

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan media sosial dalam beberapa tahun terakhir telah menghasilkan data teks dalam jumlah besar yang bersifat spontan, informal, dan merepresentasikan ekspresi emosi pengguna secara langsung. Teks-teks tersebut banyak dimanfaatkan sebagai sumber data dalam penelitian *Natural Language Processing* (NLP), khususnya pada tugas analisis sentimen dan deteksi emosi. Informasi emosional yang terkandung dalam teks memiliki peran penting dalam berbagai aplikasi, seperti pemantauan opini publik, analisis perilaku pengguna, serta sistem pendukung kesehatan mental berbasis teknologi [1, 2].

Berbeda dengan analisis sentimen yang umumnya hanya mengklasifikasikan polaritas positif, negatif, atau netral, deteksi emosi bertujuan mengidentifikasi kategori emosi yang lebih spesifik, seperti gembira, marah, sedih, takut, dan cinta. Tugas ini bersifat lebih kompleks karena emosi sering kali diekspresikan secara implisit, kontekstual, dan dipengaruhi oleh variasi bahasa informal yang lazim digunakan di media sosial [3]. Oleh karena itu, deteksi emosi berbasis teks memerlukan model representasi bahasa yang mampu menangkap konteks semantik secara mendalam.

Kemajuan signifikan dalam NLP dicapai melalui hadirnya arsitektur *Transformer*, yang diperkenalkan oleh Vaswani et al. [4]. Arsitektur ini menggantikan model berurutan seperti *Recurrent Neural Network* dengan mekanisme *self-attention* yang memungkinkan pemodelan dependensi jarak jauh antar token secara paralel. Pendekatan ini menjadi dasar bagi pengembangan berbagai model bahasa prelatih modern, termasuk BERT dan RoBERTa.

BERT (*Bidirectional Encoder Representations from Transformers*) merupakan model prelatih berbasis *Transformer encoder* yang dilatih menggunakan dua tujuan utama, yaitu *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP) [5]. RoBERTa dikembangkan sebagai penyempurnaan dari BERT dengan mempertahankan arsitektur *Transformer* yang sama, namun menghilangkan tugas NSP, menggunakan data pelatihan yang lebih besar, menerapkan *dynamic masking*, serta melakukan optimasi pada skema pelatihan [6]. Dengan demikian, persamaan utama antara BERT dan RoBERTa terletak pada arsitektur dasar

dan mekanisme *self-attention*, sedangkan perbedaannya terletak pada strategi pra-pelatihan dan optimasi pelatihan model.

Dalam konteks pengembangan model bahasa, pendekatan monolingual dan multilingual menjadi dua strategi utama yang umum digunakan. Model monolingual dilatih secara khusus menggunakan data dari satu bahasa tertentu, sehingga diharapkan mampu menangkap karakteristik linguistik lokal secara lebih mendalam. Sebaliknya, model multilingual dilatih menggunakan data dari banyak bahasa sekaligus, dengan tujuan memanfaatkan transfer pengetahuan lintas bahasa (*cross-lingual transfer*) agar dapat digunakan pada berbagai bahasa dengan satu model yang sama [7].

IndoBERT merupakan contoh model pralatih monolingual yang dikembangkan khusus untuk bahasa Indonesia dengan memanfaatkan korpus berbahasa Indonesia dalam skala besar [8]. Pendekatan monolingual ini memungkinkan IndoBERT mempelajari struktur morfologi, kosakata, dan pola sintaksis bahasa Indonesia secara lebih spesifik. Di sisi lain, XLM-RoBERTa (XLM-R) merupakan model pralatih multilingual yang dilatih menggunakan korpus *CommonCrawl* yang mencakup lebih dari 100 bahasa, termasuk bahasa Indonesia [7]. Dengan pendekatan ini, XLM-R mengandalkan kemampuan transfer lintas bahasa untuk menghasilkan representasi yang dapat digunakan pada berbagai bahasa.

Sejumlah penelitian menunjukkan bahwa IndoBERT memberikan performa yang kompetitif pada tugas klasifikasi emosi berbahasa Indonesia. Setiawan et al. melaporkan akurasi sebesar 72% pada dataset media sosial berbahasa Indonesia [9], sementara Nugroho et al. memperoleh akurasi hingga 77% pada dataset Twitter berbahasa Indonesia [10]. Namun, beberapa studi lain pada tugas klasifikasi sentimen menunjukkan bahwa model multilingual seperti XLM-R dapat menghasilkan performa yang sebanding dengan model monolingual dengan perbedaan akurasi hanya 0,1% [11]. Hal ini menimbulkan pertanyaan mengenai sejauh mana perbedaan pendekatan monolingual dan multilingual memengaruhi performa model pada tugas deteksi emosi yang bersifat lebih kompleks.

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis dan membandingkan kinerja model IndoBERT dan XLM-R pada tugas klasifikasi emosi teks berbahasa Indonesia. Penelitian ini menggunakan skema eksperimen yang terkontrol dengan dataset, konfigurasi pelatihan, dan metrik evaluasi yang identik, sehingga perbedaan performa yang muncul dapat dikaitkan secara langsung dengan karakteristik dari masing-masing model.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang diuraikan, dapat dirumuskan masalah yang ingin diselesaikan yaitu:

1. Bagaimana mengimplementasikan model deteksi emosi berbahasa Indonesia berbasis IndoBERT dan XLM-RoBERTa?
2. Bagaimana perbandingan kinerja antara model *IndoBERT* dan *XLM-RoBERTa* berdasarkan metrik akurasi, presisi, *recall*, *F1-score*?

1.3 Batasan Permasalahan

Agar pembahasan lebih terarah dan sesuai tujuan, penelitian dibatasi pada aspek-aspek berikut:

1. Penelitian hanya berfokus pada dua model transformer, yaitu IndoBERT dan XLM-RoBERTa, tanpa mempertimbangkan model lain seperti mBERT, IndoBERTweet, atau arsitektur LLM generasi terbaru.
2. Penelitian hanya mencakup lima kelas emosi (gembira, marah, sedih, takut, cinta) sesuai ketersediaan label pada dataset.
3. Penelitian hanya mengevaluasi model menggunakan metrik akurasi, presisi, recall, dan F1-score, tanpa melibatkan metrik lanjutan.
4. Penelitian tidak membahas integrasi model ke dalam aplikasi nyata seperti chatbot, sistem rekomendasi, atau sistem monitoring media sosial.
5. Penelitian ini difokuskan pada tahap klasifikasi emosi berbasis teks, tanpa membahas aspek multimodal seperti gambar, audio, atau video.

1.4 Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Mengimplementasikan model deteksi emosi teks berbahasa Indonesia menggunakan model prelatih berbasis *Transformer*, yaitu IndoBERT dan XLM-RoBERTa, melalui proses *fine-tuning* yang terstruktur.

2. Menganalisis dan membandingkan kinerja kedua model berdasarkan metrik evaluasi akurasi, presisi, *recall*, dan *F1-score* pada data uji internal.
3. Mengevaluasi kemampuan generalisasi IndoBERT dan XLM-R melalui pengujian pada dataset eksternal yang memiliki karakteristik distribusi dan gaya bahasa berbeda.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan beberapa manfaat sebagai berikut:

1. Memberikan acuan empiris mengenai efektivitas model transformer monolingual (IndoBERT) dibandingkan model multilingual (XLM-RoBERTa) pada tugas deteksi emosi teks berbahasa Indonesia.
2. Menyediakan panduan implementasi yang jelas terkait preprocessing dan tokenisasi sehingga dapat dimanfaatkan oleh peneliti atau praktisi yang mengembangkan sistem NLP berbahasa Indonesia.
3. Menambah kontribusi ilmiah terkait performa model transformer dalam domain analisis emosi, khususnya ketika dataset terbatas dan memiliki karakteristik bahasa informal khas media sosial Indonesia.
4. Memberikan dasar untuk pengembangan lanjutan, seperti peningkatan performa melalui data augmentation, optimisasi arsitektur, atau integrasi ke aplikasi berbasis emosi.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini berisi uraian mengenai latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi.
- Bab 2 LANDASAN TEORI
Bab ini membahas teori-teori yang mendasari penelitian, meliputi teori dan model emosi, media sosial sebagai sumber data, deteksi emosi berbasis teks,

arsitektur model BERT, IndoBERT, dan XLM-R, serta metrik evaluasi yang digunakan seperti akurasi, presisi, *recall*, dan *f1-score*.

- Bab 3 METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan penelitian secara rinci, meliputi sumber dan teknik pengumpulan data, tahapan pra-pemrosesan teks, arsitektur dan penerapan model IndoBERT, serta metode evaluasi kinerja model.

- Bab 4 HASIL DAN DISKUSI

Bab ini menyajikan hasil eksperimen komparasi kinerja model IndoBERT dan XLM-R dalam klasifikasi emosi teks berbahasa Indonesia. Pembahasan meliputi analisis hasil evaluasi, perbandingan kinerja kedua model berdasarkan setiap metrik evaluasi.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini memuat simpulan dari hasil penelitian yang telah dilakukan sesuai dengan rumusan masalah, serta memberikan saran untuk penelitian lebih lanjut agar dapat mengembangkan penelitian ini ke arah yang lebih luas.

