BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan ilmu komputer dalam beberapa dekade terakhir telah membawa perubahan besar dalam cara masyarakat menganalisis dan memahami data, terutama data berbasis teks. Salah satu bidang yang mengalami kemajuan signifikan adalah *Natural Language Processing* (*NLP*), yang memungkinkan komputer untuk memahami, menginterpretasikan, dan memanipulasi bahasa manusia. Salah satu aplikasi *NLP* yang semakin populer adalah analisis emosi (*emotion analysis*), yang bertujuan untuk mengidentifikasi dan mengklasifikasikan emosi yang terkandung dalam suatu teks. Selain itu, klasifikasi emosi pada data teks pidato politik juga memiliki relevansi yang tinggi dalam konteks sosial dan politik. Pidato politik seringkali digunakan sebagai alat untuk mempengaruhi pendengar, menyampaikan pesan, dan membangkitkan reaksi tertentu (Makuta, 2025). Dengan memahami emosi yang disampaikan dalam pidato politik, politisi dapat mengidentifikasi strategi komunikasi dan mengevaluasi dampaknya terhadap masyarakat. Oleh karena itu, klasifikasi emosi pada data teks pidato politik tidak hanya memiliki nilai akademis, tetapi juga memiliki implikasi praktis yang penting dalam memahami dinamika sosial dan politik.

Meskipun perkembangan *NLP* mengalami kemajuan untuk analisis emosi ternyata pendekatan tradisional berbasis leksikon tidak cukup mampu menangkap kompleksitas emosi dalam konteks pidato. Menurut Saif M. Mohammad pendekatan berbasis kamus (*lexicon-based-approach*) memiliki keterbatasan dalam memahami makna emosional yang bergantung pada konteks dan makna implisit suatu kalimat dan juga menekankan bahwa penggunaan leksikon harus dilakukan dengan hati-hati karena asosiasi emosi dalam leksikon umumnya didasarkan pada makna dominan kata dan persepsi mayoritas (Mohammad, 2020).

Penelitian oleh Adoma (2020) menunjukkan bahwa *RoBERTa* memberikan performa terbaik dibandingkan model lain seperti *BERT*, *XLNet*, dan *DistilBERT* dalam tugas klasifikasi emosi berbasis teks menggunakan dataset ISEAR berbahasa Inggris dan hasilnya *RoBERTa* menunjukkan kinerja terbaik dalam tugas pengenalan emosi berbasis teks dibandingkan model lain seperti *BERT*, *XLNet* dengan akurasi tertinggi sebesar 74.31% pada dataset ISEAR, studi tersebut menyimpulkan bahwa *RoBERTa* adalah kandidat optimal untuk klasifikasi emosi.

Penelitian tentang klasifikasi emosi berbasis teks dalam bahasa Indonesia masih relatif terbatas dibandingkan dengan bahasa Inggris. Penelitian yang dilakukan oleh Basbeth, (2024) ini membandingkan performa tiga model transformer populer yaitu *BERT*, *RoBERTa,* dan *DistilBERT*. Ketiganya diterapkan pada dataset berbahasa Indonesia (indo4B) dan diuji menggunakan pendekatan *fine-tuning* dan *hyperparameter-tuning* yang menunjukan bahwa *RoBERTa* memiliki performa terbaik dengan akurasi sebesar 90,83% dan *F1-score* mencapai 91%, mengungguli model *BERT F1-score* 90% dan *DistilBERT F1-score* 89%.

Meskipun penelitian yang dilakukan oleh Basbeth, (2024) membuktikan bahwa keunggulan *RoBERTa* dalam klasifikasi emosi teks berbahasa Indonesia secara umum tetapi pada penerapan seperti teks politik terutama pidato presiden masih jarang dilakukan. Oleh karena itu penelitian ini memanfaatkan metode *RoBERTa* sebagai pendekatan modern untuk menjawab kebutuhan klasifikasi emosi yang kontekstual dan akurat, khususnya dalam bahasa Indonesia yang digunakan untuk mengevaluasi muatan emosional dalam pidato Presiden Joko Widodo sepanjang dua periode kepemimpinannya.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah bagaimana hasil pengukuran kinerja metode *RoBERTa* dalam klasifikasi emosi data teks pidato politik berdasarkan akurasi, presisi,*recall* dan *F1-score*?

1.3 Tujuan Penelitian

Tujuan dalam penelitian ini adalah mengetahui hasil pengukuran kinerja metode *RoBERTa* dalam membangun klasifikasi emosi data teks pidato politik berdasarkan akurasi, presisi, *recall* dan *F1-score*?

1.4 Manfaat Penelitian

Penelitian ini dilakukan untuk mendalami klasifikasi emosi pada data teks pidato politik menggunakan bahasa Indonesia dengan pendekatan NLP modern dan metode berbasis *transformer* khususnya *RoBERTa,* penelitian ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Memberikan kontribusi dalam pengembangan metode klasifikasi emosi menggunakan model *NLP* berbasis *transformer*, khususnya *RoBERTa*.

2. Menambah literatur tentang *NLP* dalam analisis teks politik, terutama dalam konteks bahasa Indonesia.

1.5 Batasan Penelitian

Penelitian ini difokuskan pada klasifikasi emosi pada data teks pidato politik presiden Joko Widodo selama masa jabatannya. Adapun ruang lingkup data penelitian ini yaitu :

1. Teks hanya pidato politik yang disampaikan oleh presiden Jokowi dari tahun 2014-2024 tentang rapat terbatas.
2. Dataset yang digunakan pada penelitian ini adalah data yang diambil dari konversi video ke data teks pidato politik di youtube yang disampaikan oleh presiden Jokowi dengan memiliki 5 label yaitu *anger*, *fear*, *happy*, *love* dan *sadness*.
3. Dalam penelitian ini *labeling* dilakukan secara otomatis menggunakan *hugging fac*e yaitu *StevenLimcorn/Indonesian-roberta-base-emotion-classifier*, dan juga *pipeline*.
4. Seluruh proses dalam penelitian ini dilakukan menggunakan bahasa pemrograman *python.*

Jumlah dataset yang digunakan adalah 2952 paragraf