

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi digital telah membawa perubahan besar dalam cara bagaimana masyarakat berbelanja. *E-commerce* saat ini menjadi pilihan utama karena dianggap lebih praktis, cepat, dan memudahkan konsumen dalam membandingkan harga dan kualitas produk [1]. Salah satu *platform e-commerce* yang paling banyak digunakan adalah Shopee. Berdasarkan data *iPrice*, Shopee mencatat jumlah kunjungan tertinggi di antara *platform e-commerce* lainnya, dengan jumlah kunjungan mencapai 96 juta pada Q3 dan meningkat signifikan menjadi 129 juta pada Q4 tahun 2020, serta menempati peringkat pertama sebagai *platform e-commerce* dengan jumlah kunjungan terbanyak di kawasan Asia Tenggara, yaitu sekitar 281 juta kunjungan [2]. Tingginya minat masyarakat terhadap Shopee berdampak pada banyaknya ulasan produk yang diberikan oleh pembeli setiap harinya, sehingga membentuk kumpulan data digital dalam jumlah besar [3]. Kondisi ini menjadikan Shopee sebagai sumber data yang potensial untuk dianalisis menggunakan pendekatan *Natural Language Processing* (NLP), karena ulasan konsumen mengandung informasi penting terkait kualitas produk, pelayanan toko, serta tingkat kepuasan pelanggan [2], [4].

Salah satu kategori dengan tingkat persaingan tinggi di Shopee adalah kategori *fashion*, di mana konsumen memiliki banyak pilihan produk sebelum memutuskan untuk membeli karena banyaknya *merk*, model, warna, dan harga yang ditawarkan [5]. Dalam kondisi tersebut, ulasan produk menjadi salah satu faktor utama yang dijadikan pertimbangan oleh calon pembeli, karena berisi pendapat, perasaan, serta keluhan konsumen terhadap produk dan layanan toko, yang juga berdampak pada reputasi toko, tingkat kepercayaan calon pembeli, dan peringkat produk di *platform* [2], [4], [6]. Apabila ulasan tidak dipahami dengan baik, penjual berpotensi salah dalam mengambil keputusan, seperti menentukan strategi pemasaran, meningkatkan kualitas produk, atau menangani keluhan pelanggan, yang pada akhirnya dapat memengaruhi tingkat penjualan [6], [7].

Pada praktiknya, sebagian besar penjual masih mengandalkan *rating* bintang dan analisis sentimen sederhana yang mengelompokkan ulasan ke dalam kategori positif, negatif, dan netral. Pendekatan ini memang mampu memberikan gambaran umum tentang persepsi konsumen, tetapi belum sepenuhnya dapat merepresentasikan kondisi emosional pengguna secara mendalam, sebab dua ulasan yang sama-sama bernilai positif belum tentu mencerminkan tingkat kepuasan yang sama karena dapat mengandung emosi yang berbeda [6], [8]. Keterbatasan ini menyebabkan sulitnya mengidentifikasi tingkat kepuasan pelanggan secara akurat, karena kepuasan pelanggan tidak hanya dipengaruhi oleh polaritas sentimen, tetapi juga oleh emosi yang melatarbelakangi opini tersebut [9]. Oleh karena itu, informasi emosi perlu dipertimbangkan bersama analisis sentimen agar interpretasi terhadap pengalaman pelanggan dapat dilakukan secara lebih komprehensif dan menjadi dasar yang lebih akurat dalam mengidentifikasi tingkat kepuasan pelanggan. Emosi tersebut dapat tercermin tidak hanya dari teks, tetapi juga dari penggunaan emoji yang umum ditemukan dalam ulasan digital, untuk memperkuat makna, menunjukkan perasaan, atau mengekspresikan reaksi secara spontan dalam ulasan produk [10], [11].

Seiring dengan perkembangan NLP, arsitektur *Transformer* hadir sebagai inovasi yang banyak diterapkan untuk memahami konteks bahasa secara lebih mendalam. Namun, model BERT *multilingual* yang dilatih pada data lintas bahasa tidak dapat secara optimal menangkap kosakata, struktur gramatikal, serta kekhasan bahasa informal Indonesia seperti bahasa gaul, singkatan, dan campuran bahasa (*code-switching*) yang umum ditemukan dalam ulasan *e-commerce* [8], [12]. Oleh karena itu, IndoBERT dikembangkan sebagai model *Transformer* yang dilatih secara khusus menggunakan korpus berbahasa Indonesia berukuran besar, yang mencakup berbagai genre teks mulai dari artikel berita, Wikipedia, hingga percakapan online [7], [8]. IndoBERT terbukti mampu memahami konteks bahasa Indonesia secara lebih akurat dan telah menunjukkan performa yang unggul pada berbagai tugas klasifikasi teks, baik pada analisis sentimen maupun deteksi emosi, dengan akurasi yang kompetitif pada ulasan berbahasa Indonesia [2], [7], [8], [12], [13]. Keunggulan tersebut menjadi dasar penggunaan IndoBERT dalam penelitian ini untuk mengintegrasikan tugas klasifikasi sentimen dan emosi secara bersamaan,

karena informasi emosi dapat memberikan konteks tambahan yang membantu model memahami alasan di balik suatu opini sehingga menghasilkan representasi teks yang lebih kaya. Potensi tersebut mendorong perlunya evaluasi mengenai pengaruh informasi emosi terhadap performa klasifikasi sentimen pada ulasan produk *fashion* di Shopee.

Meskipun berbagai penelitian mengenai analisis sentimen dan deteksi emosi telah dilakukan, sebagian besar masih berfokus pada salah satu tugas secara terpisah dan belum dimanfaatkan lebih lanjut untuk mengidentifikasi tingkat kepuasan pelanggan secara langsung. Penelitian yang menggabungkan informasi sentimen, emosi dan *rating* ulasan sebagai dasar identifikasi kepuasan pelanggan, khususnya pada domain *fashion e-commerce* berbahasa Indonesia menggunakan IndoBERT, masih terbatas. Kondisi ini menunjukkan perlunya pendekatan yang mampu mengintegrasikan ketiga aspek tersebut secara sistematis guna menghasilkan identifikasi kepuasan pelanggan yang lebih komprehensif.

Berdasarkan kondisi tersebut, penelitian ini mengembangkan analisis sentimen berbasis emosi menggunakan IndoBERT pada ulasan produk *fashion* di Shopee. Model yang dikembangkan melakukan klasifikasi sentimen dan emosi secara bersamaan melalui pendekatan *multi-task learning*, untuk memperoleh pemahaman yang lebih komprehensif terhadap opini dan *respons* emosional pelanggan. Selain itu, penelitian ini juga mengevaluasi pengaruh informasi emosi terhadap performa klasifikasi sentimen. Selanjutnya, hasil klasifikasi sentimen dan emosi dikombinasikan dengan *rating* ulasan menggunakan pendekatan *rule-based mapping* untuk mengidentifikasi tingkat kepuasan pelanggan ke dalam kategori *satisfied*, *neutral*, dan *unsatisfied*. Pendekatan ini diharapkan mampu memberikan pemahaman yang lebih komprehensif terhadap pengalaman pelanggan serta memberikan kontribusi praktis bagi penjual dalam meningkatkan kualitas produk, memperbaiki layanan, dan merancang strategi promosi yang lebih tepat sasaran.

1.2. Rumusan Masalah

Berdasarkan latar belakang penelitian, maka rumusan masalah dalam penelitian ini adalah :

1. Bagaimana melakukan analisis sentimen berbasis emosi pada ulasan produk *fashion* menggunakan IndoBERT?

2. Apakah faktor emosi dapat meningkatkan performa klasifikasi sentimen menggunakan IndoBERT pada ulasan produk *fashion*?
3. Bagaimana melakukan klasifikasi tingkat kepuasan pelanggan dari hasil sentimen dan emosi yang dikombinasikan dengan *rating* ulasan ke dalam kategori *satisfied*, *neutral*, dan *unsatisfied*?

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Melakukan analisis sentimen berbasis emosi pada ulasan produk *fashion* menggunakan IndoBERT.
2. Mengevaluasi pengaruh faktor emosi terhadap performa klasifikasi sentimen menggunakan IndoBERT pada ulasan produk *fashion*.
3. Melakukan klasifikasi tingkat kepuasan pelanggan melalui kombinasi hasil analisis sentimen, emosi, dan *rating* ulasan ke dalam kategori *satisfied*, *neutral*, dan *unsatisfied*.

1.4. Batasan Masalah

Agar penelitian lebih terfokus dan terarah, maka batasan masalah dalam penelitian ini adalah :

1. Data penelitian berupa ulasan pelanggan produk *fashion* wanita pada platform *e-commerce* Shopee yang difokuskan pada ulasan dari satu *brand fashion* wanita yaitu Alunicorn Official Store sebagai sampel penelitian. Data dikumpulkan melalui proses *web scraping*.
2. Ulasan yang dianalisis adalah ulasan berbahasa Indonesia.
3. Kategori sentimen dibatasi pada sentimen, yaitu *positive*, *neutral*, dan *negative*.
4. Kategori emosi dibatasi pada emosi dasar, yaitu senang, kagum, sedih, marah, dan netral.
5. Model yang digunakan dalam penelitian ini adalah IndoBERT dengan pendekatan *multitask learning* untuk klasifikasi sentimen dan emosi pada ulasan produk *fashion*.
6. Penelitian ini hanya menganalisis teks ulasan dan tidak membahas konten non-teks seperti gambar atau video.

1.5. Manfaat Penelitian

1.5.1. Manfaat Teoritis

Penelitian ini memberikan kontribusi di bidang NLP bahasa Indonesia, khususnya dalam penerapan model IndoBERT untuk analisis sentimen dan emosi pada ulasan produk di *e-commerce*. Penelitian ini memberikan gambaran tentang pemanfaatan informasi sentimen dan emosi pada ulasan berbahasa Indonesia yang mengandung bahasa *informal* dan emoji, sehingga dapat menjadi referensi bagi pengembangan penelitian analisis teks yang lebih kontekstual di masa mendatang. Selain itu, penelitian ini juga berkontribusi dalam mengeksplorasi pemanfaatan kombinasi hasil sentimen, emosi, dan *rating* sebagai pendekatan alternatif dalam mengidentifikasi tingkat kepuasan pelanggan secara otomatis.

1.5.2. Manfaat Praktis

1. Memberikan informasi yang lebih jelas dan detail mengenai perasaan serta tingkat kepuasan pelanggan terhadap produk *fashion* wanita di Shopee berdasarkan hasil analisis sentimen, emosi dan *rating* ulasan.
2. Membantu pihak toko atau penjual dalam mengenali keluhan dan pujian konsumen secara lebih spesifik, sehingga perbaikan layanan dan kualitas produk dapat dilakukan dengan lebih tepat sasaran.
3. Membantu penjual dalam memantau perubahan sentimen, emosi, dan tingkat kepuasan pelanggan berdasarkan ulasan yang diberikan konsumen.
4. Mendukung pengambilan keputusan dalam menentukan strategi pemasaran, promosi, dan pelayanan berdasarkan hasil analisis sentimen, emosi dan identifikasi kepuasan pelanggan dari ulasan konsumen.

1.6. Kerangka Pemikiran

Kerangka pemikiran penelitian ini menggambarkan alur penelitian yang dimulai dari identifikasi permasalahan, peluang penelitian, pendekatan yang digunakan, metode pengembangan, implementasi, hingga hasil yang diharapkan dalam penelitian ini.



Gambar 1. 1 Bagan Kerangka Pemikiran

Berdasarkan Gambar 1.1, penelitian ini berangkat dari permasalahan bahwa ulasan produk *fashion* di Shopee belum dimanfaatkan secara optimal untuk mengidentifikasi kepuasan pelanggan. Analisis yang selama ini dilakukan umumnya hanya berfokus pada *rating* dan sentimen berbasis polaritas, sehingga informasi emosi pelanggan belum dimanfaatkan secara maksimal.

Di sisi lain, perkembangan NLP dan model IndoBERT memungkinkan analisis sentimen dan emosi pada ulasan berbahasa Indonesia dilakukan secara lebih kontekstual, termasuk dalam menangani karakteristik bahasa *informal* dan penggunaan emoji yang umum ditemukan pada ulasan di *platform e-commerce* seperti Shopee.

Berdasarkan peluang tersebut, analisis kepuasan pelanggan pada ulasan produk *fashion* dilakukan dengan mengklasifikasikan sentimen dan emosi menggunakan IndoBERT, yang kemudian hasilnya dikombinasikan dengan *rating* melalui pendekatan *rule-based* untuk menentukan kategori kepuasan pelanggan.

Pengembangan sistem ini mengikuti alur *Machine Learning Life Cycle* (MLLC) yang meliputi tahapan pengumpulan data, persiapan data, pelabelan dan pengolahan data, analisis data, pelatihan model, pengujian model, serta interpretasi hasil.

Pada tahap implementasi, model IndoBERT digunakan untuk mengklasifikasikan ulasan ke dalam label sentimen (*positive, neutral, negative*) serta label emosi. Kedua label tersebut, bersama dengan *rating* ulasan, selanjutnya diproses melalui sistem *rule-based* untuk menghasilkan output akhir berupa kategori kepuasan pelanggan (*satisfied, neutral, unsatisfied*) yang diperoleh dari kombinasi sentimen, emosi, dan *rating* ulasan menggunakan pendekatan *rule-based*.

1.7. Sistematika Penulisan

Sistematika penulisan tugas akhir ini disusun untuk memberikan gambaran mengenai isi dan alur pembahasan penelitian yang dilakukan. Adapun sistematika penulisan pada tugas akhir ini terdiri dari lima bab, yaitu sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, kerangka pemikiran, serta sistematika penulisan yang menjadi dasar pelaksanaan penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini berisi teori-teori dan penelitian terdahulu yang mendukung penelitian, meliputi konsep *e-commerce*, ulasan pelanggan, analisis sentimen, klasifikasi emosi, emoji, NLP, *Transformer*, IndoBERT, MLLC, serta konsep kepuasan pelanggan yang digunakan dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang digunakan, meliputi tahapan pengumpulan data, *preprocessing* data, pelabelan sentimen dan emosi, perancangan model IndoBERT, proses pelatihan dan pengujian model, metode evaluasi, serta proses identifikasi kepuasan pelanggan menggunakan pendekatan *rule-based mapping*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil implementasi dan evaluasi model yang dikembangkan. Selain itu, bab ini membahas hasil klasifikasi sentimen dan emosi, analisis pengaruh informasi emosi terhadap performa klasifikasi sentimen, serta hasil identifikasi kepuasan pelanggan berdasarkan kombinasi sentimen, emosi, dan *rating* ulasan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya.