

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

TokoTeknoPro yang sebelumnya dikenal dengan nama TokoLaptopID merupakan sebuah perusahaan yang bergerak di bidang jasa retail barang-barang teknologi. TokoTeknoPro menjual produk di berbagai platform yaitu Shopee, Tokopedia, dan juga Tiktok Shop. Dalam bisnis model retail, *inventory restocking* menjadi salah satu kunci penting. Terdapat beberapa aspek yang harus menjadi pertimbangan saat proses *restocking* seperti waktu, jumlah, dan juga produk yang akan di-*restock*. Untuk mempermudah proses tersebut, penjual dapat menggunakan fitur yang disediakan pada platform *e-commerce* seperti *filter* dan *sort* berdasarkan produk yang terlaris. Selain berdasarkan persentase angka produk yang terjual, *review* pelanggan juga menjadi pertimbangan TokoTeknoPro untuk melakukan *restock*. Hal ini dikarenakan pada transaksi *online e-commerce*, *rating* dan *review* pelanggan merupakan salah satu aspek yang akan mempengaruhi keputusan pembeli untuk membeli product di suatu toko [1]. Dikarenakan hal tersebut, penting bagi suatu toko untuk menjaga penilaian terhadap product yang akan dijual.

Rating dan *Review* pada *e-commerce* terdiri dari berbagai macam model seperti Bintang, *Text*, Gambar, ataupun Video. *Review* berupa *text* lebih bisa dipercaya oleh *customer* karena bersifat lebih detail dan personal dari satu pelanggan untuk disampaikan kepada calon pelanggan lain nantinya [1]. Dengan transaksi yang besar setiap harinya, sangat sulit bagi toko untuk melakukan analisa sentimen satu persatu pada setiap produk yang ada. Hal ini dipersulit juga dengan kondisi penilaian dengan bintang yang bisa tidak selaras dengan penilaian yang ditulis berupa *text*.

Untuk mengatasi hal tersebut saat ini banyak dikembangkan LLM (*Large Language Model*) yang dapat digunakan untuk melakukan *sentiment analysis* dalam bahasa tertentu [2]. LLM memiliki kemampuan untuk melaksanakan berbagai *task* tanpa perlu banyak pelatihan tambahan yang dikenal dengan *few-shot learning*. Pada pengaturan dengan *zero-shot* dan *few-shot*, LLM

kenunjukkan kinerja yang signifikan pada tugas sentimen analisis dibandingkan dengan model kecil yang dilatih dengan dataset khusus[10]. Penggunaan LLM untuk sentimen berbasis aspek mampu meningkatkan nilai F1 score sebesar 5.7% dibandingkan model InstructABSAs[11]. Pada penelitian terbaru oleh Tim *Data Science* Indonesia [2] dilakukan *benchmarking* IndoBERT dengan MBRET, MalayBERT, dan juga Bi LSTM-CRF dan didapati nilai F1 dari masing-masing metode yaitu *Naive Bayes* 70.95, *Logistic Regression* 72.14, BiLSTM w/ *fastText* 71.62, MBERT 76.58, MALAYBERT 82.02, dan INDOBERT 84.13. Terdapat juga penelitian mengenai uji performa ChatGPT, Gemini, dan LLaMA untuk *sentiment analysis* dengan metode pelatihan, validasi, dan pengujian [3]. Dari penelitian tersebut didapat hasil bahwa Gemini lebih konsisten dalam gaya percakapan sementara ChatGPT lebih bervariasi namun kurang konsisten. Beberapa penelitian terdahulu menunjukkan bahwa pemilihan model yang ideal bisa berbeda-beda untuk kebutuhan aplikasi yang spesifik. Maka dari itu pada penelitian ini akan dilakukan pengujian beberapa Model berbeda yaitu IndoBERT, ChatGPT, dan juga Gemini untuk mengetahui performa model terhadap *study case* TokoTeknoPro.

Penelitian ini akan berfokus pada evaluasi metodologi, strategi, atau teknik tertentu yang digunakan dalam proses pengembangan model *Transformer*. Untuk mengetahui model terbaik yang dapat digunakan untuk TokoTeknoPro maka dilakukan pengujian matriks performa pada beberapa model dengan data dari Tokopedia dan Shopee dari TokoTeknoPro.

1.2. Rumusan Masalah

Beberapa masalah yang ditemukan berdasarkan latar belakang yang telah dijabarkan oleh penulis antara lain:

- 1.2.1. Bagaimana perbandingan performa model IndoBERT, GPT, dan Gemini dalam melakukan *sentiment analysis* berbasis aspek?

1.3. Ruang Lingkup Penelitian Ruang Lingkup Penelitian

Ruang lingkup pada penelitian ini adalah sebagai berikut:

- 1.3.1. Penelitian dilakukan untuk melakukan perbandingan performa antar IndoBERT, GPT, dan Gemini dengan teknik yang paling efektif untuk masing-masing model pada tugas sentimen analisis berbasis aspek dilakukan untuk menguji teknik yang berfokus pada keefektifan model dan cara pengujian.
- 1.3.2. Dataset yang digunakan didapatkan dari hasil data *crawling* dari seluruh produk TokoTeknoPro dari platform Tokopedia dan Shopee yang masih aktif hingga tahun 2024.
- 1.3.3. Pelatihan model akan dilakukan dengan python untuk IndoBERT, environment OpenAI untuk GPT, dan environment Vertex AI untuk Gemini.
- 1.3.4. Ulasan pelanggan adalah ulasan yang menggunakan Bahasa Indonesia.

1.4. Tujuan Penelitian

Ruang lingkup pada penelitian ini adalah sebagai berikut:

- 1.4.1. Untuk mengetahui nilai dari Precision, Recall, dan F1-Score dari model finetuned IndoBERT, GPT, dan Gemini untuk sentiment analisis berbasis aspek pada dataset ulasan pelanggan TokoTeknoPro

1.5. Manfaat Penelitian

Manfaat yang bisa diberikan dari penelitian ini adalah :

- 1.5.1. Manfaat bagi peneliti adalah untuk mengetahui perbandingan performa antar model disertai metrik lama waktu pemrosesan dan harga untuk pengaplikasian pada sentimen analisis berbasis aspek dalam dataset ulasan pelanggan..
- 1.5.2. Manfaat bagi toko adalah bisa menganalisis ulasan produk secara lebih mendetail
- 1.5.3. Dengan pengembangan LLM analisa sentimen bisa dikembangkan dengan akurasi yang lebih tinggi dan skala data yang lebih besar