

# BAB I PENDAHULUAN

## 1.1 Latar Belakang Masalah

Mobilitas dan transportasi adalah aspek kunci dalam kehidupan masyarakat modern. Kendaraan bermotor menjadi tulang punggung sistem transportasi, memungkinkan manusia untuk menjalankan berbagai aktivitas dari perjalanan sehari-hari hingga pengangkutan barang [1]. Kecelakaan lalu lintas yang terjadi di jalan raya tidak hanya membawa dampak pada individu yang terlibat, tetapi juga menimbulkan implikasi sosial dan ekonomi yang luas. Dikutip dari portal berita *Bisnis.com*, Menurut Wakil Ketua Umum Bidang Keselamatan Transportasi MTI Rivian Achmad Purwanto, jumlah kecelakaan lalu lintas di Indonesia akan meningkat sebesar 6,8% pada tahun 2023 dibandingkan tahun sebelumnya.

Kecelakaan yang terjadi secara rutin membawa dampak serius bagi masyarakat, baik dalam bentuk kerugian fisik, psikologis, maupun sosial. Melihat kompleksitas dan keragaman situasi di jalan raya, menganalisis insiden-insiden ini menjadi suatu kebutuhan penting [2]. Penyelidikan yang cermat atas pola, faktor penyebab, dan potensi-potensi kecelakaan menjadi langkah krusial dalam upaya pencegahan dan perbaikan keamanan di sektor transportasi

Analisis dari berita kecelakaan di portal berita bertujuan untuk mengekstrak faktor-faktor penyebab kecelakaan yang sering muncul dalam laporan. Melalui informasi ini, dapat mengidentifikasi yang berpotensi menjadi penyebab kecelakaan, kondisi jalan atau lingkungan yang rentan terhadap insiden, serta dampaknya terhadap korban dan infrastruktur jalan. Fokusnya adalah untuk mengklasifikasikan yang menjadi penyebab terjadinya kecelakaan dan tren yang dapat membantu dalam merancang strategi keselamatan yang lebih efektif, tanpa terbatas pada satu sumber berita atau tahun tertentu. Dengan teknologi *Natural Language Processing*(NLP) yang semakin maju, pemanfaatan algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dalam menganalisis teks dari berbagai laporan insiden lalu lintas diharapkan dapat memberikan wawasan yang mendalam [3]. Hal ini tidak hanya membantu

memahami kejadian insiden secara lebih luas, tetapi juga memberikan informasi yang kritis dalam merumuskan solusi-solusi pencegahan yang lebih efektif.

Penerapan model *Natural Language Processing* seperti *Bidirectional Encoder Representations from Transformers* (BERT) dalam menganalisis insiden lalu lintas menjadi semakin penting seiring dengan perubahan perilaku pengguna jalan raya. Dari laporan media berita hingga diskusi-diskusi di platform sosial, data teks yang terkumpul memberikan keragaman informasi yang kaya, yang dapat menjadi sumber penting untuk merumuskan kebijakan-kebijakan transportasi yang lebih bijaksana. Namun, kompleksitas bahasa manusia serta kemampuan algoritma *Natural Language Processing* (NLP) yang terus berkembang menjadikan pengaplikasiannya sebagai langkah yang menantang.

BERT telah terbukti oleh Danny Sebastian, Hindriyanto Dwi Purnomo, dan Irwan Sembiring pada penelitian *BERT for Natural Language Processing in Bahasa Indonesia* yang bersumber dari jurnal IEEE menjadi algoritma yang unggul dalam pemrosesan bahasa alami. Dalam survei literatur tersebut, ditemukan 7 model BERT yang telah dilatih untuk bahasa Indonesia dan bahasa lokal Indonesia, dengan tugas downstream seperti analisis sentimen, klasifikasi, dan pembuatan ringkasan teks [4]. Sebagai perbandingan dengan algoritma lain dalam *natural language processing* lainnya yaitu LSTM, bahwa LSTM memiliki kelebihan berupa struktur yang lebih sederhana dibandingkan dengan model *transformer* seperti BERT, sehingga dapat dilatih lebih cepat dengan menggunakan sumber daya komputasi yang lebih sedikit. Namun LSTM memiliki kekurangan berupa Karena arsitekturnya yang linear, lebih sulit untuk memahami konteks global dan memproses data secara bersamaan[5]. Maka dari itu algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dipilih karena kemampuannya dalam memahami konteks teks secara mendalam.

## 1.2 Perumusan Masalah

Dari deskripsi latar belakang permasalahan yang telah disebut diatas, beberapa masalah yang dapat dapat dirumuskan adalah sebagai berikut:

1. Bagaimana implementasi model dengan algoritma *Bidirectional Encoder*

*Representations from Transformers* (BERT) untuk menganalisis laporan insiden lalu lintas melalui portal berita?

2. Bagaimana kinerja dari algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dalam menganalisis laporan insiden lalu lintas melalui portal berita?
3. Bagaimana cara model yang telah dilatih dengan algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dapat diimplementasikan dan digunakan untuk memprediksi laporan insiden lalu lintas melalui portal berita?

### **1.3 Tujuan Penelitian**

Berikut adalah beberapa tujuan penelitian ini:

1. Membangun model menggunakan algoritma *Bidirectional Encoder Representations from Transformers* (BERT) untuk menganalisis laporan insiden lalu lintas melalui portal berita.
2. Mengukur kinerja algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dalam analisis laporan insiden lalu lintas melalui portal berita untuk mengetahui akurasi dan kemampuan model
3. Membuat model yang telah dilatih dengan algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dapat diimplementasikan dan digunakan untuk memprediksi laporan insiden lalu lintas melalui portal berita.

### **1.4 Manfaat Penelitian**

Manfaat-manfaat yang diharapkan dari tuntasnya penelitian ini adalah sebagai berikut:

#### **1. Bagi Pengguna**

Penelitian ini juga diharapkan dapat berguna untuk memberikan pengguna berupa informasi yang relevan tentang insiden lalu lintas untuk upaya pencegahan yang lebih baik

## 2. Bagi penulis

Penelitian ini diharapkan dapat memberikan manfaat kepada penulis berupa pengalaman, pengetahuan baru, dan bisa memperdalam pemahaman dalam bidang Natural Language Processing.

## 3. Bagi Akademik

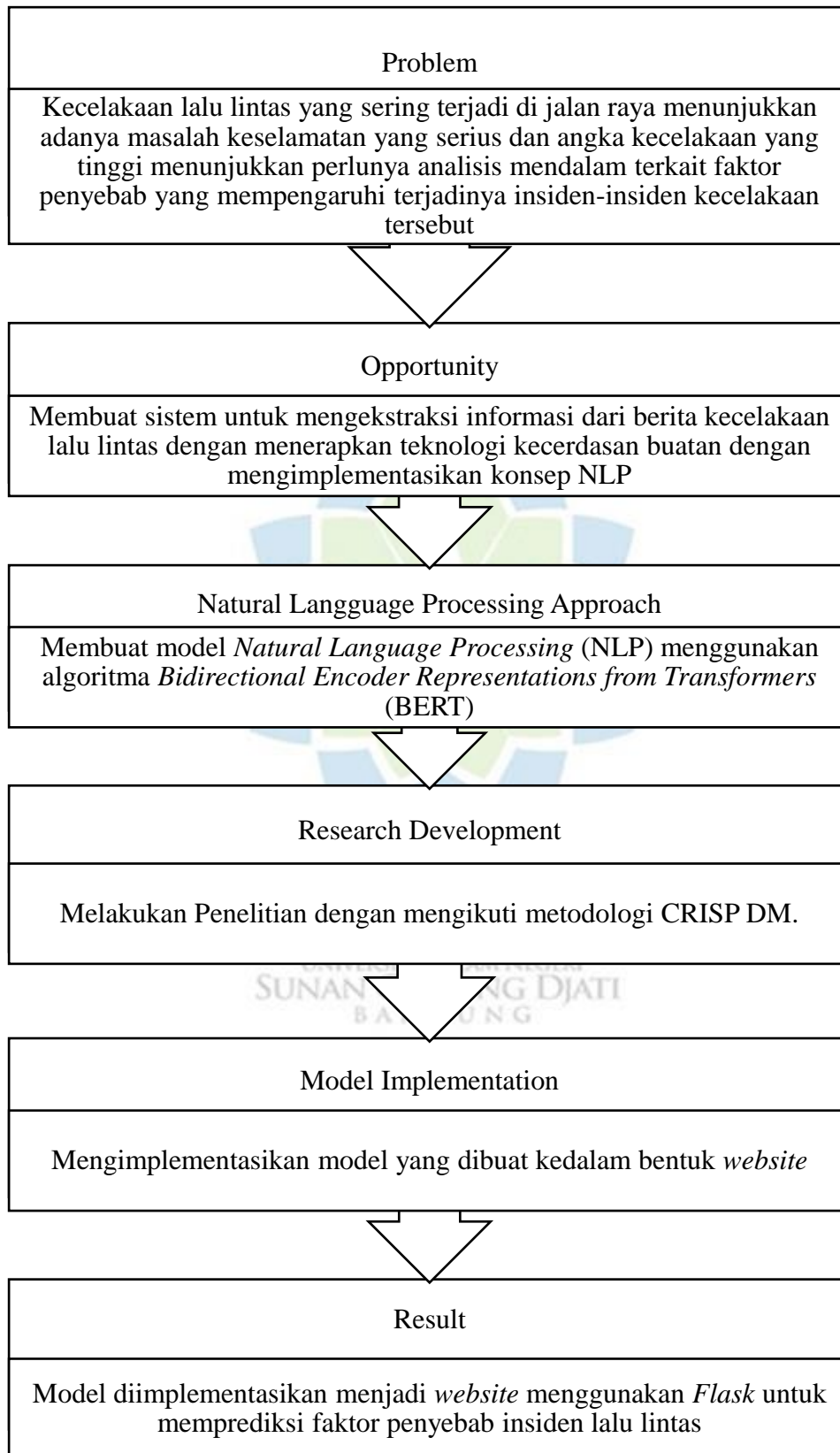
Bagi akademik, penelitian ini diharapkan mampu menjadi referensi bagi pembaca ataupun penelitian-penelitian selanjutnya agar dapat berkontribusi memajukan ilmu pengetahuan dalam bidang kecerdasan buatan, khususnya bidang *Natural Language Processing* khususnya dalam konteks analisis teks faktor terjadinya kecelakaan lalu lintas dari portal berita.

### 1.5 Batasan Masalah

Untuk memfokuskan dan mengarahkan penelitian ini, beberapa batasan masalah telah ditetapkan. Berikut adalah batasan-batasan dari penelitian ini:

1. Dataset yang digunakan pada penelitian ini berbentuk teks berbahasa Indonesia yang diambil berbagai situs portal berita seperti Detik.com, dan CNBC Indonesia terkait dengan berita kecelakaan lalu lintas.
2. Algoritma yang digunakan pada penelitian ini adalah BERT
3. Dataset yang dikumpulkan adalah berita yang dimulai dari tahun 2022, 2023, dan 2024 yang berkaitan dengan insiden kecelakaan lalu lintas di wilayah Jawa Barat.
4. Jumlah total dataset yang digunakan berjumlah 1100 berita berbentuk teks.
5. Faktor penyebab kecelakaan lalu lintas yang akan diklasifikasikan adalah faktor manusia (seperti: supir mengantuk, supir bermain *handphone* dan lain-lain), faktor kendaraan (seperti: rem blong, mesin rusak, ban pecah dan lain-lain), dan faktor lingkungan (seperti: jalan licin, jalan bergelombang, jalan rusak dan lain-lain).
6. Model yang telah dilatih akan disimpan pada platform *hugging face* agar bisa digunakan publik untuk diimplementasikan menjadi *website*.

## 1.6 Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran

## **1.7 Sistematika Penulisan**

Tugas akhir ini disusun secara sistematis dengan tujuan untuk memberikan pemahaman yang luas. Sistematika penulisan tugas akhir ini adalah sebagai berikut.

### **BAB I: Pendahuluan**

Bab I menjelaskan mengenai pendahuluan dari penelitian berupa latar belakang penelitian, rumusan masalah, tujuan dan batasan penelitian, serta kerangka pemikiran dan sistematika penulisan.

### **BAB II: Kajian Literatur**

Bab II membahas tentang kajian literatur dari pembahasan penelitian sebelumnya dan membahas teori yang relevan untuk menyelesaikan masalah penelitian.

### **BAB III: Metodologi Penelitian**

Bab III berisi rincian tentang metodologi penelitian yang digunakan dalam penelitian ini, serta penjelasan menyeluruh tentang setiap langkah dan metode yang digunakan.

### **BAB IV: Hasil dan Pembahasan**

Bab IV menguraikan pembahasan mengenai proses serta hasil yang telah dicapai dalam penelitian ini.

### **BAB V: Simpulan dan Saran**

Bab V menguraikan bagian akhir dari penelitian, termasuk kesimpulan keseluruhan yang menjawab rumusan masalah, serta memberikan saran untuk penelitian selanjutnya guna mengembangkan penelitian lebih lanjut.