

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital telah mengubah pola komunikasi masyarakat di seluruh dunia, termasuk di Indonesia. Media sosial kini menjadi salah satu ruang publik terbesar yang memfasilitasi pertukaran informasi, ekspresi opini, hingga partisipasi dalam diskusi sosial-politik. Data dari DataReportal menunjukkan bahwa hingga Januari 2025 terdapat sekitar 143 juta pengguna aktif media sosial di Indonesia, yang mencakup 50,2% dari total populasi [1]. Besarnya penetrasi media sosial tersebut menjadikannya medium penting dalam pembentukan opini publik serta dinamika demokrasi digital di Indonesia [2].

Salah satu platform yang berpengaruh dalam pergeseran komunikasi publik adalah X (sebelumnya dikenal sebagai Twitter). Karakter unggahan singkat dan penggunaan tagar (#) memungkinkan isu tertentu menyebar dengan cepat dan menjadi tren percakapan. Fenomena viralnya tagar #kaburajadulu, misalnya, mencerminkan keresahan Masyarakat terutama generasi muda terhadap situasi sosial-politik di Indonesia [3]. Tagar ini tidak hanya merepresentasikan kritik sarkastik terhadap kondisi dalam negeri, tetapi juga menjadi sarana berbagi informasi alternatif, seperti peluang kerja atau pendidikan di luar negeri, yang pada akhirnya berkembang menjadi simbol kekecewaan terhadap masa depan bangsa [4]. Fenomena ini menunjukkan bahwa media sosial berfungsi ganda yaitu sebagai ruang diskusi sekaligus kanal protes sosial.

Meskipun berbagai platform media sosial populer di Indonesia seperti Instagram, Facebook, YouTube, dan TikTok memiliki jumlah pengguna lebih besar, karakteristik konten di platform tersebut cenderung berfokus pada visual berupa foto, video, dan hiburan. Sementara itu, platform X (sebelumnya Twitter) memiliki keunikan sebagai ruang diskusi berbasis teks yang bersifat real-time dan terbuka, sehingga memungkinkan isu sosial-politik menyebar dengan cepat melalui tagar [5]. Sementara itu, laporan *Datareportal* (2025) mencatat bahwa jumlah pengguna X di Indonesia pada awal tahun 2025 mencapai sekitar 25,2 juta orang. Angka ini setara dengan 8,8% dari total populasi Indonesia, atau sekitar 11,2% dari populasi yang berusia 13 tahun ke atas sebagai audiens yang memenuhi syarat

untuk menggunakan platform tersebut. Lebih lanjut, data menunjukkan bahwa 12% dari penduduk dewasa (usia 18 tahun ke atas) di Indonesia menggunakan X, dengan distribusi gender audiens iklan sebesar 43,8% perempuan dan 56,2% laki-laki. Data ini menegaskan bahwa X memainkan peran penting sebagai salah satu medium interaksi digital dan ruang pembentukan opini publik di Indonesia [6].

Pemilihan tagar #Kaburajadulu dalam penelitian ini didasari oleh tingginya perhatian publik terhadap fenomena tersebut, khususnya di kalangan generasi muda. Tagar ini sempat menjadi tren nasional dan viral di X, sehingga menggambarkan isu publik yang aktual sekaligus menyita perhatian luas. Keberadaannya juga unik karena memadukan kritik politik, ekspresi emosional, serta wacana sosial-budaya, sehingga dapat merepresentasikan keresahan masyarakat Indonesia di ruang digital [7]. Selain itu, dibandingkan tagar lain yang cenderung bersifat sementara, #KaburAjaDulu memiliki nilai simbolik yang lebih kuat karena mencerminkan aspirasi kolektif generasi muda terhadap masa depan bangsa [8].

Oleh karena itu, penelitian ini didasari oleh tingginya perhatian publik, khususnya di kalangan generasi muda, dan relevan dijadikan objek kajian. Tagar ini bukan hanya sekadar tren, melainkan wadah bagi masyarakat untuk menyalurkan keresahan, kritik, sekaligus harapan akan kehidupan yang lebih baik, baik di dalam maupun di luar negeri [9]. Penelitian ini berkontribusi dalam memahami bagaimana media sosial membentuk opini dan ekspresi kolektif melalui penggunaan tagar. Fenomena ini penting untuk dikaji karena dapat menjadi sinyal bagi pemerintah mengenai adanya ketidakpuasan publik sekaligus menjadi masukan dalam merumuskan kebijakan dan komunikasi yang lebih peka terhadap kebutuhan masyarakat [10].

Meskipun begitu, menganalisis opini publik di media sosial bukanlah hal yang sederhana. Besarnya volume data, kecepatan penyebaran informasi, serta kompleksitas gaya bahasa mulai dari penggunaan slang, singkatan, hingga ironi menjadikan analisis manual tidak efisien dan rawan bias [9]. Oleh karena itu, pendekatan berbasis *Natural Language Processing* (NLP) menjadi salah satu metode utama dalam mengekstraksi makna dan pola sentimen dari teks media sosial.

Berbagai penelitian terdahulu telah memanfaatkan metode lexicon-based maupun algoritma machine learning konvensional seperti Support Vector Machine (SVM) dan Naïve Bayes. Hasil penelitian menunjukkan bahwa SVM mampu menghasilkan performa cukup baik dalam klasifikasi sentimen bahasa Indonesia. Namun, metode tersebut masih lemah dalam menangkap konteks kalimat, ironi, maupun ambiguitas makna yang kerap muncul di media sosial [13].

Seiring dengan kemajuan teknologi, model berbasis deep learning, khususnya arsitektur transformer, mulai banyak digunakan dalam penelitian NLP. IndoBERT, sebagai salah satu model transformer yang dilatih secara khusus menggunakan korpus berbahasa Indonesia, terbukti memiliki kinerja unggul dibandingkan metode konvensional maupun model multilingual. Penelitian terbaru menunjukkan IndoBERT mampu mencapai akurasi hingga 84% dalam klasifikasi sentimen teks bahasa Indonesia, melampaui performa SVM dan Naïve Bayes [13]. Keunggulan ini menjadikan IndoBERT sebagai model yang relevan untuk diaplikasikan pada analisis opini publik di media sosial.

Berdasarkan uraian tersebut, terdapat dua permasalahan utama yang menjadi dasar penelitian ini. Pertama, fenomena viral tagar #KaburAjaDulu yang merepresentasikan keresahan generasi muda Indonesia belum pernah dikaji secara ilmiah melalui pendekatan analisis sentimen, padahal isu ini penting untuk memahami opini publik di ruang digital. Kedua, metode analisis sentimen konvensional yang banyak digunakan masih terbatas dalam menangkap kompleksitas bahasa di media sosial, seperti ironi, ambiguitas, dan penggunaan bahasa gaul. Oleh karena itu, penelitian ini berfokus pada penerapan model IndoBERT untuk menganalisis sentimen publik terhadap tagar #KaburAjaDulu di platform X, dengan harapan dapat memberikan pemahaman yang lebih akurat dan kontekstual.

1.2 Rumusan Masalah

Perumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana implementasi model IndoBERT dalam analisis sentimen terhadap tweet dengan tagar #kaburajadulu di media sosial X?

2. Bagaimana kinerja model IndoBERT dalam mengklasifikasikan sentimen (positif, negatif) pada tweet bertagat #kaburajadulu?

1.3 Tujuan Penelitian

Adapun tujuan dari pembuatan sistem ini adalah sebagai berikut:

1. Mengimplemntasi model IndoBERT untuk melakukan analisis sentimen terhadap tweet yang mengandung tagar #kaburajadulu di media sosial X.
2. Mengetahui kinerja model IndoBERT dalam mengklasifikasikan sentimen (positif, negatif).

1.4 Manfaat Penelitian

Penulisan tugas akhir ini diharapkan mampu memberikan manfaat sebagai berikut:

1.4.1 Bagi Penulis

Penelitian ini memberikan manfaat bagi penulis, di antaranya:

- a. Memberikan pengalaman dan pengetahuan baru dalam melakukan penelitian ilmiah, khususnya dalam bidang analisis sentimen, pemrosesan bahasa alami (NLP), dan pemanfaatan model transformer-based seperti IndoBERT.
- b. Menghasilkan karya ilmiah yang dapat dijadikan sebagai referensi untuk penelitian selanjutnya terkait analisis sentimen isu sosial di media sosial.

1.4.2 Bagi Masyarakat

Penelitian ini memberikan manfaat bagi masyarakat atau pengguna, di antaranya:

- a. Memberikan wawasan mengenai persepsi publik terhadap isu sosial viral, khususnya tagar #Kaburajadulu, berdasarkan analisis data media sosial X.
- b. Membantu masyarakat memahami tren opini publik serta faktor-faktor yang memengaruhi sentimen dalam diskursus sosial di ruang digital.
- c. Menyediakan informasi berbasis data yang dapat dimanfaatkan oleh masyarakat, media, maupun pembuat kebijakan untuk merancang strategi komunikasi dan respons sosial yang lebih efektif.

1.4.3 Bagi Akademik

Penelitian ini memberikan manfaat bagi akademik atau lingkungan pendidikan, di antaranya:

- a. Memberikan kontribusi terhadap pengembangan ilmu pengetahuan dan teknologi, khususnya dalam bidang analisis sentimen berbasis deep learning dan NLP untuk bahasa Indonesia.
- b. Menambah referensi ilmiah yang relevan mengenai penerapan IndoBERT dalam analisis sentimen isu sosial di media sosial.
- c. Menjadi acuan dan inspirasi bagi peneliti lain yang akan melakukan penelitian sejenis atau lanjutan terkait analisis sentimen, isu sosial digital, maupun pengaruh media sosial dalam membentuk persepsi publik

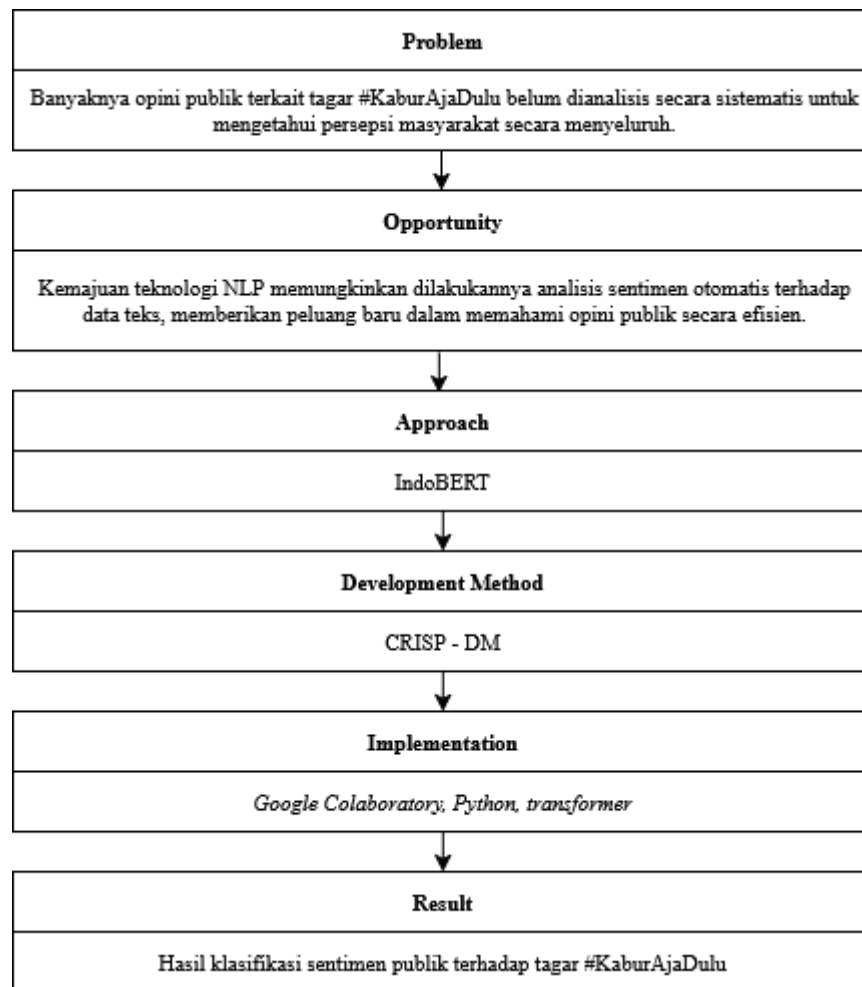
1.5 Batasan Masalah

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang diinginkan maka berikut adalah batasan masalah yang ditetapkan:

1. Data yang dianalisis terbatas pada tweet berbahasa Indonesia yang mengandung tagar #kaburajadulu, yang dikumpulkan dari media sosial X (Twitter) dalam rentang waktu 11 Februari hingga 31 Maret 2025.
2. Data bersumber hanya dari media sosial X.
3. Klasifikasi sentimen dibatasi pada dua label, yaitu positif dan negatif.
4. Model yang digunakan dalam penelitian ini adalah IndoBERT.

1.6 Kerangka Pemikiran

Kerangka pemikiran penelitian ini disusun untuk menggambarkan alur logis dari proses penelitian, mulai dari identifikasi masalah hingga diperolehnya hasil akhir. Dengan adanya kerangka pemikiran, penelitian dapat dijalankan secara sistematis, terarah, dan sesuai tujuan. Dalam penelitian ini, fokus utama adalah menganalisis sentimen publik terhadap tagar #Kaburajadulu di media sosial X (Twitter) menggunakan pendekatan berbasis *Natural Language Processing* (NLP).



Gambar 1. 1 Kerangka Pemikiran

Gambar 1.1 menunjukkan kerangka pemikiran pada penelitian ini. Kerangka pemikiran tersebut membantu pola kerja penelitian yang sistematis dengan memetakan alur logis dari permasalahan hingga hasil yang diharapkan. Kerangka penelitian dimulai dari identifikasi masalah, yaitu belum adanya analisis sistematis terhadap opini publik yang berkembang mengenai tagar #kaburajadulu. Permasalahan ini membuka sebuah *opportunity* dengan adanya kemajuan teknologi *Natural Language Processing* (NLP) yang memungkinkan dilakukannya analisis sentimen secara otomatis terhadap data teks dalam skala besar.

Untuk menangkap peluang tersebut, penelitian ini menggunakan pendekatan berbasis model IndoBERT, yaitu model bahasa berbasis *transformer* yang telah dilatih khusus untuk Bahasa Indonesia. Pendekatan ini dipadukan dengan metode pengembangan CRISP-DM (*Cross Industry Standard Process for Data Mining*) sebagai kerangka kerja dalam menyusun dan menjalankan tahapan penelitian.

Seluruh proses implementasi dilakukan menggunakan *tools* dan teknologi seperti *Google Colaboratory*, *Python*, dan pustaka *transformer* dari *HuggingFace*. Tahapan akhir dari kerangka pemikiran ini adalah hasil klasifikasi sentimen.

1.7 Sistematika Penulisan

Sistematika penulisan menyajikan deskripsi dan urutan masing-masing bab yang disusun secara sistematis untuk memberikan gambaran keseluruhan dari penelitian ini. Penulisan tugas akhir ini terdiri dari lima bab utama, yaitu:

BAB I PENDAHULUAN

Bab ini memuat gambaran umum dari penelitian, yang mencakup latar belakang, perumusan masalah, tujuan dan manfaat penelitian, ruang lingkup dan batasan penelitian, kerangka pemikiran, serta sistematika penulisan yang digunakan dalam tugas akhir ini.

BAB II KAJIAN LITERATUR

Bab ini membahas landasan teori yang relevan dengan penelitian, seperti teori mengenai analisis sentimen, *Natural Language Processing* (NLP), model IndoBERT, serta metode CRISP-DM. Selain itu, kajian terhadap penelitian-penelitian terdahulu yang memiliki keterkaitan juga disertakan sebagai acuan.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, yaitu pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*). Tahapan yang dibahas meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan deployment. Pada tahap ini juga dijelaskan metode pengumpulan data melalui *tweet-harvest* dan proses pra-pemrosesan teks.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan implementasi model analisis sentimen menggunakan IndoBERT. Dibahas pula proses pelatihan model, parameter yang digunakan, serta penggunaan *library* seperti *Python*, *Transformer* (*HuggingFace*), dan *TensorFlow*. Hasil klasifikasi sentimen terhadap tweet yang mengandung tagar *#kaburajadulu* akan dianalisis dan divisualisasikan.

BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan dari keseluruhan penelitian yang dilakukan serta saran untuk pengembangan dan penelitian lanjutan di masa mendatang, terutama dalam

konteks pemanfaatan model analisis sentimen berbasis NLP terhadap isu-isu sosial yang berkembang di media sosial.

