

BAB I

PENDAHULUAN

1.1 Latar Belakang

Industri kecantikan global saat ini menunjukkan perkembangan yang sangat pesat [1]. Tren positif ini juga tercermin di Indonesia, yang mengalami pertumbuhan konsisten dalam beberapa tahun terakhir dan menjadikannya salah satu pasar kosmetik dengan perkembangan tercepat di kawasan Asia Tenggara. Pendapatan pasar beauty and personal care di Indonesia mencapai sekitar US\$7,3 miliar pada tahun 2022 [2]. Secara lebih spesifik, data Statista Market Insights menunjukkan bahwa pendapatan segmen kosmetik di Indonesia diperkirakan mencapai US\$2,09 miliar pada tahun 2025 dengan Compound Annual Growth Rate (CAGR) sebesar 4,73% untuk periode 2025–2030 [3]. Pertumbuhan ini didorong oleh meningkatnya kesadaran masyarakat terhadap perawatan diri, perkembangan e-commerce, serta pengaruh media sosial yang semakin mempermudah konsumen dalam memperoleh informasi dan rekomendasi produk. Kondisi tersebut membuka peluang yang besar bagi merek kecantikan lokal untuk berkembang dan bersaing dengan produk internasional.

Salah satu merek lokal yang tumbuh di tengah perkembangan industri tersebut adalah Mother of Pearl (MOP), yang diluncurkan pada tahun 2021 oleh *beauty influencer* Tasya Farasya [4]. MOP dikenal sebagai salah satu merek kosmetik lokal yang lahir dari kekuatan media sosial dan berhasil menarik perhatian publik dalam waktu singkat [5]. Produk ini menjadi salah satu produk makeup yang digemari berbagai kalangan, khususnya milenial dan generasi Z [6]. Keberhasilan peluncurannya juga tercermin dari pencapaian penjualan sebesar Rp1 miliar dalam 15 menit pertama melalui platform *e-commerce* [7]. Popularitas tersebut menunjukkan bahwa MOP berhasil memanfaatkan kekuatan *influencer marketing*, *viral marketing*, dan *Electronic Word of Mouth (eWOM)* untuk membangun awareness dan kepercayaan konsumen [8]. Namun, ulasan dan opini jujur dari

pengguna lain kini menjadi penilaian utama yang sangat memengaruhi laku atau tidaknya sebuah merek kecantikan di tengah persaingan pasar yang ketat [9].

Dalam memilih produk kosmetik, konsumen produk kecantikan cenderung lebih teliti dan kritis terhadap detail kualitas produk [10]. Berbagai keluhan mengenai masalah teknis, seperti kemasan yang kurang berfungsi dengan baik atau warna produk yang menggelap saat digunakan pada kulit, sering didiskusikan secara terbuka oleh konsumen [11]. Pembahasan jujur mengenai kualitas produk MOP ini paling banyak ditemukan di platform X. Karakteristik platform tersebut memungkinkan penggunanya untuk menyampaikan opini secara bebas, cepat, dan apa adanya [12]. Selain itu, pengguna platform X sering kali menggunakan bahasa tidak baku atau percampuran bahasa Indonesia dan Inggris dalam menuliskan ulasannya, yang pada akhirnya menambah kompleksitas dalam memahami teks ulasan tersebut secara otomatis.



Gambar 1. 1 Contoh Ulasan Pengguna MOP Pertama [13]

Sebagai bukti nyata di lapangan, Gambar 1.1 menunjukkan ulasan konsumen yang mengandung opini beragam terhadap beberapa produk MOP sekaligus. Pengguna memberikan sentimen positif pada aspek tekstur dan efek menyamarkan pori atau *pore blurring* dari produk bedak tabur, serta hasil yang natural pada produk *browcara*. Namun, dalam cuitan yang sama, pengguna juga memberikan sentimen negatif pada aspek performa bedak yang kurang maksimal dalam menahan minyak, serta menilai produk pewarna bibir biasa saja.

Sementara itu, Gambar 1.2 memperlihatkan ulasan terkait produk *foundation* dan *lip tint*. Pengguna menyampaikan sentimen positif pada formula *foundation* yang minim oksidasi untuk kulit berminyak, tetapi secara bersamaan mengeluhkan aspek *packaging* yang dinilai buruk karena mudah mengelupas. Kedua contoh

cuitan ini membuktikan bahwa dalam satu ulasan, konsumen dapat memberikan sentimen yang berbeda-beda untuk setiap aspek produk yang spesifik.



Gambar 1. 2 Contoh Ulasan Pengguna MOP Kedua[14]

Berdasarkan contoh ulasan tersebut, terlihat jelas bahwa opini konsumen sering kali campur aduk. Selama ini, opini masyarakat di media sosial umumnya hanya dikelompokkan ke dalam tiga kategori sentimen umum: positif, negatif, atau netral [15]. Namun, metode analisis sentimen konvensional yang hanya menilai teks secara keseluruhan terbukti tidak cukup akurat untuk membaca ulasan yang kompleks [16]. Seperti yang terjadi pada produk MOP, sebuah teks ulasan dapat memuji kualitas formula produk sekaligus mengeluhkan kemasan di dalam satu kalimat yang sama [17]. Analisis sentimen biasa sering kali gagal menangkap perbedaan opini ini dan justru berisiko memberikan satu label sentimen yang salah. Jika hal ini dibiarkan, informasi penting mengenai kelebihan dan kekurangan spesifik dari produk tersebut akan hilang. Oleh karena itu, diperlukan sebuah metode analisis yang mampu memecah dan menilai ulasan konsumen secara lebih rinci ke dalam aspek-aspek tertentu [18].

Untuk menjawab kebutuhan analisis yang detail tersebut, pendekatan komputasional tingkat lanjut di bidang *Natural Language Processing* (NLP) menjadi solusi yang ideal [19]. Dalam permasalahan ini, *Aspect-Based Sentiment Analysis* (ABSA) dibutuhkan sebagai kerangka kerja yang tidak hanya menentukan polaritas sentimen, tetapi juga mengidentifikasi target aspek yang memengaruhinya [20]. Melalui penerapan metode ABSA, opini konsumen yang awalnya hanya berupa teks ulasan mentah dapat diolah menjadi data yang lebih terukur. Metode ini mampu menilai tingkat kepuasan pelanggan terhadap aspek-aspek spesifik secara terpisah, seperti dari segi harga, kemasan, maupun formula produk [21].

Penelitian terkait ABSA telah banyak dikembangkan, namun sering kali terbentur pada keterbatasan data berlabel yang memadai, khususnya pada ulasan produk spesifik di media sosial. Untuk mengatasi keterbatasan data tersebut, pendekatan *cross-domain* mulai diimplementasikan. Penelitian oleh [22] mengusulkan strategi domain-adaptive pre-training yang dilanjutkan dengan fine-tuning untuk meningkatkan performa model pada domain dengan sumber daya terbatas. Pendekatan lain dikembangkan oleh [23] melalui Domain Adversarial Training untuk mengurangi perbedaan distribusi antar domain tanpa memerlukan data berlabel pada domain target. Selain itu, penelitian oleh [24] menggabungkan transformer dengan BiLSTM dan prosedur *attention* untuk meningkatkan pemodelan konteks. Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada skenario *single-label classification*, sehingga belum mampu merepresentasikan ulasan yang mengandung lebih dari satu aspek dengan sentimen yang berbeda dalam satu teks.

Pendekatan multi-label ABSA kemudian mulai dikembangkan untuk menangani kompleksitas sentimen ganda tersebut. Penelitian oleh [25] menunjukkan bahwa model berbasis transformer mampu mengklasifikasikan beberapa pasangan aspek dan sentimen secara bersamaan dalam satu ulasan. Namun, penelitian ini masih terbatas pada satu domain data, sehingga belum mempertimbangkan perbedaan distribusi antar domain. Di sisi lain, data media sosial yang sarat dengan *code-mixing* juga menjadi tantangan tersendiri bagi pemodelan. Penelitian oleh [26] membuktikan bahwa pemilihan *pre-trained language* model dapat sangat berpengaruh terhadap performa pada teks *code-mixed*, meskipun hal ini belum secara khusus diterapkan pada tugas ABSA lintas domain.

Berdasarkan tinjauan tersebut, terlihat jelas adanya celah penelitian. Pendekatan *cross-domain* ABSA saat ini umumnya masih terbatas pada skenario *single-label*, sementara pendekatan *multi-label* ABSA belum mempertimbangkan perbedaan distribusi antar domain maupun menangani karakteristik teks *code-mixed*. Oleh karena itu, penelitian ini bertujuan untuk mengusulkan arsitektur *Cross-Domain*

Multi-Label Aspect-Based Sentiment Analysis (CD-ML-ABSA) berbasis transformer untuk menganalisis ulasan produk kecantikan pada media sosial. Arsitektur yang diusulkan mengintegrasikan *domain-adaptive pre-training* untuk meningkatkan representasi bahasa pada domain spesifik, serta metode klasifikasi *multi-label* untuk memprediksi sentimen pada berbagai aspek secara bersamaan dalam satu teks. Hal ini diharapkan mampu menghasilkan analisis sentimen yang komprehensif dan akurat sesuai dengan bahasa alami konsumen.

Pemilihan MOP sebagai objek penelitian sangat relevan dengan tujuan pemodelan CD-ML-ABSA. Sebagai merek kecantikan lokal yang viral dan memiliki tingkat interaksi konsumen yang masif di platform X, ulasan produk MOP memiliki karakteristik sentimen yang sangat kompleks. Konsumen MOP cenderung kritis, sering kali mengekspresikan sentimen yang berlawanan terhadap beberapa aspek sekaligus di dalam satu kalimat, dan menuliskannya menggunakan gaya bahasa *code-mixing*. Karakteristik ulasan MOP inilah yang menjadikannya objek ideal untuk menguji keandalan arsitektur CD-ML-ABSA dalam menangkap multiaspek dan multipolaritas sentimen secara bersamaan di tengah keterbatasan data berlabel.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang ada, maka penelitian ini dirumuskan dalam beberapa masalah, seperti:

1. Bagaimana hasil mengembangkan arsitektur CD-ML-ABSA berbasis *transformer* yang mampu mengidentifikasi berbagai aspek dan polaritas sentimen secara bersamaan dalam satu ulasan produk kecantikan?
2. Bagaimana hasil penerapan *domain-adaptive pre-training* dalam meningkatkan representasi bahasa pada data ulasan produk kecantikan yang bersifat *cross-domain* dan mengandung *code-mixed text*?
3. Bagaimana hasil performa model CD-ML-ABSA yang diusulkan dalam mengklasifikasikan sentimen multi-label pada berbagai aspek berdasarkan metrik evaluasi multi-label?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah yang melindungi pembahasan, sehingga pembahasan tidak terlalu luas, seperti:

1. Objek data observasi difokuskan secara eksklusif pada opini publik berbahasa Indonesia di platform X terkait merek kecantikan MOP.
2. Ekstraksi data ditarik berdasarkan pencarian kata kunci spesifik, yaitu: "review produk mother of pearl", "review concealer mother of pearl", "blush MOP ", "cushion MOP ", "mascara MOP ", "Lip MOP ", "microblur MOP ", dan "sunscreen MOP ", tanpa batasan rentang waktu tertentu.
3. Total populasi data mentah yang diekstraksi menggunakan *tweet-harvest* untuk kemudian diproses dalam tahap *preprocessing* awal berjumlah 12.547 cuitan.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Penelitian ini dilaksanakan dengan tujuan komputasional sebagai berikut:

1. Mengembangkan arsitektur CD-ML-ABSA berbasis *transformer* yang mampu mengidentifikasi berbagai aspek dan polaritas sentimen dalam satu ulasan produk kecantikan.
2. Menerapkan pendekatan *domain-adaptive pre-training* untuk meningkatkan kualitas representasi bahasa pada data ulasan produk kecantikan yang bersifat *cross-domain* dan mengandung *code-mixed text*.
3. Mengevaluasi performa model CD-ML-ABSA yang diusulkan dalam mengklasifikasikan sentimen multi-label pada berbagai aspek menggunakan metrik evaluasi multi-label.

1.4.2 Manfaat Penelitian

Adapun manfaat yang diharapkan dari hasil penelitian ini meliputi:

1. Memberikan kontribusi praktis berupa rekomendasi pendekatan pemodelan ABSA berbasis *transformer* yang mampu menangani ulasan produk kecantikan pada media sosial yang bersifat kompleks, termasuk teks yang mengandung *multi-label* serta fenomena *code-mixed text*.

2. Memberikan kontribusi teoritis dalam pengembangan metode NLP, khususnya dalam integrasi *domain-adaptive pre-training* dan klasifikasi *multi-label* dalam kerangka *cross-domain* ABSA, sehingga dapat memperkaya literatur terkait pemodelan sentimen berbasis aspek pada data berbahasa Indonesia.

1.5 Sistematika Penulisan

Sistematika penulisan penelitian ini terbagi dalam beberapa bagian, sebagai berikut:

1. BAB 1 PENDAHULUAN

Pada bagian ini berisi penjelasan mengenai latar belakang penelitian, identifikasi masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan

2. BAB 2 LANDASAN TEORI

Pada bagian ini berisi teori-teori dan konsep yang mendukung penelitian, meliputi kajian penelitian terdahulu, teori yang relevan dengan topik penelitian, metode dan algoritma yang digunakan, serta tools dan perangkat lunak yang mendukung pelaksanaan penelitian.

3. BAB 3 METODOLOGI PENELITIAN

Pada bagian ini berisi gambaran umum penelitian, termasuk pendekatan yang digunakan, sumber data, teknik pengumpulan data, kerangka kerja penelitian, serta tahapan pengolahan dan analisis data yang dilakukan.

4. BAB 4 ANALISIS DAN HASIL PENELITIAN

Pada bagian ini menyajikan hasil implementasi metode dan model yang digunakan, disertai analisis terhadap hasil eksperimen dan evaluasi kinerja model yang diperoleh.

5. BAB 5 SIMPULAN DAN SARAN

Pada bagian ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran yang dapat digunakan sebagai acuan untuk pengembangan penelitian selanjutnya.



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA