

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Industri kecantikan, terutama segmen perawatan kulit, mengalami pertumbuhan yang sangat pesat, karena kesadaran masyarakat akan kesehatan kulit semakin meningkat. Hal tersebut menjadikan produsen untuk terus berinovasi dan menawarkan berbagai klaim manfaat dan keunggulan produk. Namun, fenomena *overclaim*, yaitu klaim berlebihan yang tidak didukung dengan bukti ilmiah yang valid, semakin menjadi sorotan. *Overclaim* mengacu pada klaim berlebihan yang dibuat oleh produsen mengenai manfaat produk, yang sering kali tidak didukung oleh bukti ilmiah yang valid. Klaim ini memberikan janji tentang sesuatu yang tidak realistis dan dapat memberikan ekspektasi terhadap konsumen serta menyesatkan mereka [1].

Praktik *overclaim* ini mengakibatkan konsumen kehilangan kepercayaan terhadap merek dan industri kecantikan secara keseluruhan [1]. Konsumen yang mengandalkan produk dengan klaim yang tinggi bisa mengalami kerugian secara finansial dan juga emosional karena ekspektasi yang tidak nyata adanya. Misalnya, produk yang mengklaim mengandung 10% vitamin C, padahal hasil uji lab hanya menunjukkan 5% [2].

Pada tanggal 30 September 2024, Kepala BPOM mengumumkan akan mencabut izin edar produk *skincare* lokal yang terbukti melakukan *overclaim* [1]. BPOM menegaskan bahwa klaim yang tidak sesuai dengan kandungan atau komposisi yang tertulis di label produk akan dikenakan sanksi, termasuk peringatan dan pemanggilan secara resmi kepada pengusaha terkait. Pada tahun 2024 BPOM mendata, ada 398.603 izin edar kosmetik yang telah teregistrasi [3].

Kemudian munculnya fenomena ulasan "dokter detektif" yang membongkar klaim *skincare* yang berlebihan telah memicu perdebatan pada konsumen dan pelaku industri kecantikan [4]. Yang membuat pendapat masyarakat mengenai produk

Kecantikan tersebar luas di media sosial dan *e-commerce*. "Dokter Detektif" atau "Doktif," yaitu seorang dokter *anonim* yang sering aktif di TikTok, Doktif ini membuat konten viral yang mengungkap ketidaksesuaian antara klaim kandungan bahan pada berbagai *skincare* populer dan hasil uji lab yang didukung oleh *SIG Laboratory* [2].

Konten-konten yang dibuat oleh "Dokter Detektif" mengungkapkan bahwa beberapa produk kecantikan tidak memenuhi klaim yang tertulis pada label, terutama mengenai persentase bahan aktif. Hal tersebut membuat kesadaran konsumen akan pentingnya transparansi dalam industri kecantikan meningkat. Akun @dokterdetektif di TikTok menjadi unik karena memberikan edukasi yang diperkuat dengan hasil lab, sehingga meningkatkan tingkat kepercayaan audiens dan membantu para konsumen untuk lebih berhati-hati dalam memilih *skincare* untuk meminimalkan risiko penggunaan produk yang tidak aman [4].

Fenomena ini sangat berdampak pada tingkat kepercayaan konsumen. Sebuah laporan dari *AgilityPR* menunjukkan bahwa merek yang menghadapi isu transparansi berpotensi mengalami penurunan reputasi [2]. Konsumen cenderung beralih ke produk yang lebih transparan dan jujur dalam klaimnya. Hal ini juga mendorong konsumen untuk lebih berhati-hati dalam memilih produk *skincare*. Survei Populix menunjukkan bahwa 68% konsumen mengaku lebih berhati-hati dalam membeli produk *skincare* setelah menonton ulasan "dokter detektif" di media sosial.

Namun, ada beberapa perbedaan pendapat di kalangan masyarakat mengenai ulasan kritis ini. Ada sebagian yang mendukung edukasi ini karena dapat memberikan wawasan tentang bahan dan kegunaan produk tersebut, sementara yang lain mengkritik cara penyampaiannya yang dinilai dapat menimbulkan ketakutan secara berlebihan terhadap produk *skincare*.

Analisis sentimen menjadi metode yang efektif untuk memahami persepsi masyarakat terhadap isu *overclaim* dalam industri kecantikan. Melalui analisis sentimen, opini masyarakat dapat diklasifikasikan ke dalam kategori positif dan negatif berdasarkan data yang diperoleh dari komentar pada utas (*thread*) dari *platform X*

(sebelumnya dikenal sebagai Twitter). Saat ini, konsumen cenderung menyampaikan pendapatnya melalui *platform online* seperti media sosial dan *e-commerce* [5].

Salah satu pendekatan yang dapat digunakan adalah *Bidirectional Encoder Representations from Transformers* (BERT), sebuah model *deep learning* yang telah terbukti memiliki akurasi tinggi dalam pemrosesan bahasa alami (*Natural Language Processing/NLP*). Algoritma BERT memiliki kemampuan untuk memahami konteks kata secara lebih mendalam dibandingkan metode NLP sebelumnya, karena memproses teks secara *bidirectional*, bukan hanya dari kiri ke kanan atau sebaliknya. Berdasarkan penelitian yang dilakukan oleh P. Akter dkk., penggunaan BERT terbukti efektif dalam menganalisis sentimen konsumen serta memberikan wawasan yang dapat dimanfaatkan dalam pengambilan keputusan bisnis. Oleh karena itu, penelitian ini juga menggunakan model BERT, khususnya varian IndoBERT, untuk mengklasifikasikan sentimen masyarakat terhadap isu *overclaim* pada produk kecantikan [6].

Berbagai varian BERT telah dikembangkan untuk menyesuaikan kebutuhan spesifik bahasa dan domain. Beberapa di antaranya seperti *Multilingual BERT* (mBERT), DistilBERT, dan IndoBERT. *Multilingual BERT* dilatih dengan lebih dari 100 bahasa, sehingga cocok untuk tugas lintas bahasa, namun kurang optimal ketika digunakan untuk satu bahasa tertentu. Di sisi lain, IndoBERT merupakan varian BERT yang dilatih secara khusus dengan korpus bahasa Indonesia, sehingga memiliki keunggulan dalam memahami konteks lokal dan struktur bahasa yang khas di Indonesia.

Beberapa penelitian terdahulu menunjukkan bahwa algoritma BERT mampu memberikan hasil yang lebih akurat dibandingkan metode klasik dalam analisis sentimen. Penelitian ini menggunakan pengumpulan dataset utas dari X dengan kata kunci dokter detektif. Salah satu fokus utama penelitian ini adalah membedakan perbedaan sentimen antara akun X yang memiliki centang biru dan yang tidak memiliki centang biru. Dengan demikian, penelitian ini dapat mengidentifikasi apakah akun yang terverifikasi memiliki pola opini yang berbeda dibandingkan akun biasa.

Melalui analisis sentimen pelanggan, perusahaan dapat memperoleh wawasan mendalam tentang preferensi dan persepsi pelanggan terhadap produk. Hasil analisis tersebut dapat digunakan untuk menyempurnakan pilihan produk yang ditawarkan oleh *e-commerce*, dengan memprioritaskan atau mengoptimalkan stok produk yang mendapatkan tanggapan positif [5].

Analisis sentimen berfokus pada opini-opini yang mengekspresikan atau mengungkapkan sentimen positif atau negatif. Tujuan dari analisis sentimen yaitu untuk memetakan opini pengguna berdasarkan topik yang telah ditentukan, serta untuk mendapatkan informasi yang lebih akurat dan detail dari ulasan pengguna dari berbagai perspektif dan aspek.

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap isu *overclaim* pada produk kecantikan menggunakan algoritma BERT. Penelitian ini akan menggunakan pendekatan kuantitatif dengan mengumpulkan dan menganalisis setidaknya 3.000 ulasan dari *platform* media sosial utas dari aplikasi X untuk memastikan representasi data yang luas dan akurat. Dengan pendekatan kuantitatif, penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai persepsi publik terhadap fenomena ini. Hasil penelitian ini juga dapat menjadi bahan pertimbangan bagi pelaku usaha *skincare* dalam menyusun strategi pemasaran yang lebih transparan, bagi regulator dalam menetapkan kebijakan yang lebih ketat, serta bagi konsumen dalam mengambil keputusan yang lebih bijak terkait produk kecantikan yang mereka gunakan.

1.2 Rumusan Masalah

Mengacu pada konteks yang telah disebutkan, permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana menerapkan metode *Deep Learning* dengan algoritma BERT (*Bidirectional Encoder Representations from Transformers*) untuk program analisis sentimen terhadap isu *overclaim* pada produk kecantikan?

2. Bagaimana kinerja algoritma BERT dalam mengklasifikasikan sentimen (positif dan negatif) terhadap isu *overclaim* pada produk kecantikan?

1.3 Batasan Masalah

Untuk memastikan riset ini tetap fokus dan sesuai dengan tujuan yang telah ditetapkan, berikut beberapa batasan masalah dalam Program Analisis Sentimen terhadap Isu *Overclaim* pada Produk Kecantikan:

1. Sumber data yang digunakan berasal dari komentar di media sosial seperti X (Twitter) dengan berbahasa Indonesia.
2. Dataset yang digunakan dalam penelitian ini mencakup minimal 3.000 komentar untuk memperoleh hasil yang lebih representatif.
3. Model analisis sentimen yang diterapkan menggunakan algoritma BERT (*Bidirectional Encoder Representations from Transformers*) dan tidak membandingkan dengan algoritma lain.
4. Penelitian ini hanya berfokus pada analisis teks dan tidak mempertimbangkan elemen visual atau *non-verbal* seperti emoji, gambar, dan video.

1.4 Tujuan Penelitian

Tujuan dari pembangunan Program Analisis Sentimen terhadap Isu *Overclaim* pada Produk Kecantikan menggunakan Metode *Deep Learning* yaitu:

1. Menerapkan metode *Deep Learning* dengan algoritma BERT (*Bidirectional Encoder Representations from Transformers*) pada program analisis sentimen terhadap isu *overclaim* pada produk kecantikan.
2. Mengetahui kinerja algoritma BERT dalam mengklasifikasikan sentimen (positif dan negatif) terhadap isu *overclaim* pada produk kecantikan.

1.5 Manfaat Penelitian

Dengan dilakukan nya penelitian ini, peneliti dapat menguraikan manfaat-manfaat penelitian sebagai berikut:

1.5.1 Bagi Peneliti

Peneliti dapat menerapkan ilmu-ilmu yang telah dipelajari selama perkuliahan terutama penerapan konsep NLP dengan menggunakan model BERT pada klasifikasi sentimen.

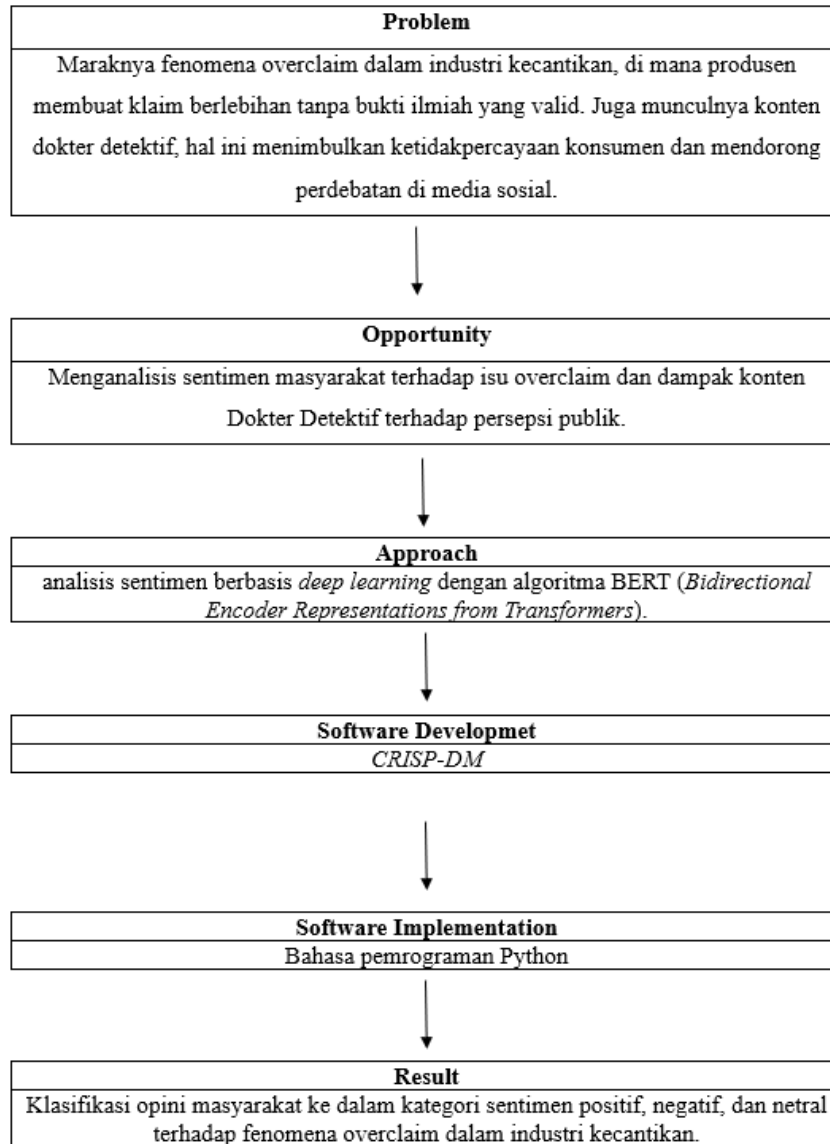
1.5.2 Bagi Universitas

- a. Menilai kemampuan mahasiswa dalam memahami dan mengimplementasikan materi pembelajaran yang telah diperoleh selama perkuliahan.
- b. Sebagai karya ilmiah yang hak ciptanya dimiliki oleh universitas serta dapat dijadikan sebagai referensi bagi penelitian selanjutnya.
- c. Memberikan gambaran mengenai kesiapan mahasiswa dalam menghadapi dunia kerja.



1.6 Kerangka Pemikiran

Kerangka pemikiran pada penelitian tugas akhir ini dapat dilihat pada gambar 1.



Gambar 1 Kerangka pemikiran

1.7 Sistematika Penulisan

Dalam menyusun penelitian ini, peneliti menggunakan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, kerangka pemikiran, metode penelitian, dan sistematika penulisan. Bagian ini bertujuan untuk memberikan gambaran awal mengenai pentingnya penelitian terhadap isu *overclaim* pada produk kecantikan dan bagaimana pendekatan analisis sentimen dengan algoritma BERT akan diterapkan.

BAB II KAJIAN LITERATUR

Bab ini menguraikan teori-teori yang mendasari penelitian, termasuk konsep isu *overclaim* pada produk kecantikan, analisis sentimen, media sosial X (Twitter), *natural language processing* (NLP), serta penjelasan tentang algoritma BERT. Selain itu, disajikan juga tinjauan pustaka dari penelitian-penelitian terdahulu yang relevan sebagai referensi dan perbandingan.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan-tahapan metodologi yang digunakan dalam penelitian, mengacu pada pendekatan CRISP-DM. Penjelasan mencakup tahap *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, serta bahan dan alat yang digunakan dalam proses penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi analisis sentimen menggunakan model BERT terhadap data komentar dari media sosial X (Twitter) terkait isu *overclaim*. Selain itu, dijelaskan juga hasil evaluasi model dengan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Pembahasan dilakukan secara mendalam untuk menginterpretasikan temuan dan hubungannya dengan tujuan penelitian.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian dan saran yang dapat diberikan kepada pihak terkait, seperti pelaku industri kecantikan, regulator, serta peneliti selanjutnya. Kesimpulan ditarik berdasarkan temuan analisis sentimen, sementara saran diarahkan pada pengembangan sistem serupa atau penelitian lanjutan.