

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

TikTok merupakan aplikasi media sosial yang memungkinkan penggunaannya untuk membuat dan membagikan video pendek dengan beragam tema dan kreativitas. TikTok menjadi sangat populer di kalangan pengguna karena memungkinkan mereka untuk dengan mudah membuat video kreatif dan unik (Susanto, 2021). Dalam beberapa tahun terakhir, Tiktok telah tumbuh menjadi salah satu platform media sosial terbesar di dunia dengan lebih dari 1 miliar pengguna aktif bulanan (Azizah et al., 2021).

Saat ini, terdapat banyak topik yang sedang viral di Tiktok. Salah satu topik yang sedang viral saat penelitian ini dilakukan adalah tentang pemilihan umum presiden dan wakil presiden (Pilpres) Indonesia yang akan dilaksanakan pada tanggal 14 Februari 2024 (Arum Wibawana, 2023). Oleh karena itu, banyak netizen yang memberikan komentarnya mengenai pilpres pada Tiktok, baik komentar yang positif ataupun komentar yang negatif (CNBC, 2023). Menurut Qadri (2020), media sosial dapat mempengaruhi opini dari publik. Sebagian besar informasi tentang Pilpres yang didapatkan oleh masyarakat berasal dari media sosial salah satunya adalah tiktok (Sakti, 2023). Oleh karena itu, komentar - komentar tersebut dapat berdampak besar dalam memengaruhi publik.

Untuk membedakan komentar - komentar tersebut ke dalam komentar positif dan negatif, dapat dilakukan dengan menggunakan analisis sentimen. Analisis sentimen merupakan teknik *mining* kontekstual yang digunakan untuk mengekstrak informasi yang berguna dan bersifat subjektif dari data yang berbasis teks (Tabinda Kokab et al., 2022). Penelitian yang dilakukan oleh Wang et al. (2020) menggunakan model BERT untuk melakukan analisis sentimen terhadap COVID-19 dengan menggunakan *dataset* yang diambil dari sosial media Sina Weibo atau yang biasa disebut sebagai Weibo untuk mengidentifikasi sentimen yang negatif, mendapatkan performa berupa akurasi sebesar 75,65% ketika dibandingkan dengan model - model lainnya, seperti SVM yang mendapatkan akurasi sebesar 70,66%, *Naïve Bayes* sebesar 66,97%, *Logistic Regression* sebesar 70,02%, CNN sebesar 71,19%, dan LSTM sebesar 57,73%. Selain itu, penelitian yang dilakukan oleh Singh et al. (2021) juga menggunakan model BERT untuk melakukan analisis sentimen terhadap dampak COVID-19 pada kehidupan sosial dengan menggunakan *dataset* yang diambil dari sosial media twitter untuk mengetahui banyaknya sentimen positif dan negatif, mendapatkan performa berupa akurasi sebesar 93,89%. Penelitian yang dilakukan oleh Boukabous & Azizi (2022) juga menggunakan model *hybrid approach* dengan menggunakan *lexicon-based* untuk menghasilkan output berupa label kemudian *output* tersebut digunakan pada model BERT untuk training. Model tersebut digunakan untuk melakukan analisis sentimen terhadap *dataset* yang diambil dari twitter untuk mendeteksi potensi ancaman keamanan, mendapatkan performa berupa akurasi sebesar 94,91% ketika dibandingkan dengan model - model lainnya, seperti SVM yang mendapatkan

akurasi sebesar 91,55%, CNB sebesar 88,17%, DT sebesar 82,46%, KNN sebesar 78,06%, *lexicon – based* sebesar 93%. Dari penelitian - penelitian sebelumnya yang telah dilakukan tersebut dapat disimpulkan bahwa model BERT memiliki akurasi yang paling tinggi ketika dibandingkan dengan model - model lainnya.

Oleh karena itu, penelitian ini akan menggunakan model BERT untuk melakukan analisis sentimen terhadap pilpres 2024 dengan menggunakan *dataset* berupa komentar yang diambil dari sosial media tiktok untuk mengetahui sentimen masyarakat terhadap pilpres 2024. Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan studi analisis sentimen pada sosial media TikTok, serta dapat memberikan wawasan yang lebih baik tentang pandangan masyarakat terhadap calon-calon presiden dan isu yang relevan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang tertulis diatas, maka rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana mengembangkan model analisis sentimen dengan BERT untuk *hashtag* #Pilpres2024 pada TikTok?
2. Bagaimana performa model analisis sentimen dengan BERT untuk *hashtag* #Pilpres2024 pada TikTok?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah penulis rumuskan diatas, maka tujuan dari penelitian ini adalah:

1. Menjelaskan cara mengembangkan model analisis sentimen dengan BERT untuk *hashtag* #Pilpres2024 pada TikTok.
2. Menganalisis performa model SA dengan BERT untuk *hashtag* #Pilpres2024 pada TikTok.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data diambil dari *hashtag* #Pilpres2024 dengan 10 konten teratas dalam rentang waktu dari tanggal 12 Desember 2023 hingga 24 Januari 2024.
2. Data diambil dengan menggunakan *web scraping*.
3. Penelitian ini menggunakan model BERT atau *Bidirectional Encoder Representations from Transformers*.

1.5 Manfaat Penelitian

Berikut adalah manfaat - manfaat yang diperoleh melalui penelitian ini:

1.5.1 Bagi Akademik

Penelitian ini dapat memberikan kontribusi dan pengetahuan dalam bidang analisis sentimen dan memberikan kesempatan kepada peneliti lainnya untuk mengembangkan penelitian ini.

1.5.2 Bagi Masyarakat

Penelitian ini dapat memberikan wawasan yang lebih baik tentang pandangan masyarakat terhadap calon-calon presiden dan isu yang relevan.

1.6 Sistematika Penulisan

Untuk mempermudah penyusunan skripsi ini, maka dibutuhkan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini akan dijelaskan beberapa poin utama yang menjadi dasar atau landasan dalam merancang sebuah penelitian. Bab ini terdiri dari beberapa sub bab diantaranya adalah latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisi penjelasan dan penyajian kerangka teori yang mendukung analisis dan pembahasan dalam penelitian, yaitu rancangan untuk menyelesaikan masalah utama yang akan dibahas. Bab ini terdiri dari beberapa sub bab diantaranya adalah landasan teori dan penelitian terdahulu.

BAB III METODOLOGI

Bab ini berisi tentang data - data dan metode yang digunakan pada saat melakukan penelitian. Bab ini terdiri dari beberapa sub bab diantaranya adalah metode pengumpulan data / penyajian data, kerangka pikir, hipotesis, tahapan penelitian, metode penyelesaian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil dari penelitian yang dilakukan dan pembahasannya. Bab ini terdiri dari beberapa sub bab diantaranya adalah hasil penelitian dan pembahasan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dan saran dari hasil penelitian. Bab ini terdiri dari beberapa sub bab diantaranya adalah kesimpulan dan saran.

