

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Data Google Trends menunjukkan pelonjakkan arus teknologi kecerdasan buatan di dunia dengan maraknya penggunaan *generative AI*. *Generative AI* telah menjadi cabang baru dari kategori kecerdasan buatan (AI) yang mengeluarkan sebuah *output* dengan cara mengumpulkan banyak data yang tersebar saat ini [1]. Hal tersebut menunjukkan bahwa perkembangan teknologi saat ini sangat pesat dan semakin memberikan kemudahan bagi manusia. Selain itu, perubahan teknologi juga akan mempengaruhi perubahan sosial dan perlu upaya lebih untuk mengarahkan teknologi tersebut dapat membentuk masyarakat yang lebih baik [2]. Berdasarkan data *Google Trends* didapatkan bahwa tahun 2022 merupakan awal pelonjakkan penggunaan *generative AI* didunia, di mana Negara Singapura menduduki posisi pertama sebagai negara terbesar penggunaan *generative AI*. Kemudian, Jakarta merupakan pengguna terbesar *generative AI* di Indonesia.

Generative AI telah bertransformasi menjadi kekuatan dalam bidang teknologi yang secara mandiri menghasilkan berbagai produk, seperti foto, teks, audio, dan sebagainya [3]. Dengan kemampuan yang dapat dilakukan oleh *generative AI*, telah memberikan banyak ruang untuk diperdebatkan dan diskusi lebih lanjut terkait dengan dampaknya terhadap kreativitas, etika, dan interaksi manusia itu sendiri. Istilah *generative AI* sendiri mengacu pada tingkat komputasi tinggi untuk menghasilkan suatu produk yang diperonalisasi oleh penggunanya dan tidak hanya bertujuan dalam hal “kreatif”, tetapi juga membantu manusia menjawab pertanyaannya [4]. Oleh karena itu, teknologi tersebut secara tidak langsung telah mengubah bagaimana manusia secara harfiah bekerja dan bagaimana manusia berkomunikasi antar satu sama lain [4]. Dilansir dari situs cnnindonesia.com, Nezar Patria, Wakil Menteri Komunikasi dan Informatika, menyatakan bahwa sebesar 79% Masyarakat di Indonesia telah berinteraksi dengan *generative AI* dan sebesar 35% perusahaan global telah menerapkan sistem AI dalam mobilitas perusahaannya. Dengan tingginya persentase penggunaan AI tersebut menjadikannya penting untuk dibahas dan ditinjau terhadap apa saja yang dapat

dihasilkan dari mewabahnya pengguna AI. Oleh karena itu, penelitian ini akan meninjau bagaimana sentimen masyarakat terkait fenomena hadirnya isu tersebut.

Analisis sentimen adalah sebuah proses analisis untuk melihat nada emosional seseorang dari sebuah teks terhadap sebuah objek atau entitas (Berupa individual, kejadian, maupun topik spesifik) [5]. Klasifikasi sentimen tersebut akan mengkategorikan sebuah pendapat seseorang bernada positif, negatif, atau netral [6]. Kemudian, sentimen tersebut akan dikelompokkan berdasarkan polaritasnya. Analisis sentimen memerlukan implementasi *Big Data* dan *DeepLearning*. Proses implementasi membutuhkan tingkat komputasi yang tinggi agar dapat memahami kompleksitas bahasa manusia beserta emosinya [7]. Dalam konteks *generative AI*, analisis sentimen disini berguna untuk meninjau opini masyarakat dan memberikan wawasan dengan melihat respons masyarakat terkait fenomena *generative AI*.

Media Sosial X merupakan layanan *microblogging* yang menjadi tempat untuk saling berbagi terkait topik atau kejadian yang sedang menjadi sorotan [7]. Dilansir dari situs databoks.katadata.co.id, menyatakan bahwa pengguna Media Sosial X di seluruh dunia mencapai 564 juta pengguna pada Juli 2023. Berdasarkan jumlah tersebut menjadikan Media Sosial X sebagai salah satu media dengan pengguna paling banyak dan pengguna paling aktif. Oleh karena itu, berdasarkan kualifikasi tersebut menjadikan Media Sosial X mampu dijadikan acuan untuk melihat sentimen pengguna terhadap *generative AI*.

Bidirectional Encoder Representations from Transformers (BERT) merupakan model *pre-trained* yang dikembangkan oleh ilmuwan di Google pada tahun 2018 [8]. BERT merupakan salah satu Teknik *neural network-based* untuk pemrosesan bahasa alami dan banyak diterapkan dalam kasus sentimen analisis [9]. Model BERT berhasil mengungguli sebelas model NLP lainnya [7] karena kemampuan *Transformer Encoder*-nya dapat menganalisis lebih dalam dan menghasilkan tingkat akurasi yang lebih baik [10]. Selain itu, BERT merupakan model pertama yang berhasil mengungguli kinerja tingkat manusia [7], di mana pemahamannya terhadap suatu konteks menjadikan hasil yang didapatkan akurat [11]. Penulisan sebuah kalimat atau kata-kata dapat menjadi sangat kompleks, model BERT disini

akan berfungsi untuk menangani apabila terdapat dependensi antar kata-kata [12] yang ditulis oleh pengguna Media Sosial X. Dengan kata lain, pusat dari model BERT *base-uncased* disini yaitu teknik *transformer* dan ketepatan pembangunan mekanismenya yang menentukan bobot relevansi setiap input yang ada [13]. Berdasarkan penjelasan tersebut, penggunaan model BERT *base-uncased* dalam penelitian ini dapat menganalisis sentimen pengguna Media Sosial X terhadap *generative AI* dengan keakuratan yang tinggi.

Pada penelitian terdahulu mengenai analisis sentimen menggunakan model BERT [13] dibahas sentimen publik selama wabah COVID-19 di China berdasarkan unggahan pada aplikasi Weibo, penelitian tersebut berhasil mendapatkan akurasi sebesar 75,65%. Meskipun akurasi tersebut tidak terlalu tinggi, namun menunjukkan penggunaan model BERT mampu menganalisis sentiment pengguna pada aplikasi Weibo terhadap wabah COVID-19. Kemudian, pada penelitian lain [14], dicapai nilai akurasi sebesar 90% ketika proses *training* dan akurasi sebesar 79% ketika *testing*. Nilai akurasi proses *testing* lebih kecil dikarenakan adanya terdapat banyak ambiguitas, seperti, bahasa gaul, kesalahan ejaan, dan sebagainya.

Oleh karena itu, peneliti tertarik untuk menerapkan model BERT *base-uncased*, yang dikenal karena kemampuannya memahami konteks dan nuansa linguistik, untuk mempelajari opini dan sentimen pengguna Media Sosial X terhadap “Produk” yang dihasilkan oleh *generative AI*. Analisis sentimen ini penting untuk memahami bagaimana maraknya *generative AI* diterima dan dipahami oleh masyarakat dalam konteks media sosial yang dinamis seperti Media Sosial X, yang seringkali mencerminkan opini publik yang lebih luas. Dengan mengamati lebih dekat reaksi pengguna Media Sosial X, peneliti dapat mengidentifikasi tren dan pola opini publik terhadap *generative AI* dengan melihat tingkat keakuratan model BERT *base-uncased* dan membandingkan dengan model Long Short-Term Memory (LSTM). Di mana, model LSTM sebagai pembanding di sini dipilih karena dapat memberikan gambaran mengenai efektivitas pendekatan klasik dengan pendekatan berbasis *transformer*. Penelitian ini diharapkan dapat mengisi ketimpangan pengetahuan yang ada dan dapat memberikan kontribusi terkait pemahaman yang dihasilkan dari penelitian ini.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, perumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana opini atau sentimen pengguna Media Sosial X terhadap *generative AI*?
2. Bagaimana hasil implementasi model Bidirectional Encoder Representations from Transformers (BERT *base-uncased*) dalam menganalisis sentimen pengguna Media Sosial X terhadap *generative AI*?
3. Bagaimana kemampuan model Bidirectional Encoder Representations from Transformers (BERT *base-uncased*) dalam menganalisis sentimen dibandingkan dengan model Long Short-Term Memory (LSTM)?

1.3. Tujuan dan Manfaat

Berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Untuk mengetahui opini atau sentimen pengguna Media Sosial X terhadap *generative AI* dengan membaginya dalam tiga kategori klasifikasi.
2. Menerapkan model Bidirectional Encoder Representations from Transformers (BERT *base-uncased*) dan melihat akurasi yang dihasilkan.
3. Melihat perbandingan akurasi dan performa model Bidirectional Encoder Representations from Transformers (BERT *base-uncased*) dengan model Long Short-Term Memory (LSTM).

1.4. Batasan Masalah

Berikut merupakan batasan masalah yang akan dibahas dalam topik penelitian ini:

1. Data yang digunakan dan dianalisis dalam penelitian didapatkan dari unggahan pada Media Sosial X sejak akhir tahun 2021 hingga 2024, di mana dalam cakupan mengenai opini terhadap *Generative AI* dengan metode *crawling*.
2. Dataset yang dihasilkan dari *crawling* pada penelitian ini dilakukan secara otomatis menggunakan model pralatih *nlptown/bert-base-multilingual-uncased-sentiment* yang tersedia pada pustaka HuggingFace. Sehingga, hasil penelitian dapat bervariasi apabila metode pelabelan yang digunakan berbeda.

3. Penelitian ini berfokus pada opini atau sentimen dari *tweet* pengguna Media Sosial X terhadap Generative AI. Kemudian, mengklasifikasikan opini tersebut ke dalam tiga kelas menggunakan model BERT *base-uncased* dan dibandingkan dengan model LSTM, dengan mengabaikan faktor eksternal lainnya yang tidak relevan dengan tujuan penelitian.
4. Data penelitian ini berfokus pada *tweet* atau pengguna Media Sosial X yang berbahasa Inggris. Sehingga, studi kasus bisa bervariasi tergantung pada ketersediaan data dan keterhubungannya dengan konteks penelitian ini.
5. Analisis penelitian akan berfokus pada opini dan sentimen pengguna Media Sosial X terhadap *generative* AI yang diidentifikasi melalui unggahan *tweet* yang ada, misalnya, sentimen yang lebih mengarah positif, netral, atau negatif. Serta, melihat kecenderungan opini atau sentimen tersebut secara linguistik dan arsitektural model.

1.5. Alur Kegiatan

Berdasarkan peninjauan dalam topik penelitian, maka berikut adalah gambaran kegiatan yang telah direncanakan:

1. Studi Literatur

Pada tahap ini, berfungsi untuk mencari dan mempelajari terhadap penelitian-penelitian yang berkaitan dengan topik sentimen analisis khususnya penelitian yang menggunakan BERT model. Selain itu, tahap ini juga akan mendalami teori-teori maupun studi terkait dengan rinci.

2. Identifikasi Topik

Pada tahap ini, topik yang diambil harus dipelajari lebih lanjut untuk melihat urgensinya dan mengeksplorasi yang berkaitan dengan topik.

3. Pengumpulan Data

Pada tahap ini, data didapatkan dengan melakukan *crawling* data pada media sosial Media Sosial X dengan lingkup atau cakupan “Produk” yang dihasilkan melalui *Generative* AI. Hasil data tersebut akan dilihat bagaimana pandangan pengguna Media Sosial X atas isu tersebut.

4. Perancangan dan Implementasi Sistem

Pada tahap ini, akan dilakukan perancangan sistem BERT yang akan dibangun dan dilanjutkan dengan mengerjakan kode program sesuai dengan sumber daya yang ada.

5. Evaluasi Performansi

Pada tahap ini, akan mengukur performa dari model yang digunakan dan akan dilakukan perbaikan apabila ditemukan hal yang tidak sesuai dengan rencana pembangunan model atau sistem.

6. Penulisan Laporan

Tahap ini merupakan tahap mendokumentasikan hasil penelitian yang telah dilakukan yang berbentuk laporan tugas akhir dan melihat manfaat potensial yang dapat dihasilkan dari penelitian tersebut.