

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Penelitian

Dalam era informasi yang semakin maju dan kompleks, pengelolaan data dan informasi yang efektif sangat penting bagi organisasi atau instansi pemerintahan untuk menjalankan operasional dan pengambilan keputusan dengan efisien, cepat dan tepat. Salah satu aspek kunci dalam pengelolaan data dan informasi adalah kemampuan untuk memantau dan menganalisis data secara *real-time*. *Command center* berfungsi sebagai pusat koordinasi dan wadah pengolahan *big data* terkait berbagai informasi [1].

*Command center* merupakan sebuah pusat pengendalian operasional yang dilengkapi dengan teknologi dan sistem informasi yang canggih. Fungsi utama *command center* adalah untuk mengumpulkan, menganalisis, dan menyajikan data secara *real-time* dari berbagai sumber yang berbeda. *Command center* juga berperan sebagai pusat komunikasi dan koordinasi untuk memonitor dan merespon peristiwa atau kejadian yang terjadi [2].

Beberapa fungsi penting dari *command center* adalah sebagai pusat kendali, *Alerting*, *Response Center*, *Monitoring*, *Decision Support System* dan *Showcase Public*. Untuk fungsi *Alerting* yaitu memberikan notifikasi atau peringatan kepada pihak terkait ketika terjadi peristiwa atau situasi yang membutuhkan perhatian atau tindakan segera. Dengan adanya sistem *alerting* di *command center*, organisasi atau instansi dapat dengan cepat mengetahui dan merespon situasi darurat, pelanggaran keamanan, atau gangguan dalam operasional mereka [2].

Selain fungsi *alerting*, *command center* juga memiliki peran penting dalam data mining atau penambangan data. Dalam konteks *command center*, *data mining* digunakan untuk menganalisis data yang dikumpulkan dari berbagai sumber dan mengidentifikasi pola, tren, atau anomali yang dapat memberikan wawasan dan informasi berharga. Dengan teknik-teknik *data mining* yang tepat, *command center*

dapat menggali informasi yang tersembunyi atau tidak terdeteksi secara manual, sehingga dapat digunakan untuk pengambilan keputusan yang lebih baik [3].

JCC yang merupakan singkatan dari Jawa Barat *Command Center* adalah pusat visualisasi dan integrasi data Jawa Barat yang dilengkapi dengan berbagai infrastruktur untuk kegiatan monitoring, koordinasi & pengambilan keputusan. JCC bertempat di kawasan Gedung Sate, Kota Bandung. Namun JCC masih belum mempunyai sistem yang dapat mengelola berbagai kejadian dari sumber berita di Jawa Barat. Hal ini diperlukan agar pihak-pihak terkait mengetahui berbagai kejadian di Jawa Barat dan pihak-pihak terkait bisa merespon kejadian tersebut melalui data yang dikumpulkan dan diolah oleh JCC [1].

Dikutip dari liputan6.com [4], tersebut dipublikasikan tanggal 09 Desember 2021 dimana disebutkan terjadi longsor di Desa Cimeong Kecamatan Banjaran kabupaten Majalengka, longsor ini mengakibatkan jalan yang menghubungkan 3 kabupaten yaitu Kabupaten Majalengka, Kabupaten Kuningan dan Kabupaten Ciamis terhambat karena longsor membuat sebagian jalan amblas dan harus diberlakukan sistem buka tutup satu arah oleh petugas. Namun dikutip dari berita yang dipublikasikan 14 April 2023 oleh pikiranrakyat.com [5], disebutkan bahwa pada tanggal 13 April 2023 longsor yang membuat sebagian jalan amblas di Desa Cimeong Kecamatan Banjaran tersebut masih belum diatasi dengan baik dan malah bertambah parah sehingga lalu lintas dari arah Cikijing ke Majalengka begitu juga sebaliknya masih diberlakukan sistem buka tutup satu arah. Dari dua berita kejadian tersebut dapat disimpulkan bahwa pemerintah Jawa Barat baik pusat maupun daerahnya lambat dalam merespon berita kejadian di Jawa Barat hal ini disebabkan karena pihak-pihak terkait tidak menerima informasi tersebut secara *uptodate*, tidak adanya sistem *alerting* yang menyebabkan berbagai instansi kemungkinan tidak mengetahui kejadian ini. Selain itu, jika berita ini muncul selama rentang waktu setahun lebih dengan frekuensi kemunculan ketika terjadi hujan lebat atau longsor yang membuat semakin parah maka pemerintah bisa menganalisis bahwa di daerah ini memang sering terjadi longsor yang menyebabkan jalan amblas sehingga

membantu pemerintah untuk menganalisis langkah apa yang harus diambil kedepannya agar lebih efektif.

Belakangan ini terdapat suatu teknologi yang mampu mengekstraksi data-data kejadian lokasi dari suatu berita. Teknologi Pemrosesan Bahasa Alami ini dapat dimanfaatkan untuk mengolah berita di Jawa Barat.

Penelitian yang berkaitan tentang *event extraction* sudah banyak dilakukan oleh para peneliti. *Event extraction* merupakan sub-bagian dari *information extraction* yang lebih luas. Dalam konteks penelitian ini, *event extraction* mengidentifikasi dan mengekstraksi informasi terkait peristiwa atau kejadian yang terjadi dalam teks. Peristiwa dapat berupa tindakan atau kejadian tertentu, dan ekstraksi peristiwa melibatkan identifikasi jenis peristiwa, peserta (participants) yang terlibat, waktu, lokasi, dan atribut lainnya yang terkait dengan peristiwa tersebut. *Event extraction* sering digunakan dalam analisis berita, penambangan berita, dan pemahaman konteks tertentu dari teks yang melibatkan serangkaian peristiwa..

*Event extraction* adalah bagian dari pemrosesan bahasa alami, bertujuan untuk menemukan, mengambil, atau memproses informasi. Sumber data untuk *event extraction* adalah teks. Untuk menghasilkan informasi, teks yang tidak terstruktur akan diubah menjadi data terstruktur. Ada banyak pendekatan yang dilakukan peneliti untuk proses ini. Sebagian besar penelitian tersebut dilakukan dalam Bahasa Inggris. Dalam makalahnya menyajikan peninjauan tren penelitian terkini, tantangan, dan peluang *event extraction* menggunakan bahasa Indonesia. Peninjauan ini menghasilkan peluang yang sangat bagus pada berbagai bidang industri dan pemerintahan yang dapat menunjang sistem Big Data [6]. Klasifikasi teks adalah teknik yang melibatkan pemberian kategori atau label yang telah ditentukan sebelumnya ke teks tertentu berdasarkan isinya. Dengan kata lain, klasifikasi teks adalah jenis *event extraction* yang melibatkan identifikasi dan *event extraction* yang relevan dari data teks yang tidak terstruktur dan menuliskannya ke kategori atau label yang telah ditentukan sebelumnya [7].

Bekele Abera Hordofa pada tahun 2020 [8] memperkenalkan model ekstraksi dan representasi peristiwa untuk artikel berita, dengan fokus pada identifikasi peristiwa penting, mengekstraksi elemen peristiwa semantik, dan mengklasifikasikan artikel ke dalam peristiwa dunia nyata dan bukan peristiwa. Model ini menggunakan analisis linguistik, menggunakan 220 kata dan frasa pemicu peristiwa, pengenalan entitas bernama, dan fitur lokal. Pengklasifikasi Entropi Maksimum digunakan untuk klasifikasi peristiwa, mencapai hasil yang menjanjikan dalam hal presisi untuk identifikasi pemicu peristiwa (67,1%), ekstraksi elemen peristiwa (69,1%), dan klasifikasi peristiwa (72%).

Budi Juarto dan Yulianto pada tahun 2022 melakukan penelitian untuk mengklasifikasi berita berbahasa Indonesia menggunakan algoritma IndoBERT. Penelitian ini menggunakan lima kategori berita untuk mengklasifikasikan berita berbahasa Indonesia, yaitu sepakbola, berita, bisnis, teknologi dan otomotif. Penelitian ini melakukan perbandingan antara algoritma IndoBERT dengan algoritma XLNET, BERT multilingual dan XLMRoberta. Hasil perbandingannya menunjukkan algoritma BERT lebih baik diantara algoritma yang lain karena algoritma ini sudah dilatih sebelumnya menggunakan dataset teks berbahasa Indonesia yang banyak jumlah katanya dengan akurasi 94.5% [9]. Pada tahun 2020, Hadiyan Kundrat Putra, dkk. Dalam jurnalnya menjelaskan mengenai tujuan dari penelitiannya yaitu penerapan metode IndoBERT untuk melakukan deteksi penggunaan kalimat abusive pada teks bahasa Indonesia, perhitungan nilai akurasi nilai F1 Score untuk masing masing kelas dan pengaruh adanya penambahan data pada dataset kelas minoritas terhadap nilai akurasi F1 Score. Hasil F1 Score model IndoBERT mengalami peningkatan setelah dilakukan adanya penambahan data pada kelas minoritas dengan nilai rata-rata F1-Score dari tiap kelas dengan 76,32% [10].

Berdasarkan uraian tersebut maka dirumuskan penelitian dengan judul **“Ekstraksi Event Berbasis Geolokasi Menggunakan BERT Topic Classification dan Spatial Knowledge Representation Pada Aplikasi Command Center”**. Dengan adanya penelitian ini diharapkan dapat menambah

inovasi untuk kegiatan monitoring kejadian di Jawa Barat, sehingga pengambilan keputusan dari berbagai instansi bisa lebih efektif dan efisien.

## 2.1 Rumusan Masalah

Rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Bagaimana mengimplementasikan kombinasi Algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dan Metode Berbasis Aturan untuk *event extraction* secara otomatis ?
2. Bagaimana kinerja dari kombinasi Algoritma *Bidirectional Encoder Representations from Transformers* BERT dan Metode Berbasis Aturan untuk *event extraction* secara otomatis ?

## 3.1 Tujuan Penelitian

Tujuan yang ingin dicapai penulis dalam melaksanakan penelitian ini adalah :

1. Implementasikan sistem klasifikasi berita berdasarkan tipe, topik dan lokasi menggunakan Algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dan Metode Berbasis Aturan.
2. Mengetahui kinerja kombinasi algoritma Algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dan Metode Berbasis Aturan untuk *event extraction* otomatis.

## 4.1 Batasan Masalah

Batasan masalah yang digunakan dalam sebuah pembahasan diharapkan agar pembahasan menjadi lebih terarah dan sesuai dengan tujuan yang akan dicapai. Adapun beberapa batasan masalah tersebut diantaranya:

1. Sistem akan memindai berita terkait Jawa Barat dari media digital yang ditentukan. Lalu mengekstrak tipe, topik dan lokasi berita dan menampilkan visualisasi spasial titik-titik lokasi berita

2. Algoritma yang akan digunakan adalah *Bidirectional Encoder Representations from Transformers* (BERT) dan Metode Berbasis Aturan.
3. Data yang digunakan untuk proses klasifikasi tipe dan topik berita merupakan dataset yang dibuat sendiri oleh penulis yang bersumber dari media digital di Jawa Barat dari awal tahap pengumpulan data menggunakan teknik *Scraping* lalu pembersihan data dan pelabelan.
4. Dataset yang digunakan untuk keperluan deteksi lokasi didapat dari website [opendata-jabarprov.go.id](http://opendata-jabarprov.go.id).
5. Label untuk topik berita yang akan digunakan diprioritaskan merupakan kejadian yang memiliki jangka waktu.

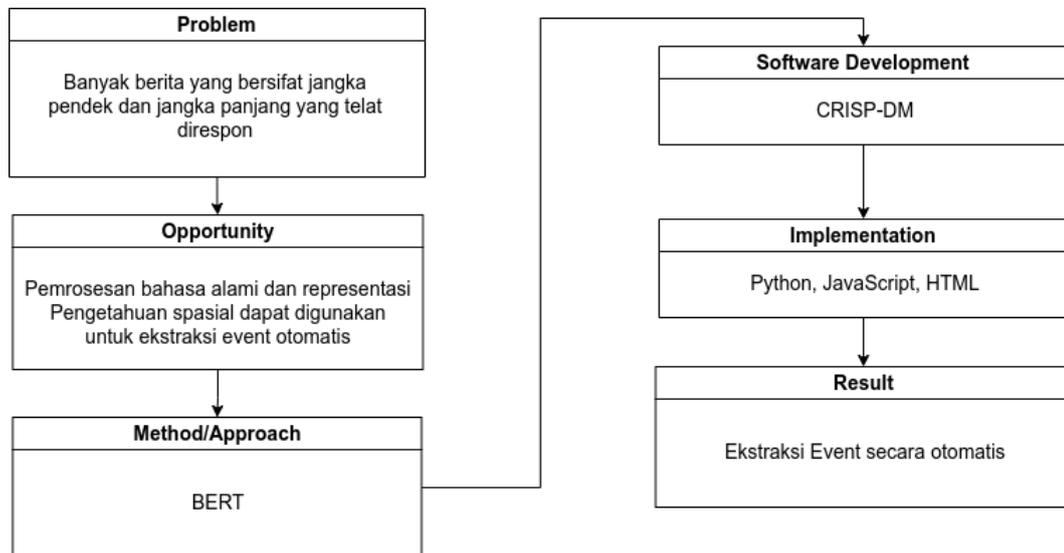
### 5.1 Manfaat Penelitian

Adapun manfaat yang didapat dari penelitian ini adalah :

1. Memberikan insight baru dari data yang sudah ada.
2. Membantu pengambilan keputusan lebih cepat dan efektif karena kejadian yang dimunculkan *realtime*.
3. Fungsi jangka pendek untuk *alerting* sebagai pengingat ke pihak-pihak terkait tentang kejadian yang terjadi dan harus segera ditangani. Fungsi jangka panjangnya untuk kegiatan *data mining* dan analisis data sehingga pemerintah bisa mendapatkan pola dan tren tentang berita kejadian sehingga bisa membantu pemerintah di masa mendatang.
4. Menjadi referensi untuk penelitian selanjutnya, khususnya yang berkaitan dengan ekstraksi event menggunakan *Bidirectional Encoder Representations from Transformers* (BERT) dan Metode Berbasis Aturan.

## 6.1 Kerangka Pemikiran

Adapun kerangka pemikiran dari penelitian ini terdapat pada Gambar 1.1 dibawah ini:



Gambar 1. 1 Kerangka Pemikiran

## 7.1 Sistematika Penulisan

Sistematika penulisan laporan ini disusun untuk menggambarkan secara umum mengenai penelitian yang penulis laksanakan.

### BAB 1 PENDAHULUAN

Bab I menjelaskan mengenai latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, kerangka pemikiran, metodologi penelitian, sistematika penulisan.

### BAB II STUDI PUSTAKA

Bab II menjelaskan dasar teori dan tinjauan pustaka yang berkaitan dengan topik penelitian yang dilakukan.

### BAB III ANALISIS DAN PERANCANGAN

Bab III menjelaskan mengenai analisis dan perancangan aplikasi yang dibangun, hasil analisis serta perancangan yang akan dimodelkan dengan UML.

#### BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab IV menguraikan mengenai perangkat keras dan perangkat lunak beserta spesifikasi yang digunakan , pengujian pengimplementasian terhadap aplikasi serta pada bab ini akan menampilkan tampilan aplikasi yang dibangun.

#### BAB V PENUTUP

Bab V menjelaskan mengenai kesimpulan dari bab-bab sebelumnya serta saran yang dipergunakan sebagai pertimbangan pengembangan sistem yang akan dibangun selanjutnya.

