

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam kehidupan sehari-hari, emosi memiliki peran yang sangat penting dalam interaksi manusia. Emosi tidak hanya diekspresikan melalui perubahan wajah, intonasi suara, atau gestur tubuh, tetapi juga dalam bahasa tertulis [1]. Secara umum, emosi dikategorikan ke dalam kelompok positif dan negatif, yang mencakup label spesifik seperti marah (*Anger*), sedih (*Sad*), takut (*Fear*), senang (*Joy*), dan cinta (*Love*). Studi mengenai deteksi emosi kini menjadi sangat relevan di berbagai bidang, mulai dari psikologi hingga ilmu komputer, karena kemampuannya dalam membantu memahami perilaku dan respons sosial masyarakat [2].

Revolusi digital telah membawa perubahan besar dalam pola komunikasi dan mengekspresikan perasaan masyarakat Indonesia. Media sosial kini menjadi ruang publik utama untuk menyampaikan opini dan emosi, seiring meningkatnya interaksi sosial secara daring melalui berbagai platform [3]. Di Indonesia, Fenomena ini tercatat melalui data terbaru per Oktober 2025, yang mencatat lonjakan signifikan identitas pengguna media sosial hingga mencapai 180 juta pengguna atau setara dengan 62,9% dari total populasi. Pertumbuhan sebesar 26,0% dibanding tahun sebelumnya ini mengindikasikan eskalasi interaksi digital yang masif. Tingginya ketergantungan masyarakat pada ranah digital juga tercermin dari durasi penggunaan media sosial yang mencapai 21 jam 50 menit per minggu, serta rata-rata penggunaan 7,7 platform media sosial berbeda oleh setiap individu setiap bulannya [4]. Meski demikian, platform berbasis teks seperti Twitter/X dan komentar YouTube menjadi wadah yang sangat produktif dalam menghasilkan data opini publik setiap detik.

Lonjakan volume data teks yang terjadi setiap hari menghadirkan peluang sekaligus tantangan. Kemampuan untuk mendeteksi emosi dari tumpukan data tersebut memberikan kontribusi strategis bagi pengambilan keputusan di berbagai sektor, mulai dari pemantauan isu politik, manajemen reputasi perusahaan, hingga

analisis kesehatan mental masyarakat [5]. Namun, menganalisis data dalam skala besar secara manual adalah hal yang mustahil karena keterbatasan waktu dan subjektivitas manusia.

Oleh karena itu, permasalahan utama yang dihadapi saat ini adalah bagaimana membangun sistem otomatis yang mampu mengklasifikasikan emosi dengan akurat dari data teks yang tidak terstruktur. Tantangan ini semakin kompleks pada data berbahasa Indonesia yang memiliki variasi gaya bahasa tinggi, mulai dari bahasa baku hingga bahasa gaul. Untuk menjawab tantangan tersebut, dibutuhkan pendekatan berbasis kecerdasan buatan, khususnya *Deep Learning*. Salah satu arsitektur yang terbukti efektif untuk menangani data teks sekuensial adalah *Long Short-Term Memory (LSTM)*, yang memiliki keunggulan dalam menangkap konteks kalimat jangka panjang.

Meskipun LSTM populer digunakan, penelitian yang menguji efektivitasnya pada dataset bahasa Indonesia yang bersifat heterogen, menggabungkan data formal dan informal, masih perlu dieksplorasi lebih lanjut. Berdasarkan hal tersebut, penelitian ini difokuskan pada analisis kinerja model LSTM untuk klasifikasi emosi. Penelitian ini menggunakan dataset gabungan yang bersumber dari GitHub, Kaggle, Twitter/X, dan YouTube dengan total 10.667 data teks. Penggabungan sumber data ini bertujuan untuk memperkaya variasi konteks bahasa, sehingga diharapkan dapat menghasilkan model yang lebih kuat dalam mengenali emosi pada teks berbahasa Indonesia.

1.2 Rumusan Masalah

Berdasar latar belakang masalah, rumusan masalah dalam penelitian ini Adalah:

1. Bagaimana kinerja model LSTM dalam mengklasifikasikan emosi teks berbahasa Indonesia, diukur menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*?
2. Bagaimana karakteristik dan pola kesalahan yang dihasilkan model LSTM dalam mengenali berbagai kategori emosi yang berbeda?

1.3 Batasan Masalah

Agar penelitian lebih terarah, Batasan masalah yang ditetapkan Adalah sebagai berikut:

1. Data yang digunakan berupa teks berbahasa Indonesia yang dikumpulkan dari empat sumber, yaitu GitHub, Kaggle, Twitter/X, dan YouTube. Seluruh data tersebut digabungkan sehingga menghasilkan total 10.667 data teks, sehingga variasi dan keragaman konteks emosi menjadi lebih luas.
2. Klasifikasi emosi terbatas pada lima kelas utama: *Anger*, *Sad*, *Fear*, *Joy*, dan *Love*.
3. Model *deep learning* yang dianalisis kinerjanya adalah *Long Short-Term Memory (LSTM)*.
4. Penelitian ini tidak membandingkan secara mendalam dengan arsitektur lain seperti *Transformer (BERT)* atau *CNN*.

1.4 Tujuan Penelitian

1. Merancang dan membangun model klasifikasi emosi menggunakan arsitektur *Long Short-Term Memory (LSTM)* yang sesuai untuk data teks berbahasa Indonesia.
2. Menganalisis dan mengukur kinerja model LSTM dalam klasifikasi emosi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.
3. Mengidentifikasi kategori emosi yang paling berhasil dan paling sulit dikenali oleh model, untuk memahami kelebihan dan kekurangan arsitektur LSTM pada dataset yang digunakan.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini dibagi menjadi dua, yaitu:

1. Manfaat Teoritis

Memberikan kontribusi pada studi literatur di bidang Pemrosesan Bahasa Alami untuk Bahasa Indonesia, khususnya dalam evaluasi kinerja model LSTM pada dataset yang beragam. Hasil penelitian ini dapat menjadi acuan bagi penelitian sejenis di masa depan.

2. Manfaat Praktis

Memberikan wawasan tentang efektivitas metode deteksi emosi otomatis yang hasilnya dapat menjadi dasar bagi perusahaan untuk mengembangkan sistem analisis umpan balik pelanggan, pemantauan merek, atau analisis sentimen publik yang lebih canggih.

1.6 Sistematika Penulisan

Sistematika penulisan yang akan diajukan pada penyusunan skripsi ini antara lain sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan secara singkat tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi teori-teori dasar yang menjadi acuan dalam penelitian, meliputi *Natural Language Processing (NLP)*, klasifikasi teks, deteksi emosi, *deep learning*, arsitektur *Long Short-Term Memory (LSTM)*, serta metrik evaluasi model.

BAB III METODE PENELITIAN

Bab ini menjelaskan langkah-langkah kerja yang dilakukan dalam penelitian secara rinci, mulai dari teknik pengumpulan data, proses pelabelan, tahapan *preprocessing* teks, perancangan arsitektur model LSTM, hingga skema pengujian dan evaluasi model.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari eksperimen yang telah dilakukan. Isi bab ini mencakup hasil kinerja model dalam bentuk metrik evaluasi, analisis *confusion matrix*, serta pembahasan mendalam mengenai temuan dan analisis kesalahan model.

BAB V PENUTUP

Bab ini berisi rangkuman dari seluruh hasil penelitian dalam bentuk kesimpulan yang menjawab rumusan masalah, serta saran untuk pengembangan atau penelitian lebih lanjut di masa mendatang.