. Ulasan positif dari para pengguna dan *beauty influencer* semakin meningkatkan popularitas dari brand ini, sehingga menjadi salah satu *skincare* yang sering dibicarakan.



Gambar 1.4 Sosial Media Ter-Populer 2024

Sejalan dengan peningkatan tersebut, berdasarkan data, Instagram menempati posisi kedua sebagai platform media sosial paling populer di Indonesia dengan nilai 85,3% dari seluruh pengguna internet di Indonesia yang berusia mayoritas 16 sampai 64 tahun [6]. Posisi pertama ditempati oleh WhatsApp dengan nilai 90,9% disusul oleh Facebook dan Tiktok, tingginya persentase penggunaan Instagram menandakan bahwa media sosial ini menjadi salah satu yang potensial dalam mengamati perilaku konsumen.

Instagram merupakan aplikasi media sosial yang memungkinkan pengguna untuk menyukai dan berkomentar pada unggahan. Aplikasi ini termasuk salah satu media sosial yang mengalami pertumbuhan pesat di tahun 2024 [7], sehingga menjadi salah satu platform yang sering digunakan oleh para pembeli untuk berbagi ulasan dan pengalaman. D'Alba tersebar di berbagai cabang negara, termasuk di Indonesia yang memiliki akun Instagram resmi bernama @dalba_indonesia.

Media sosial telah menjadi ruang utama bagi konsumen untuk menyampaikan opini, pengalaman, dan ulasan terhadap suatu produk secara terbuka dan real-time. *Marketplace* di sisi lain, menjadi kanal utama transaksi yang merepresentasikan keputusan akhir konsumen yaitu pembelian [8]. Oleh karena itu, menghubungkan

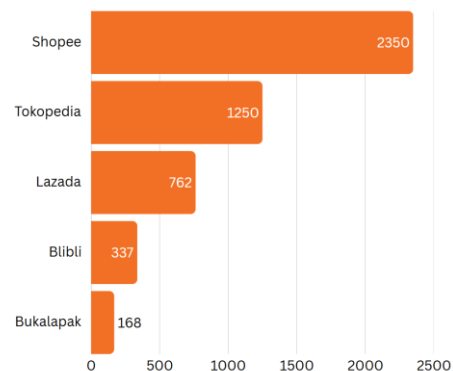
sentimen dari media sosial dengan data penjualan di marketplace menjadi relevan untuk menguji apakah opini publik memiliki pengaruh nyata terhadap performa bisnis.

Keterkaitan ini penting karena dapat membantu perusahaan memahami sejauh mana persepsi konsumen di media sosial mencerminkan perilaku pembelian mereka. Jika ditemukan korelasi yang kuat, maka analisis sentimen dapat dimanfaatkan sebagai indikator untuk memprediksi tren pasar dan mengarahkan strategi pemasaran yang lebih tepat sasaran. Sebaliknya, jika korelasi rendah, hal ini juga menjadi temuan penting bahwa ada faktor lain yang lebih dominan dalam mendorong penjualan.

Analisis sentimen di media sosial memiliki berbagai tantangan cukup kompleks, karakteristik data yang berasal dari Instagram dapat cenderung pendek, bahasa informal, campuran bahasa, dan menggunakan *emoticon*. Kegunaan analisis sentimen sendiri agar dapat diketahui pandangan seseorang cenderung ke arah opini yang positif, negatif, atau netral [9].

Dalam penelitian ini, data diambil berasal dari akun Instagram konsumen yang membuat ulasan produk D'Alba untuk mengidentifikasi sentimen pengguna terhadap produk D'Alba di pasar Indonesia. Meskipun fokus analisis sentimen dilakukan terhadap komentar di Instagram, data dari *marketplace* turut digunakan sebagai data tambahan dalam menganalisis performa penjualan produk D'Alba.

Data *marketplace* berperan dalam memberikan gambaran objektif mengenai performa penjualan produk, yang dalam konteks penelitian ini digunakan sebagai indikator untuk menguji apakah opini dan persepsi konsumen di media sosial memiliki pengaruh terhadap keputusan pembelian. Dengan demikian, data penjualan dari *marketplace* tidak menjadi variabel utama dalam analisis, tetapi menjadi acuan yang memperkuat interpretasi terhadap hasil sentimen yang ditemukan.



Gambar 1.5 E-Commerce dengan Pengunjung Terbanyak

Pada tahun 2023, Shopee dan Tokopedia menjadi dua urutan *e-commerce* tertinggi yang memiliki pengunjung terbanyak [10] sehingga pada penelitian ini digunakan data penjualan dari kedua *e-commerce* tersebut. Penelitian ini menggunakan dua algoritma dalam analisis sentimen yaitu SVM dan Naïve Bayes.

Pemilihan algoritma Support Vector Machine (SVM) dan Naïve Bayes dalam penelitian ini didasarkan pada karakteristik data serta efektivitas kedua algoritma dalam klasifikasi teks, khususnya pada domain analisis sentimen. Naïve Bayes dipilih karena memiliki kemampuan klasifikasi yang cepat dan efisien untuk teks pendek seperti komentar di media sosial. Sifat probabilistiknya memungkinkan algoritma ini bekerja dengan baik pada data berukuran besar dan sederhana, serta tetap memberikan hasil yang cukup akurat meskipun fitur-fitur bersifat independen [11].

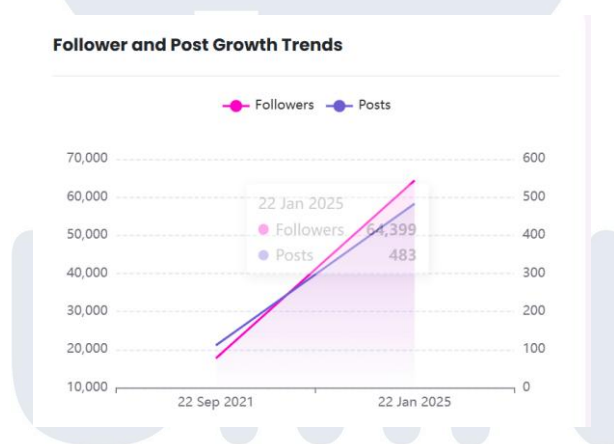
Sementara itu, algoritma SVM dipilih karena kemampuannya dalam menangani data dengan pola kompleks dan berdimensi tinggi, seperti komentar di Instagram yang mengandung variasi bahasa, emotikon, serta konteks yang tidak eksplisit. SVM dikenal sebagai algoritma yang tangguh untuk tugas klasifikasi teks karena dapat menemukan hyperplane optimal yang memisahkan kelas-kelas secara jelas, bahkan pada data yang tidak linier [12].

Meskipun kedua algoritma ini telah banyak digunakan dalam penelitian sebelumnya, namun hingga saat ini belum ada kesimpulan mutlak mengenai algoritma mana yang paling unggul secara konsisten. Setiap algoritma memiliki kelebihan dan keterbatasan, tergantung pada struktur data, metode *preprocessing*,

serta konteks aplikasi yang digunakan. Oleh karena itu, perbandingan performa antara SVM dan Naïve Bayes tetap menjadi topik yang relevan dan penting untuk diteliti, khususnya dalam konteks domain yang spesifik seperti ulasan skincare di media sosial dan keterkaitannya dengan penjualan di e-commerce.

Dengan membandingkan kedua algoritma ini, penelitian ini tidak hanya fokus pada hasil klasifikasi, tetapi juga bertujuan untuk melihat performa dan efektivitas masing-masing metode dalam menginterpretasikan opini konsumen di media sosial terhadap suatu produk, serta mengaitkannya dengan indikator bisnis berupa penjualan. Kombinasi dua pendekatan ini memberikan gambaran yang lebih utuh dan objektif dalam mengevaluasi sentimen publik terhadap produk D'Alba.

D'Alba sendiri dipilih untuk dijadikan bahan penelitian karena popularitasnya yang sedang melunjak dan sering diperbincangkan di media sosial, dilihat pada media sosial Instagramnya yang memiliki pengikut sebanyak 64ribu lebih pengguna dan mulai meningkat sejak tahun 2021 [13].



Gambar 1.6 Insight Instagram D'Alba

(Sumber: starnagage.com) [11]

Selain itu, produk D'Alba menunjukkan tingkat kepuasan pengguna yang tinggi sebagaimana terlihat pada Gambar 1.6, produk mereka memperoleh rating rata-rata 4,3 dari 5 bintang berdasarkan 508 ulasan pengguna. Sebesar 96% pengguna merekomendasikan produk D'Alba kepada orang lain, data ini memperkuat urgensi dalam meneliti lebih lanjut dominan sentimen apa yang akan muncul di media sosial Instagram.

Dengan menganalisis sentimen pengguna terhadap produk D'Alba di media sosial khususnya Instagram diharapkan hasil penelitian ini dapat memberikan wawasan bagi konsumen dan perusahaan mengenai opini publik di media sosial dapat mempengaruhi keputusan pembelian di *e-commerce* khususnya Shopee dan Tokopedia.

1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah yang bisa diambil adalah sebagai berikut:

1. Bagaimana sentimen pengguna Instagram terhadap produk skincare D'Alba?
2. Bagaimana perbandingan performa algoritma SVM dan Naïve Bayes dalam mengklasifikasikan sentimen komentar pengguna Instagram terhadap produk D'Alba?
3. Apakah terdapat korelasi antara sentimen komentar pengguna Instagram terhadap produk D'Alba dengan jumlah penjualan di Shopee dan Tokopedia?

1.3 Batasan Masalah

Agar penelitian lebih terarah dan fokus diperlukan batasan masalah yang ditetapkan untuk mencapai target analisis. Berikut batasan masalah yang ada:

1. Data komentar yang dianalisis berasal dari media sosial Instagram, khususnya pada unggahan terkait produk D'Alba dalam periode Januari 2021 hingga Januari 2025. Komentar dikumpulkan melalui teknik *web scraping exportcomments*, sehingga hanya komentar publik yang dapat diakses yang digunakan dalam analisis.
2. Data dibagi menjadi dua bagian: data latih sebesar 80% berasal dari periode Januari 2021 hingga Desember 2023, digunakan untuk melatih model. Sementara data uji sebesar 20% berasal dari periode Januari 2024 hingga Januari 2025, digunakan untuk mengevaluasi performa model.
3. Terdapat keterbatasan dalam proses pengumpulan data akibat batasan teknis *scraping tools*, seperti adanya *limit* dan tidak dapat mengakses komentar dari akun private, komentar pada *reels*, maupun *story*.

4. Penjualan produk D'Alba yang digunakan sebagai data tambahan diperoleh dari dua *marketplace* populer, yaitu Shopee dan Tokopedia. Data penjualan tidak mencakup faktor promosi, diskon, waktu transaksi yang presisi, ataupun identitas konsumen.
5. Platform *marketplace* yang dianalisis untuk data penjualan adalah Shopee dan Tokopedia. Data dikumpulkan dengan metode *scraping* berdasarkan kata kunci “*skincare D'Alba*”.
6. Penelitian ini hanya menganalisis komentar berbahasa Indonesia dan Inggris, komentar yang menggunakan bahasa selain dua bahasa tersebut tidak akan disertakan dalam proses klasifikasi.
7. Algoritma yang digunakan dalam menganalisis sentimen adalah Support Vector Machine (SVM) dan Naïve Bayes. Kedua algoritma tersebut digunakan untuk membandingkan performa klasifikasi sentimen.
8. Kategori sentimen dibagi menjadi tiga kelas: positif, negatif, dan netral. Penentuan kategori dilakukan dengan pendekatan leksikal berbasis kamus sentimen. Sehingga memiliki keterbatasan dalam menangkap konteks kalimat kompleks atau sarkasme.

1.4 Tujuan dan Manfaat Penelitian

Berdasarkan latar belakang dan rumusan masalah, maka dapat disimpulkan tujuan dan manfaat penelitian sebagai berikut:

1.4.1 Tujuan Penelitian

1. Menganalisis sentimen komentar pengguna Instagram terhadap produk *skincare D'Alba*.
2. Membandingkan performa algoritma SVM dan Naïve Bayes dalam mengklasifikasikan sentimen pengguna Instagram terhadap produk D'Alba berdasarkan metrik akurasi, presisi, *recall* dan *f1-score*.
3. Menganalisis korelasi antara hasil klasifikasi sentimen komentar pengguna Instagram dengan jumlah penjualan produk D'Alba di Shopee dan Tokopedia.

1.4.2 Manfaat Penelitian

1. Memberikan masukan bagi D'Alba dalam memahami persepsi konsumen di Indonesia sebagai dasar pengembangan produk.
2. Menambah referensi akademik terkait penerapan algoritma SVM dan Naïve Bayes dalam analisis sentimen.
3. Memberikan informasi bagi konsumen mengenai opini publik terhadap produk *skincare* D'Alba sebagai bahan pertimbangan dalam mengambil keputusan pembelian.

1.5 Sistematika Penulisan

BAB 1 PENDAHULUAN

Pada Bab 1 Pendahuluan terdapat latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan yang terdapat dalam penelitian ini.

BAB II LANDASAN TEORI

Pada Bab 2 Landasan Teori terdapat teori tentang penelitian yang berisikan kajian literatur pada penelitian terdahulu sebagai dasar pemahaman yang mendukung penelitian. Selain itu sub bab yang ada pada bab ini menyajikan konsep dan teori yang relevan terhadap penelitian.

BAB III METODOLOGI PENELITIAN

Pada Bab 3 Metodologi menjelaskan setiap langkah metodologi penelitian yang akan digunakan, seperti sumber data yang digunakan, dan alur penelitian yang telah dilakukan.

BAB IV HASIL DAN PEMBAHASAN

Pada Bab 4 Hasil dan Pembahasan menjelaskan hasil penelitian yang diperoleh berdasarkan langkah-langkah metodologi yang telah dilakukan sebelumnya meliputi hasil analisis sentimen dan hasil perbandingan kedua algoritma serta pembahasannya.

BAB V SIMPULAN DAN SARAN

Pada Bab 5 Simpulan dan Saran berisikan simpulan dari penelitian dan saran yang dapat diberikan berdasarkan dari hasil penelitian.