

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Perkembangan era digital yang pesat, membuat aplikasi perjalanan telah menjadi sarana utama bagi banyak individu untuk merencanakan dan memesan perjalanan mereka. Traveloka, sebagai salah satu platform perjalanan terkemuka di Indonesia yang memberikan layanan pemesanan tiket pesawat, hotel, kereta api, dan berbagai layanan lainnya (Traveloka, 2023). Pengguna aplikasi ini seringkali berbagi pengalaman mereka melalui ulasan atau testimoni, yang mencakup berbagai aspek seperti kualitas layanan, pengalaman pengguna, dan aspek-aspek lain yang mempengaruhi keputusan perjalanan.

Traveloka sebagai platform perjalanan terkemuka di Indonesia, memiliki jutaan pengguna yang memberikan ulasan terkait pengalaman mereka dalam menggunakan layanan aplikasi ini (Traveloka, 2023). Seiring dengan pertumbuhan pesat industri perjalanan, pengguna Traveloka aktif menyampaikan pro dan kontra melalui ulasan, mencakup berbagai aspek, mulai dari pemesanan tiket pesawat hingga penginapan di hotel. Sejalan dengan perkembangan teknologi, ulasan tersebut tidak hanya disampaikan secara langsung melalui platform aplikasi, tetapi juga melalui media sosial. Dalam konteks ini, Traveloka dihadapkan pada tantangan pemrosesan saran dan keluhan yang semula diterima melalui berbagai kanal (Traveloka, 2023). Meskipun saluran komunikasi yang melibatkan *customer service* sudah ada, namun dengan adanya ulasan, tanggapan dari pengguna dapat

disampaikan secara *real-time*, memberikan gambaran langsung terkait tingkat kepuasan pelanggan terhadap pelayanan yang diberikan oleh Traveloka.

Dataset yang telah dikumpulkan pada penelitian ini merupakan ulasan aplikasi Traveloka yang didapat dari ulasan App Store maupun Google Play yang mencakup teks dan *rating* dari pengguna. Seperti yang dijelaskan dalam artikel oleh Ashokkumar Palanivinayagam dan rekan-rekannya (Palanivinayagam et al., 2023), metode klasifikasi teks berbasis algoritma Machine Learning yang modern telah menunjukkan peningkatan signifikan dalam akurasi dibandingkan dengan metode tradisional. Oleh karena itu, penggunaan model-model ini bisa menjadi solusi lebih baik untuk mengatasi tantangan yang ada dalam analisis sentimen dari ulasan aplikasi.

Analisis sentimen terhadap ulasan pelanggan menjadi sangat penting dalam rangka memahami persepsi dan preferensi pengguna terhadap layanan yang disediakan oleh Traveloka (Syahid, 2015). Penggunaan teknologi pemrosesan bahasa alami, khususnya model BERT (Bidirectional Encoder Representations from Transformers), dapat memberikan metode analisis sentimen yang lebih canggih dan kontekstual (Chandradev et al., 2023). BERT memiliki kemampuan untuk memahami hubungan antar kata dalam kalimat dengan baik, sehingga dapat meningkatkan akurasi dalam mengevaluasi sentimen yang terkandung dalam ulasan-ulasan pelanggan (Sariwening, 2020). DistilBERT, yang merupakan versi yang lebih ringan dan lebih cepat dari BERT, dirancang untuk mempertahankan sebagian besar kapabilitas asli dengan penggunaan sumber daya yang lebih rendah. Teknik distilasi pengetahuan yang digunakan dalam pengembangan DistilBERT

memungkinkan pengurangan ukuran model sambil tetap efektif dalam memproses bahasa dan menginterpretasi konteks (Putri et al., 2023).

Dalam penelitian oleh (Fajri et al., 2022) berjudul "Membandingkan Nilai Akurasi BERT dan DistilBERT pada Dataset Twitter" DistilBERT digunakan untuk melakukan analisis sentimen pada data dari Twitter. Analisis sentimen adalah proses penggunaan *Natural Language Processing* (NLP) untuk mengidentifikasi dan mengkategorikan opini dalam teks untuk menentukan apakah sikap yang diungkapkan dalam teks tersebut positif, negatif, atau netral. Menurut penelitian ini, DistilBERT memberikan hasil yang sangat mengesankan dengan mencapai nilai akurasi 97%, presisi 99%, *recall* 99%, dan skor F1 99%. Nilai-nilai ini menunjukkan kinerja yang sangat tinggi dalam klasifikasi sentimen, yang lebih baik dibandingkan dengan BERT, yang memberikan akurasi 87%, presisi 91%, *recall* 91%, dan skor F1 89%. Keunggulan DistilBERT dalam penelitian ini mencerminkan efisiensi model dalam memproses data yang besar dengan kecepatan yang lebih tinggi sambil mempertahankan tingkat akurasi yang tinggi.

(Verdikha et al., 2023a) mengkaji penerapan model DistilBERT dan SVM untuk mengidentifikasi ujaran kebencian. DistilBERT digunakan untuk ekstraksi fitur dari tweet berbahasa Indonesia, yang kemudian di analisis menggunakan SVM dengan kernel linier dan polinomial. Karena DistilBERT menghasilkan dimensi fitur yang tinggi, teknik reduksi dimensi PCA diterapkan untuk memudahkan komputasi. Evaluasi menggunakan *10-fold cross validation* menunjukkan bahwa model dengan kernel polinomial pada 50 dimensi

memberikan *F1-Score* tertinggi 0.7857, menunjukkan efektivitas pendekatan ini dalam membedakan ujaran kebencian di media sosial.

Berdasarkan uraian diatas, penelitian ini bertujuan untuk mengembangkan model analisis sentimen menggunakan BERT dengan konteks spesifik ulasan aplikasi Traveloka di App Store dan Google Play. Penelitian ini diharapkan dapat memperoleh pemahaman yang lebih mendalam tentang kepuasan pengguna, faktor-faktor yang mempengaruhi, dan potensi perbaikan layanan yang dapat diidentifikasi dari analisis sentimen ulasan pelanggan pada aplikasi Traveloka.

## **1.2 Rumusan Masalah**

Sesuai dengan uraian pada latar belakang, maka permasalahan yang dihadapi yaitu:

1. Bagaimana cara mengembangkan model analisis sentimen ulasan Traveloka pada App Store dan Google Play menggunakan DistilBERT.
2. Bagaimana kinerja model BERT dalam mengklasifikasikan sentimen ulasan pelanggan, termasuk pengukuran tingkat akurasi dan kehandalan dalam analisis sentimen pada berbagai jenis layanan di dalam aplikasi Traveloka.

## **1.3 Tujuan Penelitian**

Tujuan penelitian ini adalah:

1. Mengetahui cara mengembangkan model analisis sentimen ulasan Traveloka pada App Store dan Google Play menggunakan DistilBERT.
2. Mengetahui kinerja model analisis sentimen dalam mengklasifikasi sentimen ulasan pelanggan, termasuk pengukuran tingkat akurasi dan

kehandalan dalam analisis sentimen pada berbagai jenis layanan di dalam aplikasi Traveloka.

#### **1.4 Batasan Masalah**

Berdasarkan latar belakang diatas, batasan penelitian ini adalah sebagai berikut:

1. Penelitian ini mengambil data ulasan Traveloka dari Apps store dan Google play store dari Kaggle (<https://www.kaggle.com/datasets/mgustiansyah/traveloka-id-application-rating-and-review-dataset>).
2. Penggunaan bahasa terbatas pada Bahasa Indonesia.
3. Menggunakan *pre-trained* model BERT.

#### **1.5 Manfaat Penelitian**

Berikut adalah manfaat-manfaat yang diperoleh melalui penelitian ini:

##### **1.5.1 Bagi Akademik**

Penelitian ini memberikan dedikasi dan pengetahuan dalam analisis sentimen. Selain itu, penelitian ini juga memberikan kesempatan kepada peneliti lain untuk mengembangkannya.

##### **1.5.2 Bagi Masyarakat**

Penelitian ini memberikan wawasan yang lebih baik kepada masyarakat terhadap aplikasi Traveloka di App Store dan Google Play.

## **1.6 Sistematika Penelitian**

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 PENELITIAN**

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, dan 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 pengumpulan data, 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.

