

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan kecerdasan buatan, khususnya model bahasa berbasis *Natural Language Processing* (NLP), telah memberikan perubahan signifikan dalam berbagai aspek kehidupan, termasuk dunia akademik [1, 2, 3]. Salah satu implementasi yang paling banyak digunakan adalah ChatGPT yang dikembangkan oleh OpenAI, yang dimanfaatkan oleh mahasiswa untuk membantu memahami materi perkuliahan, menjawab pertanyaan akademik, hingga mendukung penyusunan tugas dan karya tulis ilmiah [4, 5]. Pemanfaatan teknologi ini tidak lagi bersifat tambahan, melainkan telah menjadi bagian dari aktivitas akademik sehari-hari. Kehadiran ChatGPT juga memicu perubahan pola belajar, di mana mahasiswa mulai beralih dari penggunaan sumber konvensional seperti buku dan jurnal menuju interaksi berbasis percakapan dengan AI [6]. Perubahan ini menimbulkan beragam pandangan, mulai dari peningkatan efisiensi belajar hingga kekhawatiran terhadap ketergantungan, kedalaman pemahaman, dan integritas akademik [7].

Meskipun penggunaan ChatGPT semakin luas, pemahaman terhadap sentimen pengguna, khususnya dalam konteks akademik di Indonesia, masih belum tergambarkan secara menyeluruh [8]. Penilaian umum seperti rating pada platform distribusi aplikasi menunjukkan kecenderungan positif, namun bersifat agregat dan tidak mampu merepresentasikan opini secara spesifik maupun kontekstual [9, 10]. Selain itu, ulasan pada platform distribusi aplikasi tidak terbatas pada domain akademik sehingga belum secara khusus mencerminkan persepsi mahasiswa sebagai pengguna ChatGPT dalam kegiatan pembelajaran [10]. Di sisi lain, pengguna cenderung mengekspresikan opini, pengalaman, dan kekhawatiran mereka secara lebih terbuka di media sosial seperti Twitter/X, yang mencerminkan kondisi nyata penggunaan dalam aktivitas akademik [11]. Oleh karena itu, diperlukan pendekatan analisis sentimen berbasis teks yang mampu mengklasifikasikan opini pengguna secara otomatis dalam skala besar dan kontekstual. Analisis sentimen merupakan bagian dari *Natural Language Processing* (NLP) yang bertujuan untuk mengklasifikasikan opini pengguna ke dalam kategori positif, negatif, dan netral [12].

Beberapa penelitian terdahulu telah mengkaji analisis sentimen terhadap ChatGPT maupun pemanfaatannya dalam konteks akademik, namun masih memiliki keterbatasan. Penelitian berbasis ulasan aplikasi umumnya lebih berfokus pada performa aplikasi secara umum, seperti kualitas layanan, antarmuka, maupun respons sistem, sehingga belum secara spesifik

merepresentasikan penggunaan ChatGPT dalam aktivitas akademik [10, 13]. Sementara itu, penelitian berbasis media sosial memanfaatkan opini publik yang disampaikan pengguna melalui platform seperti Twitter/X sebagai sumber data untuk analisis sentimen [14], sedangkan penelitian terkait ChatGPT umumnya masih berfokus pada persepsi pengguna terhadap aplikasi secara umum [11]. Di sisi lain, studi persepsi mahasiswa umumnya menggunakan pendekatan survei atau wawancara dengan jumlah responden terbatas, sehingga kurang merepresentasikan opini secara luas [13]. Pendekatan analisis sentimen terdahulu umumnya menggunakan metode pembelajaran mesin klasik berbasis fitur teks seperti *Support Vector Machine* (SVM) dengan representasi *TF-IDF* [15, 16]. SVM merupakan salah satu algoritma klasifikasi yang telah banyak digunakan dalam analisis sentimen karena mampu memberikan performa yang baik pada klasifikasi teks berdimensi tinggi [17].

Namun, representasi *TF-IDF* memiliki keterbatasan dalam memahami konteks dan hubungan semantik antarkata pada teks media sosial [18]. Sebaliknya, model berbasis *Transformer* seperti IndoBERT mampu menghasilkan representasi kontekstual sehingga lebih efektif dalam memahami makna suatu kalimat [19]. Oleh karena itu, penelitian ini menggunakan IndoBERT sebagai model utama, sedangkan SVM dengan representasi *TF-IDF* digunakan sebagai model baseline untuk memberikan perbandingan terhadap pendekatan pembelajaran mesin klasik. Perbandingan kedua pendekatan tersebut dilakukan untuk mengevaluasi efektivitas model pembelajaran mesin klasik dan model berbasis *Transformer* dalam mengklasifikasikan sentimen pada dataset penelitian ini [20]. Berdasarkan kesenjangan tersebut, penelitian ini bertujuan untuk menganalisis sentimen pemanfaatan ChatGPT dalam konteks akademik menggunakan data berbahasa Indonesia, serta membandingkan performa model IndoBERT dan SVM untuk memperoleh pendekatan yang lebih efektif dalam memahami opini pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah:

1. Masih terbatasnya penelitian yang membahas analisis sentimen terhadap pemanfaatan ChatGPT dalam konteks akademik.

2. Belum diketahuinya efektivitas perbandingan antara model klasik SVM berbasis *TF-IDF* dan model modern IndoBERT berbasis *Transformer* dalam mengklasifikasikan sentimen pada data Twitter berbahasa Indonesia dalam konteks akademik.

1.3 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis sentimen terhadap pemanfaatan ChatGPT dalam konteks akademik berdasarkan data Twitter/X berbahasa Indonesia.
2. Membandingkan performa model klasifikasi SVM berbasis *TF-IDF* dengan model IndoBERT dalam klasifikasikan sentimen pada data Twitter berbahasa Indonesia dalam konteks akademik.

1.4 Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Memberikan gambaran sentimen pengguna terhadap pemanfaatan ChatGPT dalam konteks akademik berdasarkan data berbahasa Indonesia sebagai representasi opini publik secara lebih luas
2. Memberikan hasil perbandingan performa antara model SVM berbasis *TF-IDF* dan IndoBERT dalam klasifikasi sentimen, sehingga dapat menjadi acuan dalam pemilihan metode yang lebih efektif pada data berbahasa Indonesia.

1.5 Ruang Lingkup

Ruang lingkup pembahasan pada tugas akhir ini, yaitu:

1. Data penelitian yang digunakan berupa cuitan Twitter/X yang dikumpulkan pada rentang waktu Januari 2024 hingga Februari 2026 menggunakan metode *scraping*. Target pengumpulan data ditetapkan sebanyak 10.000 cuitan untuk memperoleh jumlah data yang memadai. Rentang waktu tersebut dipilih untuk mendapatkan variasi data yang lebih luas dan mendukung proses analisis sentimen terhadap pemanfaatan ChatGPT dalam konteks akademik.
2. Data yang dianalisis berupa opini pengguna dalam bentuk teks berdasarkan kata kunci terkait ChatGPT dan aktivitas akademik, seperti “ChatGPT akademik”, “ChatGPT tugas”, “ChatGPT

belajar”, ”AI akademik”, “AI skripsi” dan “ChatGPT tugas akhir”. Penelitian ini tidak membahas perilaku pengguna di luar teks yang tersedia.

3. Pelabelan sentimen dilakukan secara semi-otomatis menggunakan model *mdhugol/indonesia-bert-sentiment-classification* sebagai pelabel awal yang kemudian diverifikasi secara manual. Klasifikasi sentimen dibatasi pada tiga kategori, yaitu positif, negatif, dan netral.
4. Penelitian ini menggunakan dua pendekatan klasifikasi sentimen, yaitu model berbasis *Transformer* IndoBERT dan model klasik *Support Vector Machine* SVM dengan representasi fitur *TF-IDF*.
5. Evaluasi performa model dilakukan menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Penelitian ini tidak mencakup analisis sentimen berbasis aspek, tidak menggunakan data selain bahasa Indonesia, serta tidak membahas pengambilan data secara *real-time*.

