

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pasca gelombang demonstrasi yang berlangsung pada akhir Agustus 2025 di berbagai wilayah Indonesia, muncul sebuah gerakan tuntutan signifikan yang dikenal sebagai “17+8 Tuntutan Rakyat” [1]. Gerakan ini merangkum 17 tuntutan jangka pendek yang harus diselesaikan dalam satu minggu, ditambah 8 tuntutan jangka panjang dengan tenggat waktu satu tahun. Rangkaian tuntutan ini menyentuh berbagai isu politik dan kebijakan publik yang memiliki dimensi sosial, seperti reformasi Dewan Perwakilan Rakyat (DPR), revisi undang-undang perpajakan, penarikan militer ke barak, penguatan hak asasi manusia (HAM), hingga undang-undang perampasan aset koruptor. Media massa dan sejumlah tokoh publik berperan aktif dalam menyuarakan dan menyebarluaskan tuntutan 17+8 ini melalui berbagai platform digital, menandai tuntutan ini sebagai isu yang menarik perhatian publik secara luas [2]. Fenomena percakapan daring ini memberikan peluang untuk memahami opini publik menggunakan pendekatan komputasional, khususnya analisis sentimen berbasis *Natural Language Processing* (NLP) dengan memetakan respons masyarakat secara sistematis.

Platform X (Twitter) menjadi media sosial utama yang memfasilitasi ekspresi aspirasi publik, menurut laporan Digital 2025 merupakan salah satu platform yang paling banyak digunakan di Indonesia [3]. Platform ini memungkinkan pengguna untuk berbagi opini tekstual dengan cepat, menjadikan komentar dan cuitan warga sebagai representasi sentimen publik yang autentik. Teks yang dihasilkan di media sosial selanjutnya disebut “teks autentik”, umumnya bersifat tidak terstruktur, informal, dan mengandung bahasa non-standar, termasuk singkatan, *slang*, dan emotikon. Dalam studi analisis sentimen konvensional, teks jenis ini sering kali melalui *preprocessing* yang agresif seperti penghapusan *stopword*, *stemming*, dan normalisasi berlebihan. Praktik ini, meskipun bertujuan meringankan komputasi, berisiko menghilangkan makna atau nuansa sentimen tertentu dari pesan aslinya [4], sehingga mempengaruhi interpretasi sentimen yang sebenarnya. Kondisi ini menjadikan data teks autentik sebagai tantangan sekaligus peluang penting dalam klasifikasi sentimen, yaitu bagaimana mempertahankan integritas semantik dari pesan publik yang sebenarnya.

Untuk memahami pandangan masyarakat terhadap isu politik dan kebijakan publik, analisis sentimen sebagai bagian dari NLP telah menjadi pendekatan yang banyak diterapkan. Berbagai penelitian terdahulu telah berhasil menerapkan metode ini pada isu-isu di Indonesia. Sebagai contoh, sebuah studi mengenai opini masyarakat terhadap pemilihan presiden Indonesia menemukan bahwa *Bidirectional Long Short-Term Memory* (BiLSTM) menghasilkan performa terbaik dibandingkan LSTM dan CNN serta beberapa algoritma pembelajaran mesin tradisional lainnya dalam tugas klasifikasi sentimen politik [5]. Di sisi lain, model berbasis *Transformer*, terutama IndoBERT yang dilatih khusus pada korpus Bahasa Indonesia, juga terbukti lebih efektif dalam menangkap konteks kebahasaan yang kompleks. Hal ini ditunjukkan melalui penelitian analisis sentimen terhadap sebuah kebijakan, di mana IndoBERT secara konsisten mengungguli metode klasik dalam hal akurasi [6]. Lebih jauh, penggabungan kedua pendekatan tersebut dalam bentuk model hibrida IndoBERT-BiLSTM telah menghasilkan kinerja yang lebih kuat dan efektif, sebagaimana ditunjukkan pada studi terkait klasifikasi emosi dan sentimen umum dalam Bahasa Indonesia [7] dan juga klasifikasi sentimen ulasan TikTok, yang mampu menghasilkan performa lebih baik dibandingkan penggunaan model tunggal, dengan akurasi mencapai 81% pada *classification report* dan *average accuracy* 92.03% pada pengujian *cross-validation* [8]. Temuan-temuan ini menunjukkan bahwa pendekatan hibrida memiliki potensi kuat untuk diterapkan pada isu-isu yang lebih kompleks. Namun, penerapannya pada analisis isu politik dan kebijakan publik yang bersifat dinamis masih sangat terbatas, sehingga membuka peluang penting bagi penelitian ini.

Analisis opini publik di media sosial dapat memberikan wawasan kualitatif *real-time* mengenai persepsi dan tingkat kepuasan warga, yang berfungsi sebagai masukan kritis bagi pemangku kepentingan untuk mengevaluasi efektivitas kebijakan [9] dan mendukung keputusan [10] serta mengukur arah mobilisasi politik. Hingga saat ini, belum terdapat studi yang secara spesifik mengkaji sentimen publik terhadap gerakan “17+8 Tuntutan Rakyat”. Kesenjangan ini dapat dipahami karena fenomena tersebut tergolong baru dan spesifik, yang muncul sebagai kristalisasi dari respons spontan masyarakat pasca aksi demonstrasi pada akhir Agustus 2025 [1], [2], di mana gerakan tuntutan itu sendiri merupakan formalisasi dari ketidakpuasan publik terhadap dinamika sosial dan politik yang lebih luas. Sementara, penerapan analisis sentimen untuk menginvestigasi gerakan sosial dan politik [11] seperti “17+8 Tuntutan Rakyat” yang merupakan artikulasi langsung dari aspirasi publik terhadap kebijakan, memiliki urgensi yang krusial.

Berdasarkan hal tersebut, penelitian ini mengisi tiga celah penelitian (*research gap*) utama. Pertama, dari sisi topik: Berdasarkan penelusuran literatur (Google Scholar, Scopus, dan repositori jurnal nasional hingga November 2025) belum terdapat studi yang secara khusus mengembangkan model untuk menganalisis sentimen publik terhadap gerakan “17+8 Tuntutan Rakyat”, padahal isu seperti ini merupakan barometer penting untuk memahami aspirasi publik terhadap kinerja pemerintahan dan arah demokrasi nasional. Penelitian analisis sentimen di Indonesia umumnya berfokus pada isu yang relatif netral seperti ulasan produk dan layanan digital [12] atau kebijakan ekonomi [13]. Kedua, dari sisi metodologis: walaupun arsitektur hibrida *Transformer-Sequential* telah terbukti mencapai kinerja *state-of-the-art* dalam berbagai tugas klasifikasi teks [14], penerapannya untuk analisis sentimen pada isu Politik dan Kebijakan Publik berbahasa Indonesia yang kompleks masih relatif terbatas. Ketiga, dari sisi pengolahan data: sebagian besar penelitian terdahulu bahkan yang terbaru, masih menerapkan *preprocessing* agresif (*aggressive preprocessing*) seperti *stopword elimination* dan *stemming* dan normalisasi yang berlebihan [15], [16]. Pendekatan ini berisiko menghilangkan integritas semantik (*semantic meaning*) [17] dari teks autentik. Untuk itu, penelitian ini mengusulkan pendekatan *preprocessing* ringan guna mempertahankan keaslian semantik. Pendekatan ini didukung oleh studi empiris pada Bahasa Indonesia yang menyimpulkan bahwa *stemming* dan *stopword removal* tidak memberikan peningkatan akurasi yang signifikan dan justru berpotensi menurunkan kinerja [18]. Dengan dukungan model hibrida IndoBERT-BiLSTM, pendekatan ini diharapkan menghasilkan model analisis sentimen yang lebih *robust* dan kontekstual terhadap karakter alami bahasa publik di X.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang masalah sebelumnya, maka dapat diambil rumusan masalah yaitu:

1. Dinamika opini publik terkait 17+8 Tuntutan Rakyat memunculkan respons masif berupa teks autentik yang tidak terstruktur di media sosial X, sehingga menyulitkan pemetaan arah dukungan atau penolakan masyarakat secara sistematis.
2. Kebutuhan untuk mengidentifikasi dan mengklasifikasikan respons publik tersebut secara efektif guna memahami kecenderungan masyarakat terhadap tuntutan yang diajukan.

### 1.3 Tujuan

Tujuan dari penulisan tugas akhir ini adalah:

1. Mengembangkan model hibrida IndoBERT dan BiLSTM melalui proses *fine-tuning* untuk mengklasifikasikan sentimen teks autentik publik di platform X (Twitter) terhadap 17+8 Tuntutan Rakyat ke dalam tiga kelas: positif, negatif, dan netral.
2. Menganalisis distribusi dan kecenderungan sentimen publik terhadap 17+8 Tuntutan Rakyat berdasarkan hasil klasifikasi model yang dikembangkan, sehingga dapat menggambarkan respons masyarakat terhadap tuntutan yang diajukan.

### 1.4 Manfaat

1. Bagi peneliti dan akademisi, penelitian ini memberikan referensi empiris mengenai efektivitas pendekatan *preprocessing* ringan pada teks autentik media sosial berbahasa Indonesia, khususnya dalam menghindari penghapusan makna akibat teknik agresif seperti *stopword removal*, *stemming*, dan *lemmatization* yang berlebihan.
2. Mengevaluasi kinerja model dalam membandingkan performa terhadap pendekatan IndoBERT, BiLSTM dan hibrida IndoBERT-BiLSTM dengan data yang sama.
3. Bagi komunitas riset *deep learning* dan NLP, penelitian ini menyediakan perbandingan kinerja yang terukur antara pendekatan IndoBERT, BiLSTM, dan model hibrida IndoBERT-BiLSTM pada *dataset* yang sama, sehingga dapat menjadi acuan pemilihan arsitektur untuk tugas serupa.
4. Bagi masyarakat umum dan pemangku kepentingan publik, penelitian ini menghasilkan analisis sentimen yang menggambarkan distribusi tanggapan masyarakat terhadap 17+8 Tuntutan Rakyat secara kuantitatif, sehingga dapat digunakan sebagai bahan pemahaman dinamika opini publik terkait isu tersebut.
5. Bagi pengambil kebijakan dan lembaga pemerintah, identifikasi isu atau tuntutan yang paling sering dan intensif dibahas publik di media sosial dapat dijadikan masukan untuk memahami prioritas dan kekhawatiran masyarakat terkait gerakan 17+8 Tuntutan Rakyat.

### 1.5 Ruang Lingkup

1. Menggunakan model *pretrained* indobert-base-p2 sebagai *backbone* pemrosesan bahasa Indonesia.
2. Data yang digunakan berasal dari platform media sosial X (Twitter) dengan kata kunci "tuntutanrakyat17plus8", "17+8 Tuntutan Rakyat", "Tuntutan 17+8" dan

"tuntutanrakyat". Data dikumpulkan melalui proses *scraping* dalam rentang waktu 25 Agustus 2025 hingga 31 Desember 2025.

3. Data diklasifikasikan ke dalam 3 kategori sentimen, yaitu negatif, netral dan positif.
4. Anotasi dilakukan secara manual oleh lebih dari satu anotator untuk meminimalkan bias serta memperoleh label yang lebih akurat.

