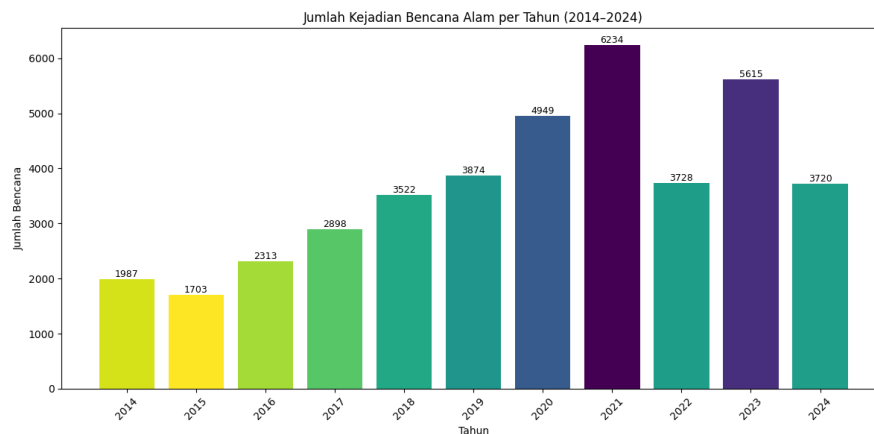


BAB I PENDAHULUAN

I.1 Latar Belakang

Indonesia merupakan negara yang terletak di pertemuan tiga lempeng tektonik besar, yaitu lempeng Indo-Australia, Eurasia, dan Pasifik, sehingga menjadikan wilayah ini sangat rawan terhadap berbagai jenis bencana alam seperti gempa bumi, tsunami, letusan gunung api, banjir, tanah longsor, serta cuaca ekstrem. Selain itu, Indonesia memiliki iklim tropis basah dengan curah hujan yang tinggi, yang turut berkontribusi terhadap risiko bencana hidrometeorologi seperti banjir dan tanah longsor. Kondisi geografis dan iklim tersebut menyebabkan Indonesia menjadi salah satu negara dengan frekuensi bencana alam tertinggi di dunia.

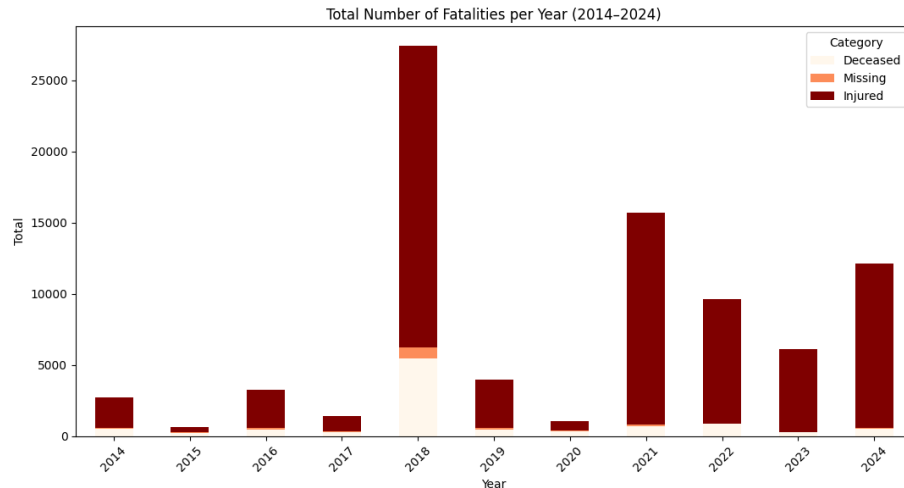
Perubahan iklim global yang dipicu oleh peningkatan emisi gas rumah kaca juga memperburuk kerentanan Indonesia terhadap bencana alam. Dampak perubahan iklim meningkatkan intensitas dan frekuensi kejadian bencana seperti banjir bandang, kekeringan, dan gelombang ekstim yang berpotensi menimbulkan kerusakan lingkungan, kerugian ekonomi, serta dampak sosial dan kesehatan yang signifikan bagi masyarakat (Haryanto dkk., 2019).



Gambar I-1. Total kejadian bencana di Indonesia selama satu dekade
Sumber data: gis.bpb.go.id

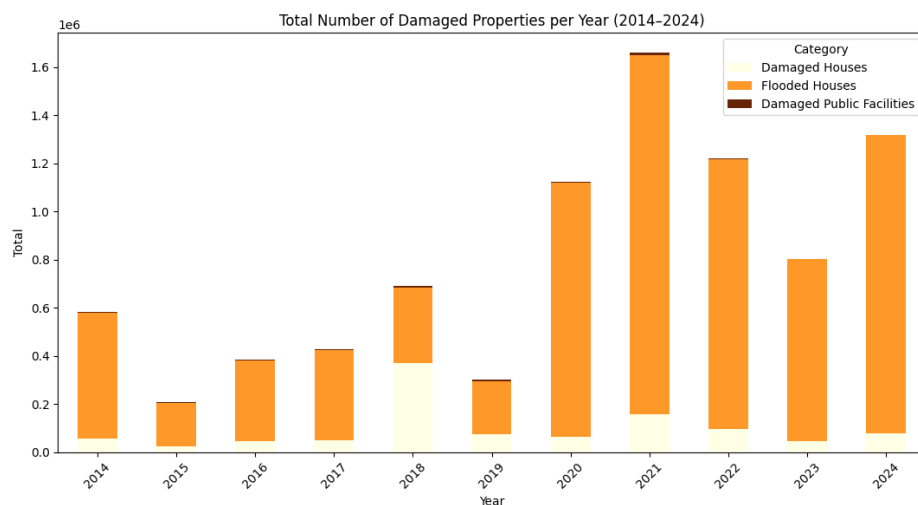
Berdasarkan data dari Badan Nasional Penanggulangan Bencana (BNPB), sepanjang tahun 2014 hingga 2024, terjadi fluktuasi signifikan dalam jumlah kejadian bencana alam di Indonesia. Seperti yang ditunjukkan pada Gambar I-1, tren jumlah kejadian bencana menunjukkan peningkatan yang cukup tajam sejak tahun 2015, dengan puncaknya terjadi pada tahun 2021 sebanyak 6.234 kejadian,

sebelum menurun secara bertahap di tahun-tahun berikutnya. Bencana yang paling umum terjadi selama periode ini adalah banjir, cuaca ekstrem, dan tanah longsor, yang termasuk dalam kategori bencana hidrometeorologi.



Gambar I-2. Jumlah korban jiwa akibat bencana per tahun
Sumber data: gis.bpb.go.id

Seperti yang terlihat pada Gambar I-2, bencana alam dalam periode ini telah menimbulkan dampak signifikan terhadap korban jiwa, termasuk korban meninggal dunia, hilang, dan terluka. Tahun 2018 mencatat jumlah korban tertinggi, yang mencapai lebih dari 27.000 jiwa, dengan proporsi terbesar adalah korban terluka. Tahun-tahun lain seperti 2021 dan 2024 juga menunjukkan jumlah korban yang relatif tinggi, mencerminkan tingginya tingkat keparahan bencana yang terjadi pada tahun-tahun tersebut.



Gambar I-3. Jumlah kerusakan akibat bencana per tahun
Sumber data: gis.bpb.go.id

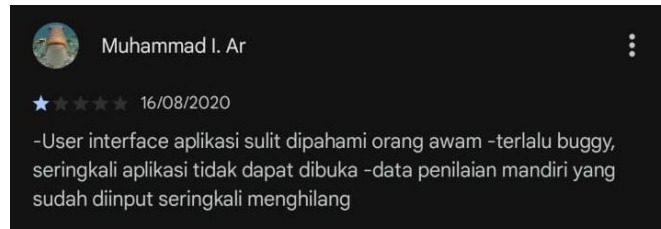
Dampak bencana tidak hanya dirasakan secara fisik oleh individu, tetapi juga secara ekonomi dan sosial melalui kerusakan infrastruktur. Seperti ditunjukkan pada Gambar I-3, jumlah rumah rusak, rumah terendam, dan fasilitas umum (fasum) rusak meningkat secara signifikan terutama sejak tahun 2018. Tahun 2021 kembali menjadi tahun dengan tingkat kerusakan tertinggi, dengan lebih dari 1,6 juta unit kerusakan tercatat, sebagian besar berupa rumah terendam akibat banjir besar di berbagai wilayah.

Menghadapi situasi yang semakin kompleks akibat bencana alam berbasis iklim, dibutuhkan salah satu aspek komunikasi yang dikenal sebagai *risk communication*. Ketika masyarakat menginginkan komunikasi mengenai risiko dan dampak yang akan terjadi, mereka akan lebih memilih untuk menerima pesan atau peringatan yang jelas dan mudah dipahami tentang risiko tersebut (Frewer, 2004).

Salah satu cara praktis dan akurat untuk memperoleh informasi terkini terkait keadaan darurat adalah dengan penggunaan aplikasi mitigasi bencana. Tujuan utamanya adalah untuk mengurangi kesalahan dalam penyampaian informasi yang dapat menyebabkan kebingungan dan tidak pasti (Putri & Hamzah, 2021). Aplikasi-aplikasi ini menyediakan beragam fitur yang memperkuat kesiapsiagaan sebelum, selama, dan setelah terjadinya bencana (Mahfudz, 2019)

Meskipun Indonesia telah memiliki beberapa aplikasi mitigasi bencana yang berguna seperti Info BMKG, Info Gempa Bumi Terkini, WRS-BMKG, EQInfo, dan sebagainya, serta aplikasi khusus InaRISK dari Badan Nasional Penanggulangan Bencana (BNPB), tetapi sayangnya, kerusakan fasilitas umum, rumah, korban luka, korban hilang, hingga korban jiwa akibat bencana belum menunjukkan penurunan yang signifikan dalam 10 tahun terakhir.

Aplikasi InaRISK berfungsi sebagai portal untuk mengkaji risiko bencana. Dengan memanfaatkan ArcGIS *server* sebagai sumber data, InaRISK menyajikan informasi mengenai wilayah rawan bencana, populasi terdampak, potensi kerugian, dan dampak kerusakan lingkungan. Aplikasi ini memiliki tujuan untuk menurunkan indeks risiko bencana sehingga memberikan langkah yang efektif dalam memitigasi bencana (Khusna dkk., 2022).

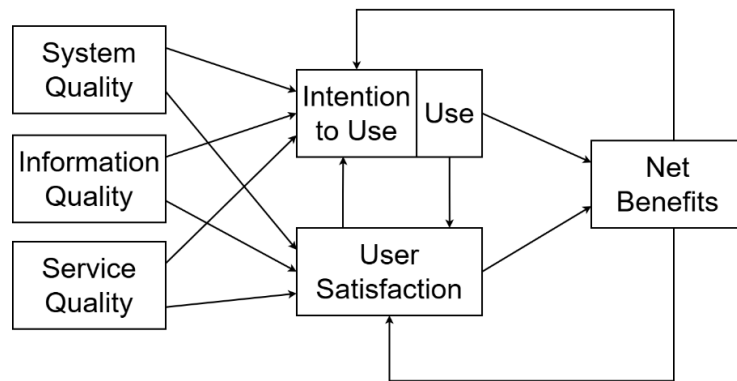


Gambar I-4. Ulasan negatif aplikasi InaRISK

Berdasarkan *review* yang dapat dilihat pada Gambar I-4 di atas, dapat dilihat bahwa tujuan aplikasi InaRISK dan kenyataan yang terjadi di Indonesia menunjukkan bahwa informasi yang terdapat dalam aplikasi memang penting dan bermanfaat. Akan tetapi, masyarakat belum dapat menangkap atau memahami informasi yang disediakan oleh aplikasi InaRISK dengan baik. Oleh karena itu, perlu dilakukan analisis terhadap aplikasi supaya dapat diketahui aspek apa saja yang perlu ditingkatkan untuk dapat menurunkan indeks risiko bencana.

Salah satu cara untuk menganalisis perspektif masyarakat terhadap aplikasi adalah melalui ulasan, yang telah menjadi salah satu cara yang populer untuk pengguna agar dapat mengungkapkan pendapat dan pengalaman mereka terhadap berbagai aplikasi layanan (Karo Karo dkk., 2023). Untuk menganalisis ulasan, diperlukan teknik yang disebut analisis sentimen atau penambangan opini, untuk memahami, mengekstrak, dan memproses data teks secara otomatis guna mendapatkan informasi mengenai sentimen yang terdapat dalam sebuah pernyataan opini. Pertumbuhan pesat dalam penelitian dan penerapan berbasis analisis sentimen menunjukkan dampak besar dan manfaatnya (Buntoro, 2017).

Meskipun analisis sentimen telah banyak diterapkan dalam berbagai konteks aplikasi layanan publik dan kebencanaan, hingga saat ini belum ditemukan penelitian yang secara spesifik membahas aplikasi InaRISK menggunakan pendekatan analisis sentimen. Hal ini menunjukkan adanya *research gap* yang penting, mengingat peran sentral InaRISK sebagai *platform* nasional dalam mitigasi risiko bencana. Oleh karena itu, penelitian ini bertujuan untuk mengisi kekosongan tersebut dengan menganalisis ulasan pengguna terhadap aplikasi InaRISK menggunakan pendekatan analisis sentimen dan pemodelan topik.



Gambar I-5 *Information system success model*
Sumber: (Petter dkk., 2013)

Untuk menganalisis ulasan tersebut secara terstruktur, penelitian ini menggunakan kerangka *Information System Success Model* (Gambar I-5) yang dikembangkan oleh Delone & Mclean (2003). Model ini telah banyak digunakan dalam berbagai konteks untuk mengevaluasi keberhasilan sistem informasi dengan enam komponen utama: *Information Quality*, *System Quality*, *Service Quality*, *Intention to Use*, *User Satisfaction*, dan *Net Benefits* (Delone & Mclean, 2003; Petter dkk., 2013). Dalam konteks aplikasi InaRISK, model ini relevan karena memberikan pemahaman menyeluruh mengenai persepsi pengguna terhadap sistem serta bagaimana persepsi tersebut memengaruhi niat penggunaan dan manfaat jangka panjang aplikasi.

Penelitian ini memfokuskan analisis pada tiga dimensi utama, yaitu kualitas informasi, kualitas sistem, dan kualitas layanan, yang dipilih karena merupakan faktor utama yang memengaruhi niat penggunaan dan kepuasan pengguna (Delone & Mclean, 2003; Petter dkk., 2013):

1. Kualitas informasi mencerminkan seberapa akurat, relevan, lengkap, dan mudah dipahami informasi yang disediakan oleh aplikasi, yang secara langsung memengaruhi kepuasan dan niat penggunaan pengguna (Delone & Mclean, 2003; Rai dkk., 2002).
2. Kualitas sistem mencakup aspek performa teknis dan desain aplikasi, termasuk kecepatan akses, kemudahan penggunaan, dan tampilan antarmuka (UI/UX). Kualitas sistem yang baik meningkatkan pengalaman pengguna dan secara signifikan memengaruhi kepuasan serta penggunaan sistem (Bakhit Jaafreh, 2017; Petter dkk., 2013).

3. Kualitas layanan berkaitan dengan seberapa baik dukungan atau respons yang diberikan oleh pengembang atau pengelola aplikasi terhadap pengguna. Kualitas layanan yang tinggi berkontribusi pada kepuasan pengguna dan niat penggunaan yang lebih besar (Delone & Mclean, 2003; Parasuraman dkk., 1985).

Penelitian oleh Sarifiyono & Lesmana (2023) serta Irma Hartiwi & Rokhayati (2024) menunjukkan bahwa beberapa kata kunci umum digunakan dalam artikel jurnal untuk mengoperasionalkan dan mengukur dimensi kualitas dalam konteks keberhasilan sistem informasi (*Information System Success*). Adapun dimensi-dimensi kualitas beserta indikatornya adalah sebagai berikut:

1. Kualitas Informasi : keakuratan, kelengkapan, ketepatan waktu, kemudahan dipahami, *up-to-date* (terbaru), format *output*, *semantic success* (keberhasilan makna).
2. Kualitas Sistem : fleksibilitas, integrasi, kemudahan pengguna, keandalan, fitur intuitif, kecepatan respon sistem, dan *technical success* (keberhasilan teknis).
3. Kualitas Layanan : respon *customer service*, jenis layanan yang ditawarkan, dukungan fisik, kesesuaian dengan keinginan dan kebutuhan pengguna, sikap dan empati petugas layanan.

Dalam konteks analisis sentimen, *Support Vector Machine* dapat dengan mudah dan efisien untuk mengklasifikasikan data sebagai label positif atau negatif berdasarkan fitur yang diekstrak dari data yang berupa teks (Haranto & Sari, 2019). Metode ini dipilih karena berdasarkan penelitian yang dilaksanakan Karo Karo dkk. (2023) dan penulis lainnya pada ulasan infoBMKG dan Buntoro (2017) pada unggahan di Twitter, SVM merupakan algoritma yang menghasilkan akurasi yang cukup tinggi. Ditambah dengan hasil penelitian analisis sentimen opini terhadap vaksin Covid-19 pada media sosial Twitter (F. Fitriana dkk., 2021), yang juga menggunakan *Support Vector Machine*, model ini efektif untuk digunakan terhadap *dataset* yang sedikit dan sangat kecil kemungkinan untuk terjadi *overfitting*.

Setelah pemodelan sentimen dengan menggunakan *Support Vector Machine*, ulasan akan diklasifikasikan lebih lanjut dengan menggunakan algoritma LDA. *Latent*

Dirichlet Allocation (LDA) dipilih karena memiliki kemampuan yang lebih baik daripada *Latent Semantic Analysis* (LSA) dalam hal klasifikasi topik (Hassan dkk., 2020). LDA akan membantu mengidentifikasi sub-aspek dari kategori utama. Dengan demikian, analisis sentimen tidak hanya sebatas klasifikasi positif atau negatif, tetapi juga menangkap dimensi informasi yang lebih mendalam tentang kekurangan atau kekecewaan pengguna terhadap suatu produk atau layanan.

I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, rumusan masalah yang mendasari penelitian ini adalah:

1. Bagaimana persebaran sentimen ulasan pengguna terhadap aplikasi InaRISK dapat diklasifikasikan ke dalam tiga aspek utama berdasarkan *IS Success Model*, yaitu kualitas informasi, kualitas sistem, dan kualitas layanan?
2. Bagaimana persebaran ulasan pengguna aplikasi InaRISK pada versi aplikasi yang paling banyak menerima ulasan negatif dan paling banyak ulasan positif?
3. Bagaimana penerapan dan tingkat koherensi metode *Latent Dirichlet Allocation* (LDA) dengan pendekatan *TF-IDF* dalam mengidentifikasi topik utama dari masing-masing aspek kualitas?
4. Bagaimana performa algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen negatif pada setiap aspek kualitas aplikasi InaRISK?

I.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menganalisis sentimen masyarakat terhadap aplikasi InaRISK untuk mengetahui aspek yang tidak memuaskan. Berpegang pada rumusan masalah sebelumnya yang telah dilandasi oleh latar belakang, maka tujuan dari penelitian ini adalah:

1. Mengklasifikasikan sentimen ulasan pengguna aplikasi InaRISK ke dalam tiga aspek utama berdasarkan kerangka *IS Success Model*, yaitu kualitas informasi, kualitas sistem, dan kualitas layanan.
2. Menganalisis persebaran ulasan pengguna pada berbagai versi aplikasi InaRISK, serta mengidentifikasi versi aplikasi yang memperoleh ulasan paling banyak dalam kategori positif maupun negatif.

3. Menerapkan metode *Latent Dirichlet Allocation* (LDA) dengan pendekatan TF-IDF untuk mengidentifikasi topik utama dari masing-masing, serta mengevaluasi tingkat koherensinya menggunakan nilai *coherence score*.
4. Mengevaluasi performa algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen pada tiga aspek kualitas InaRISK, serta menentukan konfigurasi pelatihan yang menghasilkan performa terbaik.

I.4 Batasan Penelitian

Penelitian ini memiliki sejumlah batasan untuk menjaga fokus dan kedalaman analisis agar tetap sesuai dengan tujuan yang telah ditetapkan. Oleh karena itu, penelitian ini dibatasi dengan cakupan sebagai berikut:

1. Data ulasan yang digunakan dalam penelitian diperoleh dari *platform* Google Play Store (Oktober 2017–Februari 2025), App Store (Januari 2019–Oktober 2023), serta Twitter (Agustus 2014–Maret 2025), dengan cakupan ulasan berbahasa Indonesia dan Inggris.
2. Identifikasi faktor ketidakpuasan pengguna difokuskan pada ulasan yang berkaitan dengan performa dan kualitas aplikasi InaRISK, dengan tujuan memberikan masukan bagi pengembangan aplikasi, khususnya dalam aspek penyampaian informasi kebencanaan.
3. Analisis sentimen dilakukan berdasarkan tiga aspek utama dalam kerangka *IS Success Model*, yaitu kualitas informasi, kualitas sistem, dan kualitas layanan
4. Analisis subtopik pada masing-masing aspek dilakukan menggunakan metode *Latent Dirichlet Allocation* (LDA) berbasis TF-IDF, setelah proses klasifikasi sentimen dilakukan.

I.5 Manfaat Penelitian

Penelitian ini berkontribusi dalam pengembangan aplikasi kebencanaan berbasis teknologi. Setiap pihak yang terkait dapat memanfaatkan hasil penelitian ini sebagai landasan dalam pengambilan keputusan, pengembangan strategi, atau eksplorasi keilmuan lanjutan. Berikut adalah manfaat penelitian ini berdasarkan masing-masing bidang:

I.5.1 Bagi Akademisi

Penelitian ini memberikan gambaran implementasi metode analisis data yang telah dipelajari dalam perkuliahan, khususnya metode *text mining*, TF-IDF, LDA, dan algoritma *Support Vector Machine* (SVM). Dengan studi kasus pada aplikasi kebencanaan InaRISK, akademisi dapat melihat relevansi antara teori dan praktik dalam konteks pemanfaatan data ulasan pengguna.

Selain itu, penelitian ini juga memperluas literatur dalam bidang analisis sentimen berbasis *Information Systems Success Model*. Hal ini dapat dijadikan referensi untuk pengembangan kurikulum atau bahan ajar di bidang sistem informasi, kecerdasan buatan, dan mitigasi bencana berbasis teknologi.

I.5.2 Bagi Industri

Industri yang bergerak dalam ranah kebencanaan dapat memperoleh masukan berupa rekomendasi peningkatan layanan berdasarkan analisis sentimen pengguna. Hasil penelitian ini membantu memahami persepsi pengguna terhadap kualitas informasi, sistem, dan layanan yang diberikan oleh aplikasi InaRISK.

Dengan pemahaman yang lebih baik terhadap kebutuhan dan ekspektasi pengguna, perusahaan dapat menyusun strategi pengembangan yang lebih responsif dan tepat sasaran. Hal ini pada akhirnya akan mendukung tercapainya tujuan utama aplikasi InaRISK, yaitu membantu mengurangi dampak negatif dari bencana.

I.5.3 Bagi Peneliti Selanjutnya

Laporan ini memberikan informasi tentang penelitian yang terintegrasi antara *machine learning*, *topic modelling*, dan *natural language processing* dalam konteks aplikasi kebencanaan. Struktur metodologi dan tahapan analisis yang digunakan dapat dijadikan acuan bagi penelitian selanjutnya dalam bidang yang serupa.

Selain sebagai referensi, penelitian ini juga membuka peluang eksplorasi lanjutan terhadap pendekatan algoritmik yang lebih kompleks, penggunaan data dengan domain yang berbeda, atau pengembangan sistem rekomendasi berbasis masukan pengguna. Dengan demikian, penelitian ini berkontribusi dalam membangun fondasi penelitian yang berkelanjutan di bidang sistem informasi publik dan mitigasi risiko bencana.

I.6 Sistematika Penulisan

Tugas akhir ini disusun secara sistematis agar pembaca dapat memahami alur pemikiran, metode yang digunakan, serta hasil dan kontribusi penelitian secara menyeluruh. Sistematika penulisan terdiri dari enam bab utama, yang dijabarkan sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang yang menjadi dasar pentingnya penelitian, terutama terkait banyaknya ulasan negatif terhadap aplikasi InaRISK. Selanjutnya, rumusan masalah disusun sebagai panduan penelitian untuk melakukan analisis.

Tujuan penelitian kemudian akan menjelaskan hasil konkret yang ingin dicapai, sedangkan batasan penelitian disusun untuk memperjelas ruang lingkup dan fokus analisis, termasuk sumber data yang digunakan. Bab ini ditutup dengan penjelasan mengenai manfaat penelitian, yang ditujukan bagi akademisi, praktisi industri, dan peneliti selanjutnya.

BAB II TINJAUAN PUSTAKA

Bab ini memuat teori-teori yang mendasari penelitian, antara lain teori analisis sentimen, metode TF-IDF, SVM, dan LDA, serta teori lainnya yang digunakan untuk pemrosesan data dan ekstraksi topik. Pemahaman teori ini penting untuk memastikan metode yang digunakan relevan dengan tujuan penelitian.

Selain itu, bagian ini menyajikan kajian literatur dari penelitian sebelumnya yang relevan. Tujuannya adalah untuk menunjukkan posisi penelitian ini dalam konteks keilmuan yang lebih luas, sekaligus mengidentifikasi celah penelitian yang ingin diisi oleh tugas akhir ini.

BAB III METODOLOGI PENYELESAIAN MASALAH

Bab III akan menjelaskan mengenai langkah-langkah sistematis dalam menyelesaikan masalah. Penjelasan akan memuat informasi mengenai *Knowledge Discovery in Databases* yang merupakan proses pengumpulan dan pengolahan data, serta metode evaluasi yang diterapkan dalam penelitian ini.

BAB IV PENYELESAIAN MASALAH

Bab ini menyajikan detail proses pengolahan data dan eksperimen yang dilakukan. Dimulai dari tahapan pembersihan dan persiapan data teks, *splitting data*, kemudian dilakukan transformasi data ke bentuk numerik menggunakan TF-IDF agar dapat diproses oleh algoritma *machine learning*.

Setelah itu, dilakukan klasifikasi sentimen untuk setiap aspek kualitas menggunakan algoritma SVM, serta pengelompokan ulasan negatif dalam tiap aspek menjadi subtopik dengan metode LDA. Hasil dari bab ini akan menjadi dasar untuk mengevaluasi dan memahami persepsi pengguna secara lebih terstruktur dan mendalam.

BAB V VALIDASI, ANALISIS, HASIL, DAN IMPLIKASI

Bab ini menampilkan hasil pengujian dan evaluasi terhadap model yang telah diterapkan. Performa model SVM dinilai menggunakan metrik seperti *precision*, *recall*, dan *F1-score* untuk mengidentifikasi kombinasi parameter terbaik dalam membedakan sentimen negatif pada tiap aspek kualitas.

Di sisi lain, hasil topik dari LDA dievaluasi menggunakan *coherence score* untuk mengetahui sejauh mana kualitas topik yang terbentuk. Bab ini juga menampilkan visualisasi berupa grafik dan tabel untuk memperjelas interpretasi data dan memudahkan pembaca memahami temuan penelitian.

BAB VI KESIMPULAN DAN SARAN

Bab ini merangkum jawaban atas rumusan masalah berdasarkan hasil penelitian yang telah dilakukan. Kesimpulan yang disusun mencakup temuan utama terkait persepsi pengguna terhadap aplikasi InaRISK berdasarkan aspek kualitas informasi, sistem, dan layanan.

Selain itu, disampaikan pula saran pengembangan aplikasi berdasarkan hasil klasifikasi dan topik yang diperoleh. Rekomendasi ini diharapkan dapat membantu pengembang dalam meningkatkan kualitas dan efektivitas aplikasi InaRISK sebagai alat mitigasi bencana yang lebih adaptif terhadap kebutuhan pengguna.