

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Aktivitas *e-commerce* di Indonesia menunjukkan bahwa perdagangan digital telah menjadi bagian penting dari ekonomi digital nasional. Laporan Kinerja Perdagangan Melalui Sistem Elektronik (PMSE) Indonesia Tahun 2025 mencatat nilai transaksi *e-commerce* Indonesia tahun 2024 sebesar Rp1.288,93 triliun, meningkat 17,08% dibandingkan tahun 2023, serta melibatkan 4,40 juta unit usaha PMSE [1]. Pada skala ekonomi digital, laporan *e-Conomy SEA 2025* memperkirakan *Gross Merchandise Value* (GMV) ekonomi digital Indonesia mendekati US\$100 miliar, dengan *e-commerce* sebagai kontributor terbesar sebesar US\$71 miliar [2]. Besarnya aktivitas transaksi dan jumlah pelaku usaha tersebut membuat ulasan produk menjadi sumber data tekstual yang relevan untuk memahami pengalaman konsumen pada platform *e-commerce*.

Ulasan produk merupakan konten yang dihasilkan konsumen yang memuat pengalaman konsumen terhadap kualitas produk, layanan penjual, proses pengiriman, dan kesesuaian barang yang diterima. Informasi dalam ulasan dapat membantu calon konsumen menilai produk sebelum pembelian, sedangkan penjual dan platform dapat menggunakannya sebagai masukan untuk memahami respons pengguna terhadap produk dan layanan. Kajian terbaru menunjukkan bahwa ulasan daring berperan dalam keputusan pembelian karena memuat informasi yang dianggap berguna, diagnostik, dan dipengaruhi oleh konteks sosial pembaca ulasan [3, 4]. Dengan demikian, ulasan produk idealnya tidak berhenti sebagai teks mentah, tetapi dapat diolah menjadi informasi terstruktur yang menggambarkan pengalaman konsumen secara lebih rinci.

Analisis ulasan tidak selalu cukup dilakukan melalui klasifikasi sentimen berbasis polaritas. Sentimen positif, negatif, atau netral dapat menggambarkan arah umum penilaian, tetapi belum tentu menjelaskan bentuk respons afektif yang muncul di balik penilaian tersebut. Kajian pemrosesan bahasa alami membedakan analisis sentimen dan analisis emosi karena keduanya memiliki fokus label dan tujuan interpretasi yang berbeda [5, 6]. Pada teks berbahasa Indonesia, pemisahan tugas tersebut juga digunakan dalam evaluasi klasifikasi sentimen dan emosi, misalnya melalui label polaritas untuk sentimen dan label emosi untuk ekspresi

afektif yang lebih spesifik [7].

Klasifikasi emosi relevan untuk ulasan produk karena ulasan dengan polaritas serupa dapat memuat respons afektif yang berbeda. Ulasan bernada negatif dapat menunjukkan *anger* ketika konsumen menyampaikan protes terhadap barang rusak, *sadness* ketika produk tidak memenuhi harapan, atau *fear* ketika ulasan memuat kekhawatiran dan peringatan risiko. Sebaliknya, ulasan bernada positif dapat menunjukkan *happy* ketika konsumen merasa puas atau *love* ketika terdapat ekspresi suka dengan intensitas yang lebih kuat. Perbedaan tersebut menunjukkan bahwa label emosi seperti *anger*, *fear*, *happy*, *love*, dan *sadness* dapat memberi informasi yang lebih spesifik daripada kategori positif, negatif, atau netral [8].

Dalam konteks ulasan produk berbahasa Indonesia, PRDECT-ID merupakan dataset publik yang relevan untuk klasifikasi emosi. Dataset ini berisi 5.400 ulasan produk berbahasa Indonesia dari 29 kategori produk Tokopedia dan menyediakan anotasi emosi untuk lima kelas, yaitu *anger*, *fear*, *happy*, *love*, dan *sadness* [8, 9]. PRDECT-ID menyediakan dasar empiris bagi kajian klasifikasi emosi pada domain ulasan produk Indonesia. Namun, ketersediaan dataset tetap perlu diikuti oleh pemilihan model yang sesuai dengan karakteristik bahasa Indonesia serta evaluasi yang memperhatikan performa dan efisiensi pelatihan.

Salah satu model pra-latih yang relevan untuk pemrosesan bahasa alami berbahasa Indonesia adalah IndoBERT Base dalam kerangka IndoNLU. IndoNLU menyediakan sumber daya dan *benchmark* untuk evaluasi pemahaman bahasa Indonesia, sedangkan IndoBERT dilatih pada korpus bahasa Indonesia berskala besar [10]. Relevansi IndoBERT pada domain ulasan produk ditunjukkan oleh studi klasifikasi emosi berbasis BERT pada PRDECT-ID dan ulasan *e-commerce* Indonesia [11, 12]. Dengan dasar tersebut, IndoBERT Base dapat digunakan sebagai model dasar untuk menguji adaptasi model pada tugas klasifikasi emosi ulasan produk berbahasa Indonesia.

Adaptasi model pra-latih untuk klasifikasi teks dapat dilakukan melalui *full fine-tuning*, yaitu pembaruan seluruh parameter model selama pelatihan. Pendekatan tersebut dapat menghasilkan performa yang kompetitif, tetapi pembaruan semua parameter meningkatkan kebutuhan komputasi ketika eksperimen dijalankan pada banyak konfigurasi atau pada sumber daya terbatas [13, 14]. Oleh karena itu, evaluasi model klasifikasi tidak hanya perlu melihat performa prediksi, tetapi juga efisiensi pelatihan.

Kebutuhan efisiensi tersebut mendorong penggunaan *Parameter-Efficient Fine-Tuning* (PEFT), yaitu strategi adaptasi model pra-latih dengan jumlah

parameter latih yang lebih kecil dibandingkan *full fine-tuning* [14]. Salah satu metode PEFT yang banyak dikaji adalah *Low-Rank Adaptation* atau LoRA. LoRA membekukan bobot utama model pra-latih dan melatih matriks tambahan berderajat rendah selama adaptasi model [13]. Pada studi aslinya, LoRA dilaporkan dapat mengurangi jumlah parameter latih dan kebutuhan memori GPU pada skenario model bahasa besar dibandingkan *full fine-tuning* [13]. Meskipun demikian, performa dan efisiensi LoRA tidak dapat diasumsikan selalu lebih baik pada seluruh model, dataset, dan tugas sehingga evaluasi empiris tetap diperlukan pada konteks yang spesifik.

Penelitian terdahulu telah mengkaji klasifikasi emosi ulasan produk Indonesia menggunakan metode klasik, model berbasis BERT, dan model bahasa besar pada skenario tertentu. Kajian berbasis SVM menunjukkan relevansi metode klasik pada klasifikasi emosi ulasan *e-commerce* [15]. Kajian berbasis BERT, IndoBERT, DistilBERT, dan model bahasa besar juga menunjukkan bahwa domain ulasan produk Indonesia telah digunakan dalam evaluasi klasifikasi emosi [11, 12, 16]. Di sisi lain, LoRA telah digunakan dalam beberapa konteks klasifikasi teks dan klasifikasi emosi, termasuk tugas emosi dan empati serta tugas teks berbahasa Indonesia di luar domain ulasan produk Tokopedia [17, 18, 19, 20]. Berdasarkan literatur yang ditinjau, celah penelitian terletak pada masih terbatasnya kajian yang mengevaluasi beberapa konfigurasi LoRA pada IndoBERT Base untuk klasifikasi lima emosi pada dataset ulasan produk Tokopedia berbahasa Indonesia, sekaligus membandingkannya secara langsung dengan *full fine-tuning* menggunakan ukuran performa dan efisiensi pelatihan yang sama.

Berdasarkan uraian tersebut, penelitian ini perlu dilakukan untuk mengevaluasi penerapan LoRA pada IndoBERT Base dalam klasifikasi lima emosi pada dataset ulasan produk Tokopedia berbahasa Indonesia. Evaluasi diarahkan pada pengaruh variasi *rank*, *alpha*, dan *target modules* terhadap performa klasifikasi, serta perbandingan LoRA dengan *full fine-tuning* berdasarkan metrik performa dan efisiensi pelatihan. Dengan arah tersebut, penelitian ini menempatkan LoRA sebagai pendekatan adaptasi yang perlu diuji secara empiris, bukan sebagai metode yang diasumsikan lebih unggul sejak awal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam skripsi ini adalah sebagai berikut.

1. Bagaimana penerapan LoRA melalui variasi *rank*, *alpha*, dan *target modules* memengaruhi performa IndoBERT pada klasifikasi emosi dataset ulasan produk Tokopedia berbahasa Indonesia berdasarkan *macro-F1*?
2. Konfigurasi LoRA mana yang memberikan keseimbangan terbaik antara *validation macro-F1* dan efisiensi sumber daya berdasarkan *trainable parameters*, waktu pelatihan, dan *peak GPU memory*?
3. Bagaimana perbandingan performa dan efisiensi konfigurasi LoRA terpilih terhadap *full fine-tuning* IndoBERT Base pada *test split*?

1.3 Batasan Permasalahan

Batasan permasalahan diuraikan sebagai berikut.

1. Data penelitian dibatasi pada dataset gabungan ulasan produk Tokopedia berbahasa Indonesia yang terdiri dari PRDECT-ID hasil *refinement* dan data tambahan hasil *scraping* Tokopedia. Dataset ini tidak diklaim mewakili seluruh ulasan yang tersedia pada platform Tokopedia.
2. Label yang digunakan diposisikan sebagai label target otomatis berbantuan sistem melalui Label Studio dengan panduan anotasi eksplisit. Label tersebut digunakan sebagai target pelatihan dan evaluasi, tetapi tidak diklaim sebagai label referensi manual final yang telah divalidasi manusia.
3. Eksperimen pencarian konfigurasi difokuskan pada variasi parameter LoRA, yaitu *rank* (r), *alpha* (α), dan *target modules* pada modul proyeksi *attention* serta *dense* yang diuji pada IndoBERT. *Hyperparameter* pelatihan lain ditetapkan sebagai konfigurasi tetap sesuai rancangan eksperimen.
4. Metrik performa yang digunakan meliputi *accuracy*, *macro-precision*, *macro-recall*, dan *macro-F1*, dengan *macro-F1* sebagai metrik utama. Evaluasi efisiensi difokuskan pada jumlah dan persentase *trainable parameters*, waktu pelatihan, serta *peak GPU memory*.
5. Ruang lingkup tidak mencakup *deployment*, pengujian *real-time*, integrasi model ke sistem *e-commerce*, pengujian pada dataset lain di luar dataset akhir penelitian, analisis bisnis Tokopedia, atau klaim generalisasi ke seluruh domain dan platform *e-commerce*.

1.4 Tujuan Penelitian

Tujuan penelitian dirumuskan sebagai berikut.

1. Menganalisis penerapan LoRA melalui variasi *rank*, *alpha*, dan *target modules* terhadap performa IndoBERT pada klasifikasi emosi dataset ulasan produk Tokopedia berbahasa Indonesia berdasarkan *macro-F1*.
2. Menentukan konfigurasi LoRA yang memberikan keseimbangan terbaik antara *validation macro-F1* dan efisiensi sumber daya berdasarkan *trainable parameters*, waktu pelatihan, dan *peak GPU memory*.
3. Membandingkan performa dan efisiensi konfigurasi LoRA terpilih dengan *full fine-tuning* IndoBERT Base pada *test split*.

1.5 Manfaat Penelitian

Manfaat penelitian ini dibagi menjadi manfaat teoritis dan manfaat praktis sebagai berikut.

1. Manfaat teoritis
 - (a) Memberikan kontribusi terhadap pengembangan kajian *Natural Language Processing* bahasa Indonesia, khususnya pada tugas klasifikasi emosi teks ulasan produk berbahasa Indonesia.
 - (b) Menambah kajian empiris mengenai penggunaan IndoBERT Base untuk klasifikasi teks berbahasa Indonesia dengan lima kelas emosi, yaitu *anger*, *fear*, *happy*, *love*, dan *sadness*.
 - (c) Memberikan gambaran mengenai penerapan *Parameter-Efficient Fine-Tuning*, khususnya LoRA, pada model bahasa pra-latih untuk tugas klasifikasi emosi.
 - (d) Memberikan pembandingan akademik antara LoRA dan *full fine-tuning* IndoBERT dari sisi performa klasifikasi dan efisiensi penggunaan sumber daya komputasi.
 - (e) Menjadi rujukan awal bagi penelitian selanjutnya yang mengkaji model hemat parameter pada klasifikasi teks berbahasa Indonesia, terutama dengan memperhatikan keterbatasan label target otomatis.

2. Manfaat praktis

- (a) Memberikan acuan eksperimen bagi mahasiswa atau praktisi akademik yang ingin menerapkan LoRA pada IndoBERT untuk tugas klasifikasi teks berbahasa Indonesia.
- (b) Memberikan pertimbangan dalam memilih metode *fine-tuning* yang lebih efisien berdasarkan jumlah *trainable parameters*, waktu pelatihan, dan penggunaan memori GPU.
- (c) Membantu memahami hubungan antara performa model dan efisiensi komputasi melalui metrik *accuracy*, *macro-precision*, *macro-recall*, *macro-F1*, *confusion matrix*, dan *error analysis*.
- (d) Memberikan gambaran awal mengenai pengolahan ulasan produk berbahasa Indonesia secara lebih rinci melalui klasifikasi emosi, tanpa mengklaim bahwa model yang dihasilkan siap digunakan langsung pada sistem produksi.

1.6 Sistematika Penulisan

Sistematika penulisan disusun sebagai berikut.

- Bab 1 PENDAHULUAN
Bab ini membahas latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- Bab 2 LANDASAN TEORI
Bab ini membahas landasan teori dan penelitian terkait, meliputi pemrosesan bahasa alami, klasifikasi teks, ulasan produk Tokopedia sebagai domain dataset penelitian, PRDECT-ID, klasifikasi emosi, label target otomatis, prapemrosesan teks, tokenisasi, *Transformer*, BERT, IndoBERT, *fine-tuning*, PEFT, LoRA, fungsi kerugian, metrik evaluasi, dan efisiensi parameter.
- Bab 3 METODOLOGI PENELITIAN
Bab ini menjelaskan alur metodologi dari pembentukan data sampai evaluasi model. Pembahasan mencakup sumber data, pengambilan data tambahan Tokopedia melalui *scraping* semi-manual, pelabelan otomatis berbasis panduan, kontrol kualitas label otomatis, prapemrosesan, pembentukan

dataset gabungan penelitian, pembagian *train/validation/test*, audit kebocoran data, tokenisasi, konfigurasi *full fine-tuning* dan LoRA, pemilihan model berdasarkan *validation macro-F1*, metrik evaluasi, analisis kesalahan, lingkungan eksperimen, dan batasan metodologis.

- Bab 4 HASIL DAN DISKUSI

Bab ini menyajikan hasil pembentukan dataset, performa baseline *full fine-tuning*, performa seluruh konfigurasi LoRA, pemilihan LoRA terpilih dan LoRA efisien, perbandingan performa dan efisiensi, evaluasi pada *test split*, analisis *classification report*, *confusion matrix*, pembahasan rumusan masalah, batasan interpretasi hasil, dan ringkasan temuan.

- Bab 5 SIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan berdasarkan rumusan masalah, hasil pengujian dan pembahasan yang telah dilakukan. Bab ini juga memuat saran untuk penelitian selanjutnya.

