

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan era digital mendorong intensitas interaksi masyarakat di berbagai platform media sosial, sehingga menghasilkan akumulasi data dalam jumlah besar yang bersifat tidak terstruktur [1]. Di antara ragam platform yang ada, platform Threads menuai popularitasnya berkat tokoh-tokoh seperti artis dan *influencer* media sosial yang menggunakan platform tersebut, selain itu platform ini juga unggul dengan kemudahan aksesnya yang terintegrasi dengan platform Instagram. Threads menawarkan ruang baru yang ideal untuk berbagi pembahasan dan pembaruan secara *real time* [2].

Namun demikian, teks yang dihasilkan di platform Threads khususnya dalam bahasa Indonesia mengandung kompleksitas linguistik yang tinggi, mencakup penggunaan bahasa gaul (*slang*), singkatan, serta ketidakbakuan ejaan [3]. Kondisi tersebut menyebabkan analisis sentimen secara manual menjadi tidak efisien dan rentan terhadap bias subjektif. Cakupan analisis dalam penelitian ini tidak dibatasi pada sentimen positif, negatif, atau netral, melainkan diarahkan pada tugas klasifikasi emosi (*Emotion Classification*) dengan mengacu pada delapan kelas emosi dasar Plutchik [4]. Pendekatan yang lebih terperinci ini dapat memberikan pemahaman yang lebih mendalam terhadap opini masyarakat.

Sebagai konteks pengujian kemampuan model klasifikasi emosi, penelitian ini menggunakan data teks unggahan publik yang berkaitan dengan bencana banjir di Sumatera pada November 2025 [5]. Peristiwa ekstrem tersebut tidak hanya mengakibatkan kerusakan infrastruktur secara fisik, tetapi juga memicu diskusi yang signifikan di ruang digital. Bencana berskala besar ini dipilih sebagai dasar pembentukan *dataset* karena secara alami menimbulkan respons emosional publik yang kompleks dan beragam. Keberagaman emosi yang terkandung dalam data tersebut menghasilkan kumpulan data yang beragam dan ideal untuk mengevaluasi batas performa arsitektur kecerdasan buatan.

Seiring dengan kemajuan di bidang NLP, arsitektur berbasis Transformer telah menjadi standar terkini (*state of the art*) dalam pemrosesan bahasa alami berkat kemampuannya menangkap relasi antarkata melalui mekanisme *self attention* [6]. Meski demikian, masih terdapat perdebatan mengenai arsitektur mana

yang paling sesuai untuk mengolah data teks media sosial yang mengandung banyak *noise*. Penelitian ini membandingkan dua pendekatan, pre trained model IndoBERT sebagai model monolingual yang dirancang khusus untuk memahami konteks bahasa Indonesia [7], serta XLM-RoBERTa sebagai model multilingual berskala besar yang mampu memproses 100 bahasa secara bersamaan [8]. Perbandingan ini bertujuan untuk menilai apakah spesialisasi bahasa lokal menghasilkan kinerja yang lebih unggul dibandingkan generalisasi lintas bahasa dalam menangani variasi bahasa pada platform Threads.

Salah satu studi terkait klasifikasi emosi yang dilakukan oleh Ameer I. et al. [9] Mengevaluasi sepuluh model yang mencakup model *deep learning* (LSTM dan Bi-LSTM) serta *transfer learning* (XLNet, DistilBERT, RoBERTa, dan RoBERTa-MA). Seluruh model diuji pada dua korpus, yaitu SemEval-2018 E-c (bahasa Inggris) dan Ren-CECps (bahasa Cina). Pada SemEval 2018 Task 1C, kinerja terbaik diperoleh oleh RoBERTa-MA (RoBERTa-Multi-Attention) dengan akurasi 62.4%, micro F_1 74.2%, dan *macro* F_1 60.3%. Adapun pada dataset Ren-CECps, tiga model menunjukkan performa terbaik, yaitu XLNet-MA (eXtra Large Neural Network with Multi Attention) dengan akurasi 45.6%, micro F_1 48.3%, *macro* F_1 41.6%, DistilBERT-MA (DistilBERT with Multi Attention) dengan akurasi 42.5%, micro F_1 49.2%, *macro* F_1 43.0%, serta RoBERTa-MA dengan akurasi 43.6%, micro F_1 49.7%, dan *macro* F_1 42.4%.

Model berbasis Transformer kerap menghadapi keterbatasan interpretabilitas akibat sifatnya yang menyerupai *black-box*, di mana proses dari teks *input* hingga prediksi emosi sulit untuk dijelaskan secara transparan [10]. Keterbatasan ini menjadi kelemahan yang signifikan, terutama ketika *output* model digunakan sebagai dasar pengambilan keputusan. Untuk mengatasi permasalahan tersebut, integrasi metodologi *Explainable Artificial Intelligence* (XAI) menjadi sangat relevan. Penelitian ini memanfaatkan metode SHAP (*SHapley Additive exPlanations*) untuk menguraikan proses pengambilan keputusan model [11].

Penelitian yang dilakukan oleh Najm S. dan Mukhtar H. [12] Menyajikan perbandingan model transformer pada tugas klasifikasi emosi serta menginterpretasikan hasil model dengan menggunakan SHAP. Penelitian ini membandingkan tiga model transformer (DistilBERT, BERT, dan RoBERTa) dalam pengklasifikasian 6 emosi (*anger, disgust, fear, joy, sadness, dan surprise*) menggunakan dataset GoEmotions. Hasilnya, model RoBERTa berhasil meraih skor akurasi tertinggi sebesar 83,55% dengan *test macro* F_1 82.51%, diikuti model lainnya DistilBERT (*accuracy*=82.35%, *test macro* F_1 =81.56%), dan BERT

(*accuracy*=81.87%, *test macro F₁*=80.56%).

Penelitian lainnya, oleh Loekito J. et al. [13] Mengusulkan model IndoBERT dengan *Valence–Arousal–Dominance* (VAD) untuk menganalisis dialog konseling berbahasa Indonesia. Dataset terdiri dari 8070 dialog konseling, yang dikumpulkan pada rentang waktu tahun 2021 hingga 2024. Hasilnya, arsitektur *VAD-Calibrated* IndoBERT berhasil meraih skor *macro F₁* tertinggi sebesar 94.23%, yang merupakan hasil dari peningkatan sebesar 2.48% jika dibandingkan dengan model *baseline* biasa.

1.2 Rumusan Masalah

1. Bagaimana perbandingan performa (*Accuracy, Precision, Recall, dan Macro-Averaged F1-Score*) antara model IndoBERT dan XLM-RoBERTa dalam mengklasifikasikan delapan emosi dasar Plutchik pada dataset unggahan platform Threads yang membahas krisis kebencanaan banjir Sumatera?
2. Bagaimana metode SHAP (*SHapley Additive exPlanations*) dapat diimplementasikan untuk memberikan interpretabilitas (Global dan Local Explanations) dalam mengungkap kontribusi setiap token yang mendasari keputusan klasifikasi dari model Transformer yang bersifat *black-box*?

1.3 Batasan Permasalahan

1. Dataset teks diperoleh secara eksklusif melalui teknik *web scraping* pada platform media sosial Threads.
2. Pengumpulan data dibatasi pada unggahan publik yang berkaitan dengan peristiwa banjir di wilayah Sumatera selama periode November 2025.
3. Pelabelan dan klasifikasi teks dilakukan dengan acuan 8 (delapan) emosi dasar dalam teori *Wheel of Emotions* Robert Plutchik, meliputi *Anger* (Marah), *Anticipation* (Antisipasi), *Disgust* (Muak), *Fear* (Takut), *Joy* (Sukacita), *Sadness* (Sedih), *Surprise* (Terkejut), dan *Trust* (Percaya).
4. Proses pelabelan emosi utama pada dataset dilakukan secara otomatis menggunakan pendekatan berbasis leksikon, tanpa melibatkan validasi *Inter-Annotator Agreement*.

5. Pemodelan *Natural Language Processing* (NLP) difokuskan pada evaluasi komparatif dua *Pre-trained Language Model* (PLM) berbasis arsitektur Transformer, yaitu IndoBERT sebagai model monolingual dan XLM-RoBERTa sebagai model multilingual.

1.4 Tujuan Penelitian

1. Membandingkan performa (Accuracy, Precision, Recall, dan *macro-Averaged F1-Score*) antara model IndoBERT dan XLM-RoBERTa dalam melakukan klasifikasi delapan emosi dasar Plutchik pada dataset teks unggahan platform Threads terkait peristiwa banjir Sumatera.
2. Mengimplementasikan metode SHAP (SHapley Additive exPlanations) untuk memberikan interpretabilitas model (Global dan Local Explanations) serta mengidentifikasi bobot kontribusi setiap token (kata) yang mendasari keputusan klasifikasi dari arsitektur *black-box* tersebut.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat, baik secara keilmuan maupun praktis, yang diuraikan dalam poin-poin berikut:

1. Menyajikan evaluasi komparatif mengenai performa model *Transformer monolingual* (IndoBERT) dan *multilingual* (XLM-RoBERTa) dalam tugas klasifikasi emosi berbasis data teks tidak terstruktur dari platform Threads.
2. Mendemonstrasikan efektivitas penerapan *Explainable Artificial Intelligence* (XAI), khususnya metode SHAP, untuk menginterpretasikan keputusan prediksi dari model *deep learning* yang bersifat *black-box*, sehingga hasil klasifikasi teks menjadi lebih transparan dan dapat dipertanggungjawabkan secara linguistik.
3. Memberikan kontribusi literatur akademis dan referensi komprehensif bagi sivitas akademika Program Studi Informatika Universitas Multimedia Nusantara, khususnya untuk penelitian lanjutan yang berfokus pada alur pemrosesan (*pipeline*) data media sosial dan implementasi *Explainable AI* pada ranah *Natural Language Processing* (NLP).

1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini disusun berdasarkan struktur berikut:

- Bab 1 PENDAHULUAN

Bab ini menguraikan latar belakang masalah mengenai urgensi analisis emosi publik pada media sosial Threads, khususnya terkait peristiwa Banjir Sumatera. Selain itu, bab ini juga mendefinisikan rumusan masalah, batasan penelitian, tujuan yang ingin dicapai, manfaat penelitian, serta rancangan sistematika penulisan laporan.

- Bab 2 LANDASAN TEORI

Bab ini memaparkan konsep-konsep dan kajian literatur akademik yang menjadi dasar dalam penyelesaian masalah. Topik utama yang meliputi *Natural Language Processing* (NLP), arsitektur model Transformer, *Explainable AI* (khususnya SHAP), teori delapan emosi dasar Dr. Robert Plutchik yang menjadi landasan bagi sistem pelabelan leksikon (NRClex/EMOLEX), serta definisi dari model yang digunakan dalam penelitian.

- Bab 3 METODOLOGI PENELITIAN

Bab ini menjabarkan kerangka kerja dan tahapan komputasi sistematis yang diterapkan dalam penelitian. Uraian mencakup skenario pengumpulan data mentah (*web crawling*), alur *text preprocessing*, proses pelabelan data multikelas, tahapan *fine-tuning* model, hingga evaluasi dan interpretasi model.

- Bab 4 HASIL DAN DISKUSI

Bab ini mendokumentasikan hasil dari implementasi kedua model. Diskusi difokuskan pada perbandingan metrik kinerja klasifikasi untuk kedelapan kelas emosi dasar dan interpretasi kedua model secara global dan lokal menggunakan SHAP.

- Bab 5 SIMPULAN DAN SARAN

Bab ini menyajikan seluruh temuan pengujian untuk menjawab rumusan masalah yang telah ditetapkan. Pada bab ini juga disajikan rekomendasi dan saran pengembangan bagi penelitian selanjutnya terkait klasifikasi emosi pada data teks dan *Explainable AI*