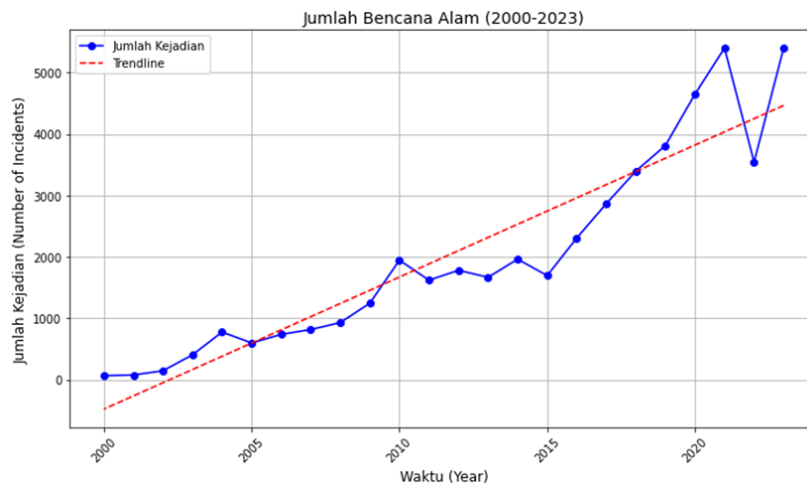


BAB I PENDAHULUAN

Pada bab pendahuluan akan dijelaskan mengenai gambaran umum dari topik penelitian yang dibahas. Bab pendahuluan memuat beberapa bagian yakni latar belakang, perumusan masalah, tujuan penelitian, batasan-batasan, serta potensi manfaat dari penelitian.

I.1 Latar Belakang

Indonesia merupakan salah satu negara yang tergolong rawan dari segi bencana alam. Hal ini dikarenakan letak Indonesia yang berada pada daerah cincin api pasifik atau *ring of fire*. Letak Indonesia yang berada pada daerah *ring of fire* ini menyebabkan Indonesia sering mengalami bencana alam seperti gempa bumi dan erupsi gunung api. Selain itu, kondisi topografi Indonesia yang memiliki banyak daerah dengan kontur pegunungan menjadi salah satu faktor penyebab terjadinya tanah longsor (Setiawan dkk., 2022).



Gambar I-1. Jumlah Kejadian Bencana Alam

Sumber: Data Informasi Bencana Indonesia (DIBI) BNPB

Berdasarkan Gambar I-1, jumlah bencana alam yang terjadi di Indonesia selama 20 tahun terakhir secara garis besar mengalami kenaikan. BNPB mencatatkan telah terjadi sebanyak 47.822 kali bencana alam sejak tahun 2000 hingga tahun 2023. Adapun jenis bencana alam yang dicatat oleh BNPB adalah banjir, tanah longsor, gelombang pasang dan abrasi, cuaca ekstrem, kebakaran hutan dan lahan,

gempa bumi, erupsi gunung api, gempa bumi dan tsunami, serta letusan gunung api.

Tabel I-1. Dampak Korban dan Kerusakan Rumah akibat Bencana Alam

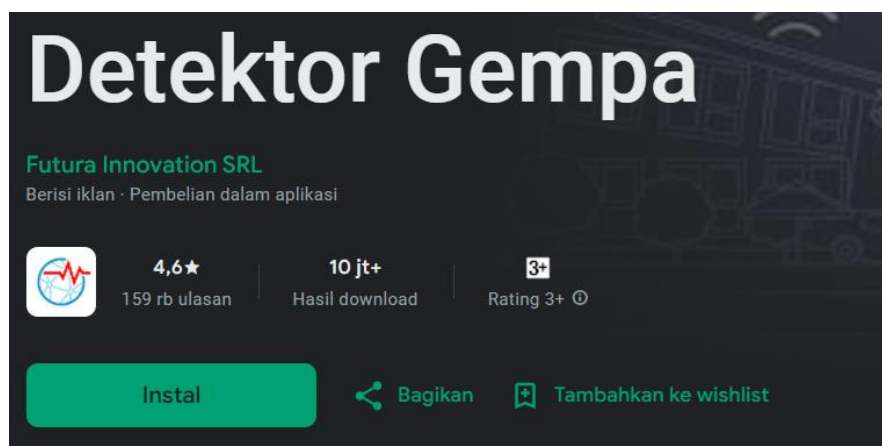
Waktu	Korban		Kerusakan Rumah		
	Meninggal	Menderita	Rusak berat	Rusak sedang	Rusak ringan
2023	325	8.340.746	4.870	6.079	36.245
2022	858	5.492.456	20.205	23.213	51.985
2021	728	7.581.247	21.335	27.936	109.388
2020	376	5.873.967	13.240	9.694	42.810
2019	478	5.693.171	15.813	14.565	43.345

Berdasarkan data dari BNPB pada Tabel I-1 juga tercatat bahwa jumlah total korban jiwa akibat bencana alam selama 5 tahun terakhir dari 2019 hingga tahun 2023 adalah sebanyak 2.765 jiwa. Selain mengakibatkan korban jiwa, bencana alam juga memberikan dampak kerusakan pada rumah penduduk yakni sebanyak 75.463 rumah yang mengalami rusak berat.

Salah satu faktor yang menyebabkan besarnya dampak dari bencana yang terjadi di Indonesia adalah keterlambatan informasi kepada penerima atau masyarakat. Keterlambatan informasi dapat mempengaruhi besarnya dampak bencana karena dapat menyebabkan terlambatnya masyarakat dalam melakukan evakuasi bencana terhadap anggota keluarga dan harta benda (Syahri & Sahrul, 2019). Info BMKG sebagai salah satu media informasi yang digunakan BMKG memiliki peran dalam melakukan mitigasi bencana. Oleh karena itu, evaluasi terhadap aplikasi Info BMKG diperlukan untuk menilai kekurangan Info BMKG dalam melakukan perannya dalam mitigasi bencana.

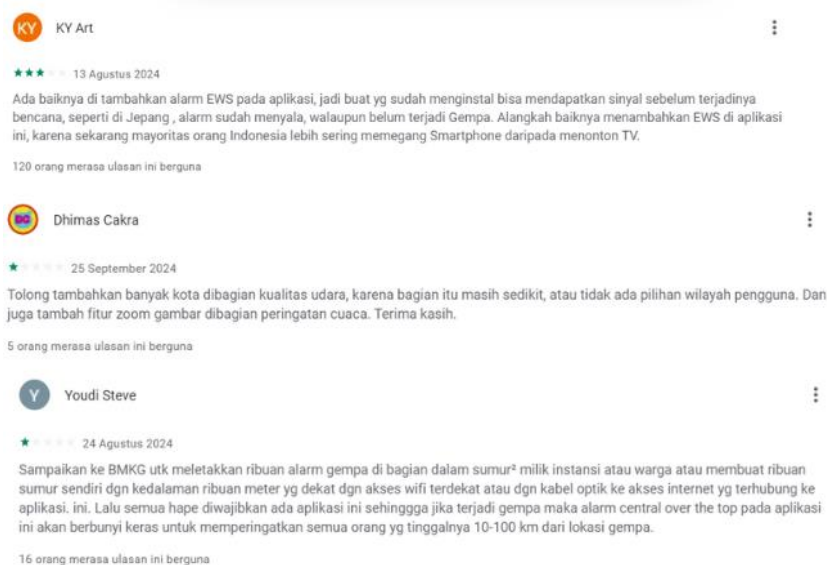
Hingga tanggal 19 Oktober 2024, jumlah unduhan aplikasi Info BMKG di *platform* Play Store mencapai lebih dari 5 juta unduhan. Jumlah penduduk Indonesia di tahun 2024 mencapai 282 juta penduduk. Hal ini menunjukkan

bahwa penduduk yang melakukan unduhan terhadap aplikasi mitigasi bencana masih sedikit yakni kurang dari 5% dari jumlah keseluruhan penduduk Indonesia. Dengan jumlah unduhan yang sedikit akan berakibat pada terbatasnya jangkauan informasi terkait cuaca dan bencana yang diperoleh penduduk. Sedangkan, dalam penanganan kebencanaan penyebaran informasi yang cepat dan akurat menjadi kunci penting (Syarif dkk., 2014). Oleh karena itu diperlukan penggalian informasi terkait ulasan pengguna yang masih menjadi kekurangan pada aplikasi Info BMKG yang menyebabkan jumlah pengunduh yang masih sedikit.



Gambar I-2. Aplikasi Detektor Gempa (Sumber: Google Play Store)

Berdasarkan Gambar I-2, Aplikasi Detektor Gempa memiliki rating 4.6 dengan jumlah unduhan lebih dari 10 juta. Aplikasi Detektor Gempa merupakan aplikasi yang memiliki fitur serupa dengan Info BMKG yakni fitur peringatan dini gempa. Dengan fitur serupa, aplikasi Detektor Gempa diunduh oleh lebih banyak pengguna dan memiliki rating yang lebih baik dibandingkan dengan aplikasi Info BMKG yang menunjukkan bahwa Info BMKG masih memerlukan evaluasi untuk dapat meningkatkan kualitas aplikasi Info BMKG dengan memperbaiki kekurangan yang dimiliki dan mempertahankan hal yang telah dirasa baik oleh pengguna.



Gambar I-3. Review Pengguna Info BMKG Play Store (Sumber: Google Play Store)

Berdasarkan Gambar I-3, ditemukan berbagai kritik dan saran dari pengguna aplikasi Info BMKG di *Google Play Store* yang menunjukkan adanya ruang untuk pengembangan dan perbaikan lebih lanjut. Masukan-masukan ini mencerminkan aspek-aspek tertentu dari aplikasi yang perlu ditingkatkan agar dapat memenuhi kebutuhan dan harapan pengguna dengan lebih baik. Oleh karena itu, evaluasi terhadap aplikasi sangat diperlukan. Evaluasi ini bertujuan untuk mengidentifikasi kelemahan dan kelebihan yang ada serta menemukan solusi yang tepat guna meningkatkan kualitas dan pengalaman pengguna secara keseluruhan.

Analisis ulasan atau komentar pengguna merupakan metode yang sering digunakan untuk melakukan evaluasi terhadap kelebihan dan kekurangan suatu aplikasi. Umpan balik negatif dapat menunjukkan kekurangan aplikasi dan umpan balik positif dapat digunakan untuk menilai seberapa baik aplikasi dalam memenuhi kebutuhan dan harapan dari pengguna (Oyebode dkk., 2020).

Data ulasan atau komentar memiliki karakteristik atau tipe data berupa teks. *Natural Language Processing* (NLP) menjadi teknik yang tepat untuk mengolah data dengan tipe teks. Teknik-teknik NLP memungkinkan mesin untuk memproses bahasa manusia dengan menerjemahkan bahasa manusia ke format yang dapat dipahami oleh mesin (Rajput, 2019).

Ulasan suatu aplikasi bukanlah data yang berjumlah sedikit yang dapat dianalisis secara manual (Olagunju dkk., 2020). Analisis sentimen sebagai bagian dari *Natural Language Processing* (NLP) dapat digunakan untuk mengidentifikasi dan mengenali emosi pengguna dari data ulasan (Mehta & Pandya, 2020).

Untuk melakukan analisis terhadap data teks, khususnya data yang berasal dari *review* dan media sosial, diperlukan adanya langkah-langkah *preprocessing* data untuk mempersiapkan data. Data teks berkemungkinan mengandung *noise* yang dapat mempengaruhi dan mengganggu hasil analisis. Dengan melakukan *preprocessing* data teks, data akan dibersihkan dari karakter-karakter yang tidak memiliki arti atau tidak bermanfaat, dinormalisasi sehingga data teks hanya mengandung informasi yang bermanfaat bagi analisis (Duong & Nguyen-Thi, 2021).

Penelitian terdahulu yang telah dilakukan terhadap aplikasi Info BMKG masih menggunakan model-model tradisional seperti SVM dan random forest. Dalam penelitian ini model modern CatBoost dipilih karena performa yang lebih baik dibandingkan dengan algoritma atau model *boosting* lainnya seperti AdaBoost, XGBoost, dan Gradient Boosting. Selain itu, pada penelitian sebelumnya analisis topik ulasan dilakukan dengan cara tradisional yakni dengan menggunakan frekuensi kemunculan kata. Pada penelitian ini akan digunakan BERTopic sebagai algoritma *topic modelling* yang menggunakan *embedding* berbasis *transformer* sehingga dapat lebih memahami konteks pada ulasan.

CatBoost merupakan salah satu model *boosting* yang dapat dimanfaatkan dalam melakukan analisis sentimen (Efat dkk., 2023). CatBoost memiliki kelebihan di dalam analisis sentimen karena CatBoost telah menyediakan parameter yang dapat digunakan untuk melakukan pengolahan *text features* termasuk tokenisasi. CatBoost juga cocok digunakan karena CatBoost dapat memperoleh performa yang lebih tinggi dengan jumlah data latih yang sedikit dibandingkan dengan model lainnya.(Yang, 2020). Menurut Rahman dkk., (2023), hasil pengujian terhadap beberapa model *boosting* untuk melakukan analisis sentimen menunjukkan bahwa CatBoost memiliki performa terbaik di antara model lainnya,

seperti AdaBoost, XGBoost, dan Gradient Boosting, dengan memperoleh skor F1 sebesar 85,8%.

Evaluasi terhadap aplikasi akan menggunakan *D&M IS Success model* untuk membagi review ke dalam topik seperti *system quality* dan *information quality*. *D&M IS Success model* merupakan salah satu model yang dapat digunakan untuk menilai keberhasilan suatu sistem informasi dalam hal penggunaan teknologi dan hasilnya (Sabeh dkk., 2021).

Analisis sentimen menghasilkan *output* berupa emosi dari suatu teks yakni positif, netral, atau negatif. Untuk mendapatkan informasi lebih dari ulasan mengenai aspek yang menjadi kekurangan pada aplikasi melalui ulasan dengan sentimen negatif, *topic modelling* dapat digunakan. *Topic model* merupakan model yang dapat digunakan untuk menemukan struktur semantik tersembunyi dari suatu data teks (Dieng dkk., 2020). BERTopic merupakan salah satu model yang dapat digunakan dalam melakukan *topic modelling*. Topik-topik yang dihasilkan oleh BERTopic dapat mengungkap informasi terkait subtopik yang terdapat dalam D&M model.

I.2 Perumusan Masalah

Sesuai dengan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah yang menjadi dasar penelitian adalah sebagai berikut:

- a. Bagaimana distribusi sentimen masyarakat terhadap aplikasi Info BMKG berdasarkan ulasan pengguna pada Play Store?
- b. Bagaimana performa model CatBoost dalam melakukan klasifikasi topik dan analisis sentimen dengan menggunakan data ulasan pengguna aplikasi Info BMKG?
- c. Apa saja subtopik yang teridentifikasi BERTopic untuk topik *system quality* dan *information quality* pada aplikasi Info BMKG?

I.3 Tujuan Penelitian

Adapun tujuan yang mendasari penelitian ini sesuai dengan rumusan masalah yang ada yakni:

- a. Mengetahui distribusi sentimen masyarakat terhadap aplikasi Info BMKG berdasarkan ulasan pengguna pada Play Store.
- b. Mengetahui performa model CatBoost dalam melakukan klasifikasi topik dan analisis sentimen terhadap data ulasan pengguna Info BMKG pada Google Play Store.
- c. Mengetahui subtopik yang menjadi kelebihan dan kelemahan untuk topik *system quality* dan *information quality* pada aplikasi Info BMKG dengan mengimplementasikan BERTopic.

I.4 Batasan Penelitian

Batasan dari penelitian ini yaitu:

- a. Analisis sentimen dilakukan dengan menggunakan 2 sentimen yaitu positif dan negatif.
- b. Analisis sentimen dan *topic modelling* dilakukan dengan data ulasan pengguna dalam Bahasa Indonesia.

I.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

- a. Bagi pengembang aplikasi Info BMKG, penelitian ini dapat bermanfaat untuk mengidentifikasi kekurangan dalam aplikasi Info BMKG serta menjadi opsi untuk melakukan evaluasi terhadap aplikasi Info BMKG.
- b. Bagi peneliti, penelitian ini dapat bermanfaat untuk memperdalam pengetahuan peneliti dalam bidang *data mining* khususnya dengan menggunakan *Natural Language Processing* (NLP) dengan melakukan analisis sentimen dan *topic modelling*.
- c. Bagi peneliti lain, penelitian ini dapat bermanfaat untuk mengetahui performa model CatBoost dan BERTopic dalam melakukan klasifikasi topik dan analisis sentimen serta *topic modelling* dengan menggunakan data ulasan pengguna aplikasi dari *platform* Google Play Store.

I.6 Sistematika Laporan

Penelitian ini disusun berdasarkan sistematika penulisan sebagai berikut:

- a. BAB I Pendahuluan, menjelaskan mengenai latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penelitian.
- b. BAB II Landasan Teori, berisi uraian teori-teori dan studi literatur yang relevan dengan topik penelitian. Pada bab ini akan dijelaskan teori terkait analisis sentimen, topic modelling, model yang digunakan, proses pembersihan data, dll. Selain itu akan diuraikan penelitian-penelitian terdahulu yang berkaitan dengan penelitian. Penelitian terdahulu dikaji untuk menunjukkan celah penelitian yang ingin dijawab serta sebagai pembanding untuk menilai penelitian.
- c. BAB III Metode Penyelesaian Masalah, menjelaskan metode penelitian yang digunakan dalam penelitian serta langkah-langkah sistematis yang digunakan dalam melakukan analisis dan menyelesaikan masalah yang telah diidentifikasi sebelumnya.
- d. BAB IV Penyelesaian Permasalahan, berisi penjelasan terkait proses analisis yang mencakup pengumpulan data, pemahaman data, dan preprocessing data, data mining yang mencakup proses modelling untuk klasifikasi topik dan analisis sentimen, diikuti dengan melakukan *topic modelling* untuk memperoleh informasi terkait subtopik untuk setiap topik. Selain itu dilakukan evaluasi dan visualisasi data untuk mengevaluasi model.
- e. BAB V Validasi, Analisis, Hasil, dan Implikasi, menjelaskan hasil dari pembuatan model klasifikasi topik dan analisis sentimen dengan menggunakan model CatBoost, serta *topic modelling* dengan menggunakan model BERTopic. Evaluasi model yang telah dikembangkan juga dijelaskan dengan menggunakan *confusion matrix*, *classification report*, dan *hierarchical clustering*, serta *Coherence Score*.
- f. BAB VI Kesimpulan dan Saran, menjelaskan ringkasan hasil yang diperoleh dari penelitian yang dilakukan serta saran yang dapat menjadi rekomendasi bagi peneliti selanjutnya untuk dapat mengembangkan penelitian dan memperbaiki kelemahan yang terdapat pada penelitian.