

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi informasi dan komunikasi di era digital telah memfasilitasi penyebaran dan konsumsi informasi secara instan melalui internet dan platform media sosial [1, 2]. Dalam sektor politik, pemanfaatan media sosial untuk menyebarkan opini dan berita telah menciptakan tantangan baru, terutama terkait persebaran konten yang mengandung bias atau ketidakakuratan. Kondisi ini berpotensi memengaruhi opini publik dan stabilitas demokrasi [3]. Studi di Indonesia menunjukkan bahwa partisipasi politik di media sosial semakin meningkat dan memberikan dampak signifikan terhadap proses kebijakan pemerintah [4, 5, 6].

Dalam konteks analisis sentimen politik di Indonesia, metode konvensional seperti *Support Vector Machine* dan *Naive Bayes* yang menggunakan teknik ekstraksi fitur berbasis *TF-IDF* atau *CountVectorizer* menghadapi keterbatasan signifikan. Keterbatasan ini terutama muncul dalam menangkap konteks linguistik yang kompleks pada penggunaan bahasa informal, campuran bahasa, dan istilah khusus politik yang dominan di platform media sosial [7]. Hal ini menunjukkan kebutuhan akan pendekatan yang lebih canggih untuk analisis sentimen.

Platform X (sebelumnya Twitter) telah menjadi arena sentral pertukaran gagasan politik di Indonesia menjelang Pemilu 2024, dengan karakteristik unik berupa pesan singkat, penggunaan hashtag, dan viralitas yang tinggi [8]. Data menunjukkan bahwa lebih dari 70% diskusi politik online terjadi di platform ini, dengan volume mencapai jutaan tweet per hari selama masa kampanye [9]. Dinamika komunikasi politik pada Platform X telah memunculkan polarisasi dan fragmentasi opini publik yang kompleks, di mana kandidat dan pendukung aktif menyebarkan kampanye dan mengkritisi lawan politik mereka [10]. Kompleksitas ini menjadi landasan utama penelitian untuk mengembangkan sistem analisis sentimen yang lebih akurat.

Pendekatan berbasis *Bidirectional Encoder Representations from Transformers (BERT)*, khususnya *IndoBERT*, menawarkan solusi dengan kemampuannya menghasilkan representasi teks yang kontekstual dan mendalam [11]. Dalam konteks bahasa Indonesia, *IndoBERT* telah menunjukkan keunggulan

dalam klasifikasi teks politik dengan akurasi mencapai 87-92%, jauh melampaui metode konvensional yang rata-rata hanya mencapai 75% [12, 13, 14]. Model ini mampu memahami nuansa linguistik lokal, termasuk slang politik, sarkasme, dan kombinasi bahasa yang sering muncul dalam diskusi politik di Platform X [15]. Keunggulan teknis ini menjadikan *IndoBERT* sebagai kandidat utama untuk implementasi analisis sentimen politik yang akurat.

Meskipun demikian, tantangan fundamental dalam implementasi model AI untuk analisis politik adalah kurangnya transparansi dalam proses pengambilan keputusan. Transparansi menjadi sangat penting mengingat sensitivitas isu politik dan potensi bias algoritmik yang dapat memengaruhi hasil analisis. Tanpa transparansi yang memadai, kepercayaan terhadap hasil analisis sentimen politik sulit dibangun, terutama di kalangan pemangku kebijakan dan masyarakat umum.

Pendekatan *generative AI* menawarkan solusi inovatif dengan kemampuannya menyediakan penjelasan naratif berdasarkan *attention weights* dan fitur-fitur bahasa yang menonjol dalam teks [16, 17, 18, 19]. Berbeda dengan metode penjelasan yang kompleks, pendekatan naratif lebih mudah dipahami oleh pengguna non-teknis, seperti analis politik dan pembuat kebijakan. Penelitian terbaru menunjukkan bahwa penjelasan naratif dapat meningkatkan tingkat kepercayaan terhadap hasil analisis hingga 40% dibandingkan sistem tanpa penjelasan [20]. Kemampuan ini menjembatani kesenjangan antara kompleksitas teknis model AI dan kebutuhan praktis pemangku kepentingan.

Berdasarkan tinjauan menyeluruh terhadap permasalahan dan peluang teknologi, penelitian ini merumuskan penggunaan model *IndoBERT* yang telah di-*fine-tune* khusus pada data berbahasa Indonesia di Platform X, dipadukan dengan modul *generative AI* untuk menyajikan penjelasan naratif yang komprehensif. Pendekatan terpadu ini diharapkan mampu meningkatkan akurasi analisis sentimen dalam diskusi politik pra-pemilu di Platform X. Peningkatan akurasi ini dicapai melalui kemampuan *IndoBERT* dalam memahami konteks linguistik yang kompleks, yang krusial untuk mengklasifikasikan sentimen dalam teks politik yang sering kali mengandung bahasa informal, slang, dan nuansa khusus. Proses *fine-tuning* pada dataset tweet politik berbahasa Indonesia memungkinkan model untuk mengenali pola dan fitur khusus domain, sehingga meningkatkan presisi klasifikasi sentimen dibandingkan metode konvensional. Sementara itu, modul *generative AI* menyediakan penjelasan naratif yang komprehensif atas prediksi model, memastikan transparansi dan kemudahan interpretasi hasil analisis bagi pengguna non-teknis. Dengan demikian, pendekatan ini tidak hanya meningkatkan

akurasi, tetapi juga memberikan transparansi melalui penjelasan yang mudah dipahami, serta menghadirkan wawasan yang lebih tajam mengenai persepsi publik terhadap isu politik terkini. Kontribusi utama penelitian ini terletak pada pengembangan kerangka kerja analisis sentimen yang unggul dari segi akurasi, transparan, dan dapat dipertanggungjawabkan dalam konteks analisis politik di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian ini diarahkan untuk menjawab permasalahan sebagai berikut:

1. Bagaimana melakukan analisis sentimen menggunakan model *IndoBERT*?
2. Bagaimana *generative AI* dapat diintegrasikan untuk menghasilkan penjelasan naratif atas hasil analisis sentimen tersebut?

1.3 Batasan Permasalahan

Bagian ini menyajikan batasan-batasan yang diterapkan dalam penelitian agar fokus dan ruang lingkup penelitian lebih terarah. Penelitian ini memiliki batasan sebagai berikut:

1. Data yang dianalisis hanya berasal dari Platform X, dengan periode pengumpulan data terbatas pada Oktober 2022 hingga April 2023, sehingga hasil penelitian mencerminkan kondisi sentimen publik pada platform tersebut dalam rentang waktu tersebut.
2. Penelitian hanya fokus pada implementasi model *IndoBERT* yang telah di-*fine-tune* serta integrasi *generative AI* untuk penjelasan naratif, tanpa mempertimbangkan model atau pendekatan alternatif lainnya.
3. Analisis sentimen dibatasi pada klasifikasi biner, yaitu sentimen positif dan negatif, tanpa memasukkan kategori netral atau nuansa sentimen lainnya.
4. Penggunaan model *BERT* terbatas pada varian *IndoBERT*, yang dioptimalkan untuk bahasa Indonesia, tanpa eksplorasi terhadap varian *BERT* lainnya.
5. Akurasi analisis sentimen dapat dipengaruhi oleh keterbatasan ukuran dataset yang digunakan (10.000 tweet per kandidat) serta kompleksitas bahasa

informal dan slang dalam tweet, yang mungkin tidak sepenuhnya ditangkap oleh model *IndoBERT* meskipun telah di-*fine-tune*.

1.4 Tujuan Penelitian

Berdasarkan permasalahan yang telah diuraikan, penelitian ini memiliki tujuan sebagai berikut:

1. Menerapkan model *IndoBERT* untuk melakukan analisis sentimen terhadap dataset tweet terkait calon presiden Indonesia dari platform media sosial X (Twitter) menjelang pemilu 2024.
2. Mengevaluasi kinerja sistem analisis sentimen menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk memastikan kehandalan model dalam mengklasifikasikan sentimen publik terhadap kandidat presiden.
3. Mengintegrasikan pendekatan *generative AI* untuk menghasilkan penjelasan naratif yang transparan mengenai basis pengambilan keputusan model, sehingga meningkatkan interpretabilitas hasil analisis sentimen politik.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Kontribusi Keilmuan

Membuka peluang pengembangan lebih lanjut di bidang *Natural Language Processing* (NLP), terutama terkait pemanfaatan *IndoBERT* yang telah di-*fine-tune* dan pendekatan *generative AI* untuk analisis sentimen politik.

2. Pengembangan Metodologi

Menyediakan kerangka analisis yang dapat diadaptasi untuk berbagai kebutuhan riset, dengan meliputi proses pengumpulan data, *preprocessing*, pelatihan model, serta validasi kinerja menggunakan metrik terstandar.

3. Aplikasi di Dunia Praktis

Menyediakan landasan bagi pengembangan sistem pemantauan opini publik dalam lingkup politik atau isu lain yang bersifat dinamis. Sistem hasil penelitian ini dapat dikembangkan lebih lanjut untuk mendukung proses penyusunan strategi kampanye, pemetaan pendapat masyarakat, atau pemantauan kebijakan.

4. Peningkatan Akurasi dan Efektivitas Analisis

Meningkatkan keandalan proses pengambilan keputusan berbasis data, melalui model yang mampu menangkap konteks linguistik secara mendalam serta mengolah volume data yang besar dan bervariasi.

1.6 Sistematika Penulisan

Laporan penelitian ini disusun dalam lima bab, yang masing-masing membahas aspek penting dari penelitian, sebagai berikut:

- **BAB 1 PENDAHULUAN**

Menguraikan konteks dan latar belakang penelitian mengenai analisis sentimen opini publik di media sosial terkait kandidat presiden dalam Pemilu 2024 menggunakan model *IndoBERT* serta integrasi *generative AI* untuk menghasilkan penjelasan naratif yang interpretatif. Bab ini juga mencakup rumusan masalah, batasan penelitian, tujuan, dan manfaat penelitian bagi analisis sentimen politik.

- **BAB 2 LANDASAN TEORI**

Menyajikan tinjauan pustaka tentang teori-teori yang menjadi dasar penelitian, meliputi analisis sentimen, karakteristik Platform X sebagai sumber data, konsep *transformer*, pengembangan dan penerapan model *IndoBERT*, serta peran *generative AI* dalam meningkatkan interpretabilitas hasil analisis sentimen.

- **BAB 3 METODOLOGI DAN PERANCANGAN SISTEM**

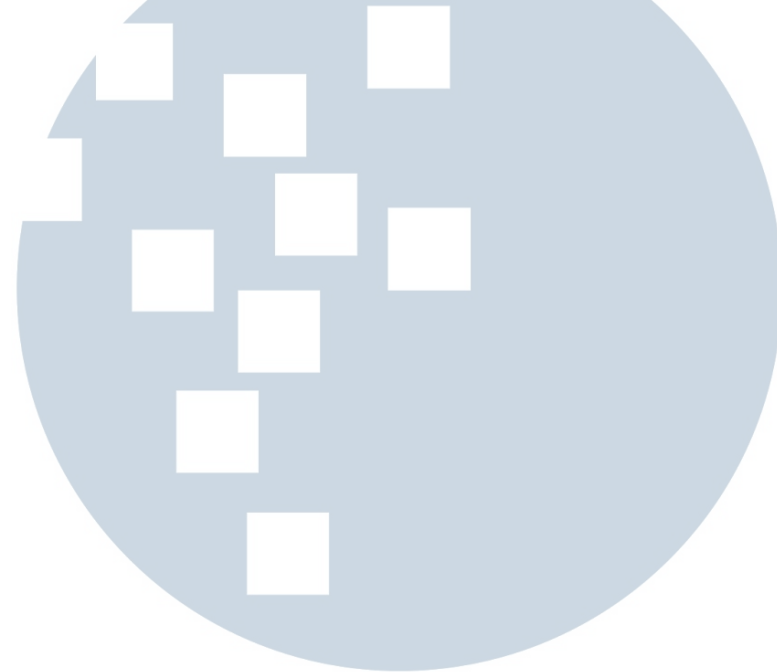
Menjelaskan metodologi penelitian, desain sistem, dan alur implementasi yang mencakup pengumpulan data tweet dari Platform X, tahapan *preprocessing*, *fine-tuning* model *IndoBERT* untuk klasifikasi sentimen, serta integrasi *generative AI* untuk menghasilkan narasi yang menjelaskan output model secara kontekstual.

- **BAB 4 HASIL DAN DISKUSI**

Menyajikan hasil eksperimen, termasuk implementasi dan evaluasi performa model *IndoBERT* menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, distribusi sentimen kandidat presiden, serta penjelasan naratif berbasis *generative AI*. Bab ini juga membahas temuan penelitian secara mendalam dan implikasinya dalam konteks politik Indonesia.

- **BAB 5 KESIMPULAN DAN SARAN**

Merangkum kesimpulan berdasarkan hasil analisis sentimen dan evaluasi sistem, serta memberikan saran untuk pengembangan sistem analisis sentimen dan penelitian lanjutan di masa depan.



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA