BAB I
PENDAHULUAN

1.1 Latar Belakang

Transformasi digital telah secara fundamental mengubah paradigma komunikasi publik, menjadikan media sosial dan secara khusus platform X (sebelumnya Twitter) sebagai sebuah fenomena yang melampaui fungsi awalnya sebagai jaringan sosial. Kini, X telah berevolusi menjadi repositori data publik yang masif sekaligus berfungsi sebagai barometer sosio-politik yang esensial, beroperasi secara *real-time*. Volume data yang dihasilkan oleh platform ini sangat besar; data yang meningkat dari 500 juta *tweet* per hari pada tahun 2013 kini diprediksikan telah melonjak menjadi 644 juta per hari (Asgari-Chenaghlu dkk., 2021). Namun, signifikansi data ini tidak hanya terletak pada jumlahnya yang mentah, tetapi pada fungsinya sebagai cerminan langsung dan sering kali tanpa filter dari perbincangan, sentimen, serta respons kolektif masyarakat terhadap peristiwa dunia nyata. Kombinasi unik antara volume masif dan keragaman topik inilah yang membuat X menjadi sumber data primer tak ternilai bagi penelitian ilmiah yang bertujuan untuk membedah dan memahami dinamika opini publik.

Pada periode Juli–September 2025, Indonesia menyaksikan diskusi publik yang sangat intensif. Gelombang unjuk rasa di berbagai daerah, yang dipicu oleh serangkaian rancangan kebijakan pemerintah mencakup kenaikan PPN 12%, kenaikan tunjangan DPR, *coretax*, penerapan Makan Bergizi Gratis (MBG), hingga isu Uang Kuliah Tunggal (UKT), menghasilkan volume percakapan digital yang masif. Namun, opini publik yang terekspresikan ini tidak bersifat monolitik. Sebuah label sentimen agregat, seperti "80% negatif", menjadi ambigu dan tidak dapat ditindaklanjuti (*non-actionable*). Analisis semacam ini gagal menjawab pertanyaan fundamental: apa yang menyebabkan sentimen negatif tersebut? Apakah publik lebih marah terhadap aspek 'PPN', aspek 'UKT', atau aspek 'Tunjangan DPR'?.

Tantangan utama dalam menganalisis opini publik dari sumber data ini bersifat ganda. Pertama, tantangan ini terletak pada volume dan kompleksitas data itu sendiri, yang menunjukkan karakteristik klasik *big data*. Karakteristik ini

awalnya diformulasikan sebagai model 3-V: *Volume*, yang merujuk pada jumlah data dalam skala sangat besar; *Velocity*, yang mengindikasikan kecepatan tinggi dalam generasi dan transfer data, seperti pada *data streaming* dari Twitter; dan *Variety*, yang menggambarkan keragaman tipe data yang dikumpulkan. Seiring berjalannya waktu, model ini diperluas menjadi 5-V untuk memberikan gambaran yang lebih komprehensif, dengan menambahkan *Value*, yaitu tantangan untuk mengekstraksi informasi yang bermakna dan berharga dari tumpukan data mentah, dan *Veracity*, yang mengacu pada kebenaran dan akurasi data. Kualitas data ini menjadi tantangan kritis tersendiri, misalnya di X, di mana data dapat sangat bervariasi dari cuitan harian berkualitas rendah (*noise*) hingga laporan peristiwa berkualitas tinggi dari akun berita terverifikasi (*signal*). Keseluruhan karakteristik ini menuntut arsitektur pemrosesan *just-in-time* agar data dapat dipantau dan dilacak secara efektif untuk menghasilkan wawasan yang akurat (Asgari-Chenaghlu dkk., 2021).

Tantangan kedua, yang sering kali bersinggungan dengan yang pertama, adalah kompleksitas linguistik data itu sendiri. Dalam konteks Indonesia, data cuitan di platform X sarat dengan teks informal, bahasa gaul (*slang*), dan fenomena campur kode (*code-mixing*), yakni penggunaan dua bahasa atau lebih (misalnya bahasa Indonesia dan bahasa Inggris) secara bergantian dalam satu kalimat atau cuitan. Kecenderungan ini telah menjadi kebiasaan yang lazim dan sering dijumpai dalam berbagai percakapan pada platform X di Indonesia, terutama di kalangan generasi muda (Barik dkk., 2021).

Akibat dari tantangan ganda ini skala *big data* dan kerumitan linguistik adalah analisis manual untuk menyaring jutaan opini ini tidak lagi memadai. Metode manual tidak hanya sangat tidak efisien dan memakan waktu, tetapi juga sangat rentan terhadap bias subjektif dari analis (Murshed dkk., 2022). Oleh karena itu, untuk mengekstrak wawasan yang bermakna dari data tekstual mentah tersebut, penerapan pendekatan komputasional, khususnya pemrosesan bahasa alami, bukan lagi pilihan, melainkan sebuah kebutuhan esensial.

Salah satu aplikasi fundamental NLP dalam konteks ini adalah analisis sentimen, yang bertugas mengekstraksi dan mengklasifikasikan sentimen (positif,

negatif, netral) secara otomatis (Suhaeni dkk., 2025). Dalam konteks bahasa Indonesia, studi literatur yang sudah dilakukan menemukan bahwa salah satu aplikasi analisis sentimen paling populer adalah untuk menganalisis respons publik terhadap layanan dan kebijakan pemerintah. Berbagai studi telah berhasil menerapkan analisis sentimen untuk mengukur sentimen publik Indonesia terhadap kebijakan pemerintah seperti *Omnibus Law* (Sukma dkk., 2020), pemindahan Ibu Kota Nusantara (Azis dkk., 2024), dan Program Makan Bergizi Gratis (Suhaeni dkk., 2025).

Meskipun demikian, analisis sentimen tradisional memiliki keterbatasan yang fundamental, dengan memberikan label sentimen tunggal pada sebuah cuitan, metode ini sering kali terlalu umum dan gagal menangkap granularitas opini yang kompleks. Dalam realitasnya, sebuah cuitan dapat berisi pujian terhadap "aspek A" dari sebuah kebijakan, namun sekaligus berisi kritikan tajam terhadap "aspek B". Analisis sentimen konvensional mungkin akan memberikan label cuitan ini sebagai "netral" atau "campuran", sehingga menghilangkan wawasan penting yang justru paling dibutuhkan oleh pembuat kebijakan (de Araújo dkk., 2024).

Untuk mengatasi keterbatasan ini dan menggali lebih dalam, bidang NLP telah mengembangkan *Aspect-Based Sentiment Analysis* (ABSA). ABSA adalah metode yang lebih mendalam dan canggih. Metode ini tidak hanya mengidentifikasi sentimen keseluruhan, tetapi juga mengidentifikasi aspek atau fitur spesifik dari entitas yang dibahas, serta mengidentifikasi sentimen yang terkait secara partikular dengan setiap aspek tersebut (Perwira dkk., 2025). Sebagai contoh, hasil analisis ABSA dapat memetakan opini publik menjadi wawasan yang jauh lebih tajam, seperti: {Aspek: "Implementasi Kebijakan", Sentimen: "Negatif"} dan {Aspek: "Tujuan Kebijakan", Sentimen: "Positif"}.

Tantangan inti dan langkah paling kritis dalam *pipeline* ABSA adalah langkah ekstraksi aspek, terutama ketika aspek-aspek tersebut tidak diketahui sebelumnya (*unsupervised*). Untuk tugas ini, pemodelan topik telah menjadi solusi standar karena kemampuannya untuk menganalisis korpus teks besar dan menemukan struktur tematik, topik, atau aspek laten di dalamnya (Acheadeth dkk., 2022). Namun, pemilihan model topik menjadi krusial. Model yang paling

fundamental dan diadopsi secara luas adalah *Latent Dirichlet Allocation* (LDA). Namun, berbagai studi secara konsisten menunjukkan bahwa kinerja LDA menurun drastis ketika diterapkan pada teks pendek seperti cuitan di media sosial X (Tamzila dkk., 2025).

Hal ini disebabkan oleh masalah *data sparsity* (kekurangan data) di mana konteks dan kookurensi kata yang minim dalam satu cuitan membuat model probabilistik LDA kesulitan menemukan topik yang koheren secara statistik (Weisser dkk., 2023). Sebagai alternatif yang dirancang khusus untuk mengatasi masalah ini, model seperti *Gibbs Sampling Dirichlet Multinomial Mixture* (GSDMM) dikembangkan untuk klasifikasi teks pendek. GSDMM beroperasi dengan asumsi yang lebih sesuai untuk cuitan, yakni bahwa setiap dokumen pendek cenderung berfokus pada satu topik. Beberapa studi komparatif telah menunjukkan bahwa GSDMM mengungguli LDA dalam tugas klasifikasi teks pendek (Santakij dkk., 2024).

Sementara pendekatan LDA dan GSDMM berupaya untuk menyempurnakan metodologi probabilistik, sebuah revolusi fundamental datang dari model neural atau berbasis *embedding*, dan yang paling menonjol adalah *BERTopic* (Grootendorst, 2022). Tidak seperti LDA atau GSDMM yang mengandalkan frekuensi *bag-of-words*, *BERTopic* mengambil langkah yang berbeda secara fundamental. Model ini memanfaatkan model *transformer* (seperti BERT) untuk membuat *document embeddings* yang menangkap makna semantik teks secara kontekstual. Model ini kemudian menggunakan *pipeline* reduksi dimensi (UMAP) dan *clustering* (HDBSCAN) untuk mengelompokkan dokumen berdasarkan kedekatan semantik, bukan hanya frekuensi kata (Ma dkk., 2025). Sejumlah besar studi komparatif pada data teks pendek, termasuk ulasan pelanggan (Krishnan & Kennedyraj, 2023), komentar YouTube (Nanayakkara & Thennakoon, 2024), keluhan lalu lintas (Tamzila dkk., 2025), dan teks berita (Asas dkk., 2025), telah menunjukkan bahwa *BERTopic* secara konsisten mengungguli model tradisional seperti LDA dalam menghasilkan topik yang lebih koheren dan bermakna. Bahkan, telah ada upaya untuk lebih mengoptimalkan *BERTopic* untuk teks pendek yang ambigu dengan menambahkan modul-modul kustom (Feng dkk.,

2025).

Di sinilah letak kesenjangan penelitian (*research gap*) yang ingin diisi oleh studi ini. Meskipun ada bukti kuat yang menunjukkan bahwa GSDMM mengungguli LDA dalam skenario teks pendek (Santakij dkk., 2024), dan bukti kuat lainnya yang menunjukkan bahwa BERTopic juga mengungguli LDA (Tamzila dkk., 2025), masih terdapat kekurangan dalam evaluasi komparatif yang komprehensif dan langsung yang mempertemukan ketiga model ini (LDA vs. GSDMM vs. BERTopic) secara bersamaan.

Kesenjangan ini menjadi lebih dalam ketika difokuskan pada konteks spesifik, yakni teks bahasa Indonesia yang informal dan penuh campur kode dari media sosial, serta domain opini kebijakan pemerintah yang kompleks dan sarat nuansa. Penelitian ini bertujuan mengisi celah tersebut dengan membandingkan secara langsung kinerja LDA (sebagai *baseline* tradisional), GSDMM (sebagai *baseline* probabilistik teks pendek), dan BERTopic (sebagai *baseline* neural) untuk menentukan model mana yang paling efektif berfungsi sebagai modul ekstraksi aspek dalam *pipeline* ABSA.

Setelah aspek berhasil diekstraksi oleh salah satu dari tiga model tersebut, proses selanjutnya adalah melakukan klasifikasi sentimen pada aspek tersebut secara akurat. Mengingat bahasa media sosial di Indonesia sarat dengan bahasa gaul, salah ketik, dan bahkan sarkasme (Rosid dkk., 2024), di mana makna sangat bergantung pada konteks, model *machine learning* tradisional seperti *Naïve Bayes* atau SVM, meskipun masih umum digunakan dalam penelitian di Indonesia (Azis & Wahyudi, 2024), sering kali kesulitan menangkap nuansa kontekstual yang kompleks ini.

Sebagai terobosan, model *transformer* seperti BERT (*Bidirectional Encoder Representations from Transformers*) telah menjadi *state-of-the-art* dalam NLP (Devlin dkk., 2018). Keunggulan utama BERT adalah kemampuannya memahami konteks secara *bidirectional* (dua arah), memungkinkannya membedakan makna kata berdasarkan keseluruhan kalimat (Sjoraida dkk., 2024). Varian BERT yang dilatih khusus untuk bahasa Indonesia, seperti *IndoBERT* atau *IndoBERTWEET*, telah terbukti mencapai kinerja superior dalam berbagai tugas

klasifikasi teks di Indonesia (Koto dkk., 2021) dan akan dimanfaatkan dalam penelitian ini.

Berdasarkan tantangan dan tinjauan tersebut, penelitian ini mengusulkan sebuah arsitektur *pipeline* ABSA yang mengintegrasikan, pertama, pemodelan topik (baik GSDMM, LDA, maupun *BERTopic*) untuk ekstraksi aspek secara *unsupervised*, dan kedua, model BERT (khususnya *IndoBERTWEET*) untuk klasifikasi sentimen pada aspek-aspek yang telah diekstraksi. Fokus utama dan orisinalitas penelitian ini terletak pada evaluasi komparatif dari modul ekstraksi aspek. Penelitian ini tidak hanya sekadar menerapkan BERT untuk ABSA, tetapi secara spesifik akan mengevaluasi dan membandingkan bagaimana tiga metode ekstraksi aspek yang berbeda secara fundamental, GSDMM yang dioptimalkan untuk teks pendek, LDA sebagai model probabilistik yang sudah umum digunakan, dan *BERTopic* yang memanfaatkan model *transformer* memengaruhi kinerja *end-to-end* dari sistem ABSA berbasis BERT.

Penelitian ini diharapkan dapat memberikan kontribusi teoretis dan praktis. Secara teoretis, penelitian ini akan memberikan wawasan empiris mengenai metode pemodelan topik mana (GSDMM, LDA, atau *BERTopic*) yang lebih superior untuk melakukan klasifikasi teks pendek bahasa Indonesia di media sosial, yang berfungsi sebagai *input* bagi model klasifikasi sentimen *transformer* (BERT). Sementara itu, secara praktis, penelitian ini akan menghasilkan model *ABSA* yang efektif untuk bahasa Indonesia sekaligus menyajikan wawasan *granular* (per aspek) mengenai sentimen publik terhadap "Kebijakan Pemerintah 2025", yang dapat menjadi masukan berbasis data yang berharga bagi pemerintah dan pemangku kepentingan di masa yang akan datang.

1.2 Tujuan Penelitian

Berdasarkan latar belakang masalah yang telah dijabarkan, penelitian ini memiliki beberapa tujuan utama yang ingin dicapai, yaitu:

1. Melakukan ekstraksi aspek (topik) dari data teks media sosial X dengan rentang waktu Januari sampai dengan Desember 2025 menggunakan tiga model pemodelan topik yang berbeda secara fundamental, yakni *Latent Dirichlet Allocation* (LDA), *Gibbs Sampling Dirichlet Multinomial Mixture*

(GSDMM), dan *BERTopic*.

2. Melakukan evaluasi komparatif secara kuantitatif untuk menilai kinerja ketiga model tersebut dalam menangkap makna kontekstual data teks media sosial X.
3. Mengimplementasikan *pipeline* ABSA menjadi sebuah sistem informasi yang mampu menyajikan wawasan secara lebih informatif dan dapat ditindaklanjuti bagi pemangku kebijakan.

1.3 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat teoretis (akademis) dan manfaat praktis bagi pemangku kepentingan terkait. Secara teoretis, penelitian ini memberikan bukti empiris mengenai perbandingan metodologis dalam pemodelan topik untuk teks pendek. Penelitian ini tidak hanya membandingkan model probabilistik tradisional (LDA) dengan model probabilistik khusus teks pendek (GSDMM), tetapi juga mengikutsertakan model neural *state-of-the-art* (*BERTopic*) yang berbasis *embedding*. Secara spesifik, penelitian ini berkontribusi pada pemahaman tentang pendekatan mana (probabilistik-tradisional, probabilistik-teks-pendek, atau *neural-embedding*) yang lebih superior untuk ekstraksi aspek dari teks media sosial berbahasa Indonesia yang *noisy* dan *sparse*.

Selain itu, penelitian ini menyajikan dan mengevaluasi sebuah *pipeline* ABSA terintegrasi yang menggabungkan berbagai pendekatan pemodelan topik (probabilistik dan neural) sebagai pengekstraksi aspek dengan model *transformer* canggih (BERT) sebagai pengklasifikasi sentimen. Kerangka kerja ini dapat menjadi rujukan metodologis untuk penelitian serupa. Terakhir, penelitian ini diharapkan dapat menambah khazanah studi NLP untuk bahasa Indonesia, khususnya pada domain analisis opini kebijakan.