

Implementasi *Chatbot* Cuaca Berbasis SBERT dan LLM *Gemini* Menggunakan *OpenWeather* API

Hanif Luthfi Irfanudin¹, Mohamad Yusuf^{2*}

^{1,2*} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia.

Corresponding Email: mhd.yusuf@mercubuana.ac.id^{2*}

Histori Artikel:

Dikirim 22 November 2025; *Diterima dalam bentuk revisi* 20 Desember 2025; *Diterima* 1 Januari 2025; *Diterbitkan* 10 Januari 2026. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Penelitian ini membahas pengembangan dashboard chatbot cuaca berbasis Natural Language Processing (NLP) yang mampu menerima query berbahasa Indonesia secara alami dan menampilkan informasi cuaca secara interaktif. Sistem mengombinasikan Sentence-BERT (SBERT) paraphrase-multilingual-MiniLM-L12-v2 sebagai mesin kesamaan teks, Large Language Model (LLM) Gemini 2.5-flash sebagai generator ringkasan bahasa alami, serta OpenWeather API sebagai sumber data cuaca real-time. Pendekatan zero-shot semantic similarity digunakan tanpa fine-tuning, dengan penentuan intent berbasis cosine similarity dan threshold 0,5 yang dioptimalkan melalui pengujian beberapa nilai ambang. Metode pengembangan yang digunakan adalah model Waterfall dengan tahapan analisis kebutuhan, perancangan arsitektur, implementasi modul NLP dan integrasi API, pengembangan frontend React dashboard, serta pengujian black-box fungsional dan evaluasi metrik performa. Hasil pengujian menunjukkan bahwa SBERT dengan threshold 0,5 menghasilkan akurasi klasifikasi intent sebesar 90% pada 20 skenario uji, yang meningkat menjadi 100% setelah dikombinasikan dengan mekanisme rule-based auto-adjust dan fallback. Ekstraksi entitas kota dengan sliding window 1–3 kata terhadap `city.list.json` mencapai akurasi 100% pada 15 query berkota, sementara metrik makro menghasilkan precision 0,9286, recall 0,90, dan F1-score 0,8958. Integrasi OpenWeather API dan Gemini memungkinkan penyajian ringkasan cuaca yang natural, informatif, serta visualisasi data cuaca real-time dalam bentuk grafik interaktif pada dashboard berbasis React.

Kata Kunci: Chatbot Cuaca; SBERT; Cosine Similarity; LLM Gemini; OpenWeather API.

Abstract

This study discusses the development of a Natural Language Processing (NLP)-based weather chatbot dashboard capable of receiving natural Indonesian language queries and displaying weather information interactively. The system combines Sentence-BERT (SBERT) paraphrase-multilingual-MiniLM-L12-v2 as a text similarity engine, Gemini 2.5-flash Large Language Model (LLM) as a natural language summary generator, and the OpenWeather API as a real-time weather data source. A zero-shot semantic similarity approach is used without fine-tuning, with intent determination based on cosine similarity and a threshold of 0.5 optimized through testing several threshold values. The development method used is the Waterfall model with stages of requirements analysis, architectural design, NLP module implementation and API integration, React dashboard frontend development, as well as functional black-box testing and performance metric evaluation. The test results show that SBERT with a threshold of 0.5 produces an intent classification accuracy of 90% in 20 test scenarios, which increases to 100% after being combined with rule-based auto-adjust and fallback mechanisms. City entity extraction with a 1–3 word sliding window against `city.list.json` achieved 100% accuracy on 15 city-based queries, while macro metrics yielded a precision of 0.9286, a recall of 0.90, and an F1-score of 0.8958. The integration of the OpenWeather API and Gemini enables the presentation of natural, informative weather summaries and visualization of real-time weather data in the form of interactive graphs on a React-based dashboard.

Keyword: Weather Chatbot; BERT; Cosine Similarity; LLM Gemini; OpenWeathermAP.

1. Pendahuluan

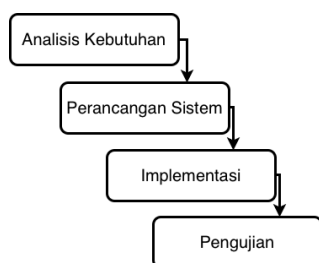
Cuaca memiliki pengaruh signifikan terhadap berbagai aktivitas manusia, mulai dari transportasi, pertanian, pariwisata, hingga mitigasi bencana. Informasi mengenai kondisi cuaca yang akurat dan mudah dipahami menjadi kebutuhan penting dalam kehidupan modern. Menurut Sopyandi dkk. (2024), penyajian data cuaca yang cepat dan terstruktur membantu pengguna dalam mengambil keputusan harian, khususnya di negara beriklim tropis seperti Indonesia yang memiliki pola musim hujan dan kemarau. Keberadaan layanan data terbuka seperti *OpenWeather API* memungkinkan pengembang untuk mengakses data cuaca waktu nyata (*real-time weather data*) yang dapat diintegrasikan ke dalam aplikasi berbasis web maupun *dashboard* interaktif (Bharath dkk., 2024). Perkembangan teknologi informasi dan komunikasi telah membuka peluang besar dalam penyampaian informasi berbasis data terbuka. Melalui integrasi antarmuka pemrograman aplikasi (*Application Programming Interface* atau API), data meteorologi kini dapat diolah dan divisualisasikan secara dinamis. Namun, tantangan utama dalam pengembangan sistem informasi cuaca terletak pada bagaimana sistem tersebut dapat memahami dan menanggapi pertanyaan pengguna yang diajukan dengan bahasa alami. Hal ini menjadi bidang yang terus berkembang dalam penelitian *Natural Language Processing* (NLP) (Amur dkk., 2023). Dalam beberapa tahun terakhir, kemajuan *machine learning* berbasis model transformer seperti *Bidirectional Encoder Representations from Transformers* (BERT) dan turunannya *Sentence-BERT* (SBERT) telah meningkatkan kemampuan komputer dalam memahami makna semantik teks. Model ini efektif digunakan dalam berbagai tugas seperti klasifikasi teks, *semantic similarity*, dan *intent classification* (Lee & Hsiang, 2020; Holis dkk., 2025). SBERT mampu menghasilkan representasi vektor kalimat (*sentence embedding*) yang lebih efisien dibandingkan pendekatan tradisional berbasis *bag-of-words* atau *TF-IDF* (Rashid & Ahmed, 2025). Di sisi lain, perkembangan *large language model* (LLM) generatif seperti Gemini menghadirkan kemampuan baru dalam menghasilkan respons bahasa alami yang lebih luwes, kontekstual, dan adaptif terhadap kebutuhan pengguna. Hardiman dkk. (2024) menunjukkan bahwa kombinasi SBERT dengan Gemini dapat meningkatkan kualitas dialog sistem berbasis bahasa alami karena keduanya saling melengkapi: SBERT unggul dalam penentuan kesamaan semantik, sementara Gemini unggul dalam generasi teks yang alami.

Meskipun telah banyak penelitian terkait pemanfaatan *chatbot* dan *dashboard* berbasis data cuaca, sebagian besar sistem masih mengandalkan interaksi berbasis menu atau formulir. Hal ini membatasi fleksibilitas pengguna dalam mengajukan pertanyaan dengan struktur kalimat bebas. Penelitian oleh Peyton dan Unnikrishnan (2023) menegaskan bahwa sistem *chatbot* yang tidak memanfaatkan pendekatan semantik sering kali gagal mengidentifikasi maksud pengguna ketika variasi bahasa meningkat, terutama pada bahasa non-Inggris seperti Bahasa Indonesia. Dengan demikian, diperlukan sistem yang mampu memahami makna kueri secara semantik tanpa bergantung pada pola kalimat tertentu. Selain itu, integrasi *zero-shot learning* menjadi pendekatan yang menjanjikan karena memungkinkan sistem melakukan klasifikasi intent tanpa memerlukan proses *fine-tuning* yang mahal secara komputasi (Owen & Suhartono, 2025). Pendekatan *zero-shot semantic similarity* dengan SBERT memungkinkan deteksi intent berbasis skor kesamaan semantik antara kueri pengguna dan kumpulan kalimat rujukan. Mekanisme ini dapat dikombinasikan dengan *rule-based auto-adjustment* untuk memperbaiki prediksi berdasarkan indikator linguistik seperti jumlah entitas kota yang terdeteksi atau keberadaan kata kunci perbandingan. Penelitian ini memfokuskan diri pada pengembangan *chatbot* cuaca berbasis SBERT dan LLM Gemini yang terintegrasi dengan *OpenWeather API*. Sistem yang dikembangkan bertujuan untuk: (1) mengklasifikasikan intent kueri cuaca berbahasa Indonesia secara otomatis, (2) mengekstraksi entitas kota secara presisi dari teks pengguna, (3) mengambil data cuaca waktu nyata dari *OpenWeather API*, serta (4) menghasilkan ringkasan dan visualisasi cuaca dalam bentuk *dashboard* berbasis React yang interaktif. Dengan menggabungkan FastAPI sebagai *backend service*, React sebagai antarmuka pengguna, SBERT sebagai mesin analisis semantik, serta Gemini sebagai generator respons, penelitian ini diharapkan dapat menghadirkan sistem yang mampu memahami bahasa alami pengguna secara lebih manusiawi. Pendekatan ini tidak hanya menekankan akurasi klasifikasi intent, tetapi juga menilai aspek keterpahaman hasil keluaran sistem oleh pengguna.

Berbeda dari sebagian penelitian sebelumnya yang mengandalkan *fine-tuning* dataset bahasa Inggris, penelitian ini menerapkan pendekatan *zero-shot* untuk Bahasa Indonesia dengan tujuan menilai sejauh mana model SBERT pralatih dapat beradaptasi tanpa pelatihan tambahan. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kinerja sistem dalam klasifikasi intent serta ketepatan ekstraksi entitas kota. Secara keseluruhan, penelitian ini berupaya menunjukkan bahwa integrasi antara SBERT, LLM Gemini, dan *OpenWeather API* dapat menghasilkan sistem *chatbot* cuaca yang mampu memahami pertanyaan pengguna dalam Bahasa Indonesia dan memberikan keluaran yang informatif, adaptif, serta efisien dalam konteks aplikasi berbasis web interaktif.

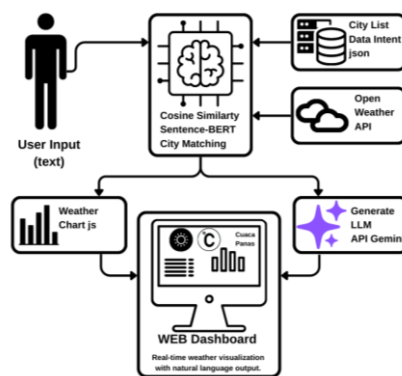
2. Metode Penelitian

Model pengembangan perangkat lunak yang digunakan adalah Waterfall, meliputi fase analisis kebutuhan, perancangan, implementasi, dan pengujian yang terdokumentasi secara berurutan.



Gambar 1. Model Proses Penelitian

Tahap analisis kebutuhan mencakup perumusan kebutuhan fungsional, meliputi pemrosesan kueri cuaca untuk satu kota, perbandingan dua kota, prakiraan cuaca beberapa hari ke depan, serta penanganan kueri di luar ranah cuaca. Tahap ini juga menetapkan kebutuhan nonfungsional yang berkaitan dengan kemudahan penggunaan dan waktu respons sistem. Tahap perancangan sistem menyusun arsitektur yang terdiri atas *frontend* React, *backend* FastAPI, modul *similarity engine* SBERT, integrasi *OpenWeather API*, dan modul LLM Gemini. Selain itu, pada tahap ini dilakukan penyusunan dataset intent berbahasa Indonesia serta pemilihan *city.list.json* sebagai basis referensi nama kota.



Gambar 2. Framework Arsitektur

Ekstraksi kota dibatasi pada pencocokan terhadap daftar *city.list.json*, sehingga variasi penulisan atau ambiguitas di luar daftar tersebut tidak dipertimbangkan dalam penelitian ini. Implementasi *Natural Language Processing* (NLP) menggunakan model *Sentence-BERT paraphrase-multilingual-MiniLM-L12-v2* untuk menghasilkan embedding dari kueri dan dataset intent. Klasifikasi intent dilakukan

dengan menghitung *cosine similarity* antara embedding kueri dan setiap label pada dataset intent, kemudian memilih skor tertinggi dengan penerapan ambang (*threshold*) kemiripan. Ambang awal 0,5 dipilih sebagai nilai moderat untuk menyeimbangkan sensitivitas klasifikasi dan risiko kesalahan pada pendekatan *zero-shot* tanpa *fine-tuning*. Nilai ambang ini kemudian divalidasi melalui pengujian beberapa nilai *threshold* lainnya (0,3; 0,4; 0,5; 0,6; 0,7) untuk menentukan nilai ambang yang optimal. Mekanisme *rule-based auto-adjust* diterapkan sebagai proses pasca-pemrosesan (*post-processing*) untuk mengoreksi intent berdasarkan jumlah kota yang terdeteksi, serta menangani kasus *low-confidence*. Sementara itu, ekstraksi kota dilakukan menggunakan teknik *sliding window* 1–3 kata terhadap daftar nama kota yang terdapat dalam *city.list.json*. Dataset intent berbahasa Indonesia disusun secara manual, terdiri dari empat kelas intent, yaitu *SINGLE_CITY*, *FORECAST_CITY*, *COMPARE_CITIES*, dan *UNKNOWN*. Secara keseluruhan, dataset ini mencakup sekitar 200 pola kalimat, dengan distribusi ± 90 pola untuk *COMPARE_CITIES*, ± 70 pola untuk *SINGLE_CITY*, ± 30 pola untuk *FORECAST_CITY*, dan ± 20 pola untuk *UNKNOWN*. Struktur dataset intent disusun dalam format JSON. Contoh pola kalimat yang digunakan disajikan pada Tabel 1.

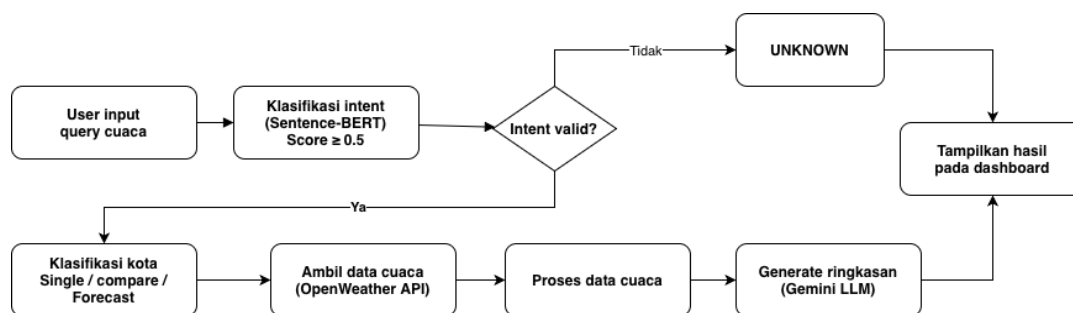
Tabel 1. Intent dataset

No	Teks Query	Label Intent
1	Gimana keadaan cuaca di <city1>?	SINGLE_CITY
2	Cuaca di kota <city1> bagaimana?	SINGLE_CITY
3	Prakiraan cuaca besok di <city1>?	FORECAST_CITY
4	Besok hujan tidak di <city1>?	FORECAST_CITY
5	Lebih panas mana sih <city1> sama <city2>?	COMPARE_CITIES
6	Bagaimana perbandingan cuaca <city1> dan <city2>?	COMPARE_CITIES
7	Siapa nama presiden Indonesia?	UNKNOWN
...		
200	Tolong buat puisi	UNKNOWN

Dataset intent disusun secara *template-based* menggunakan bahasa Indonesia sehari-hari dengan placeholder <city1> dan <city2> untuk generalisasi nama kota. Dataset ini terdiri dari sekitar 200 pola kalimat dan digunakan sebagai referensi embedding dalam pendekatan *zero-shot* SBERT, tanpa *fine-tuning*. Hasil klasifikasi intent yang dihasilkan oleh SBERT digunakan sebagai prediksi awal. Selanjutnya, penyesuaian dilakukan menggunakan aturan yang tercantum pada Tabel 2 melalui mekanisme *rule-based* sebagai proses *post-processing*. Tahapan ini diterapkan setelah perhitungan nilai *cosine similarity* untuk mengoreksi prediksi ketika hasil awal belum sesuai dengan konteks kueri pengguna, sehingga menghasilkan *Final Intent* yang lebih akurat.

Tabel 2. Mekanisme Rule-Based Auto-Adjust Intent

No	Rule	Kondisi	Final Intent	Contoh
1	Kata kunci perbandingan	Query mengandung kata <i>bandingkan, vs, versus, antara, beda</i>	COMPARE_CITIES	“Bandingkan cuaca Jakarta vs Bandung”
2	Jumlah kota ≥ 2	Jumlah kota terdeteksi ≥ 2 melalui ekstraksi berbasis <i>sliding window</i>	COMPARE_CITIES	“Cuaca Jakarta dan Bandung panas mana?”
3	Satu kota terdeteksi	Jumlah kota terdeteksi = 1	SINGLE_CITY	“Cuaca di Jakarta sekarang?”
4	Fallback UNKNOWN	Tidak terdeteksi kota, tidak ada kata kunci perbandingan, score $< 0,5$	UNKNOWN	“Buat puisi sedih?”



Gambar 3. Flowchart sistem

OpenWeather API diintegrasikan melalui endpoint proxy untuk data cuaca saat ini (*current weather*) dan prakiraan cuaca 5 hari setiap 3 jam (*5-day/ 3-hour forecast*), yang digunakan untuk mengelola intent *SINGLE_CITY*, *FORECAST_CITY*, dan *COMPARE_CITIES*. Data cuaca yang diperoleh kemudian diformat dan diringkas menggunakan Gemini 2.5-flash, dengan prompt yang dirancang khusus untuk menghasilkan penjelasan singkat, ramah, dan informatif dalam bahasa Indonesia. Antarmuka *frontend* dashboard dibangun dengan React dan Chart.js, yang menampilkan input kueri, ringkasan cuaca, serta grafik tren suhu dan perbandingan cuaca antar kota. Pengujian dilakukan dengan metode *black-box testing*, menggunakan 20 skenario uji yang mencakup empat kelas intent (*SINGLE_CITY*, *FORECAST_CITY*, *COMPARE_CITIES*, *UNKNOWN*), dengan pengukuran metrik *accuracy*, *precision*, *recall*, dan *F1-score* (makro), serta akurasi ekstraksi kota. Evaluasi *threshold similarity* dilakukan dengan menghitung performa klasifikasi intent pada berbagai nilai ambang (*threshold*) dan menganalisis pengaruhnya terhadap jumlah kueri yang diklasifikasikan sebagai *UNKNOWN*.

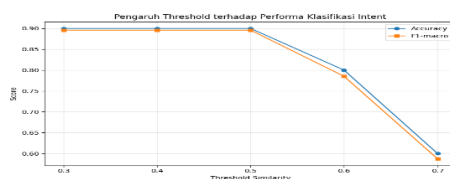
Tabel 3. Tabel pengujian

Jenis Intent	Jumlah Pengujian	Keterangan
SINGLE_CITY	5	Menguji data cuaca kota tunggal
COMPARE_CITY	5	Menguji perbandingan 2 kota
FORECAST_CITY	5	Menguji prakiraan 5 hari
UNKNOWN	5	Menguji fallback Gemini
TOTAL	20	

3. Hasil dan Pembahasan

3.1 Hasil

Implementasi *dashboard* cuaca menghasilkan antarmuka berbasis React yang memungkinkan pengguna untuk memasukkan kueri dalam bahasa Indonesia dan memperoleh ringkasan cuaca serta visualisasi grafik untuk satu kota, prakiraan beberapa hari, atau perbandingan antara dua kota. Modul *backend* FastAPI mengorkestrasi pemanggilan *SBERT similarity engine*, ekstraksi kota, integrasi *OpenWeather API*, dan pemanggilan Gemini LLM, sehingga seluruh alur, mulai dari kueri hingga visualisasi, berjalan secara otomatis.



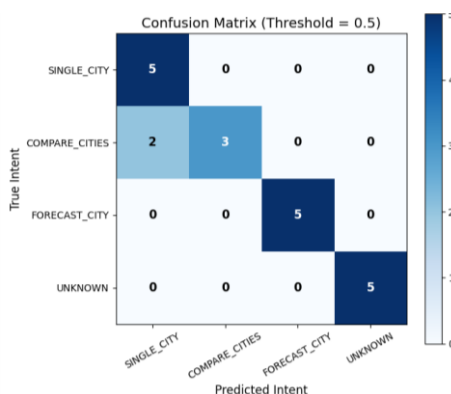
Gambar 4. Confusion matrix

Pengujian *threshold similarity* menunjukkan bahwa nilai antara 0,3 hingga 0,5 memberikan performa klasifikasi intent yang stabil, dengan akurasi 0,90 dan *F1-macro* sekitar 0,8958. Sebaliknya, peningkatan *threshold* menjadi 0,6 dan 0,7 mengakibatkan penurunan akurasi, masing-masing menjadi 0,80 dan 0,60, karena lebih banyak kueri yang diklasifikasikan sebagai *UNKNOWN*. Berdasarkan hasil tersebut, *threshold* 0,5 dipilih sebagai nilai optimal karena menyeimbangkan sensitivitas dan ketepatan klasifikasi sebelum penerapan mekanisme *rule-based auto-adjust*.

Tabel 4. Performa klasifikasi berdasar threshold similarity

Threshold	Accuracy	Precision (macro)	Recall (macro)	F1-macro
0.3	0.90	0.9286	0.90	0.8958
0.4	0.90	0.9286	0.90	0.8958
0.5	0.90	0.9286	0.90	0.8958
0.6	0.80	0.8562	0.80	0.7852
0.7	0.60	0.8462	0.60	0.5873

Pada 20 skenario uji representatif yang mencakup empat kelas intent, SBERT dengan *threshold* 0,5 mencapai akurasi klasifikasi intent sebesar 90% (18/20) sebelum koreksi berbasis aturan (*rule-based*). Dua kasus *COMPARE_CITIES* awalnya salah diprediksi sebagai *SINGLE_CITY*, namun berhasil dikoreksi oleh aturan “ ≥ 2 kota \rightarrow *COMPARE_CITIES*”, sehingga akurasi akhir sistem hibrida SBERT + *rule-based* mencapai 100% (20/20), dengan akurasi ekstraksi kota 100% pada 15 kueri yang mengandung nama kota. Ringkasan metrik menunjukkan precision makro 0,9286, recall 0,90, *F1-macro* 0,8958, dan rata-rata *cosine similarity* 0,7199, yang mengindikasikan performa sistem yang kuat dalam pemrosesan kueri cuaca berbahasa Indonesia.



Gambar 5. Confusion matrix threshold similarity 0.5

Dari *confusion matrix* pada Gambar 5, kesalahan utama terjadi pada kueri perbandingan dua kota yang diprediksi sebagai *SINGLE_CITY* akibat kemiripan skor *similarity* yang tinggi. Pendekatan *zero-shot* tanpa *fine-tuning* kurang sensitif terhadap jumlah entitas kota, namun kesalahan tersebut berhasil dikoreksi oleh mekanisme *rule-based* yang berbasis jumlah kota. Oleh karena itu, *threshold* 0,5 dipilih sebagai nilai optimal.

Tabel 5. Hasil uji akurasi SBERT

No	Query	Intent True	Intent Pred	Cosine Max	Cities True	Cities Pred	Status
1	Cuaca sekarang di Jakarta gimana?	SINGLE_CITY	SINGLE_CITY	0.6548	Jakarta	Jakarta	Berhasil
2	Lagi hujan nggak di	SINGLE_CITY	SINGLE_CITY	0.7092	Bandung	Bandung	Berhasil

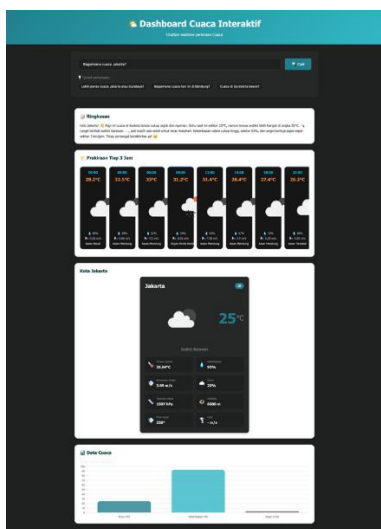
	Bandung?						
3	Info cuaca di Cikarang	SINGLE_CITY	SINGLE_CITY	0.7931	Cikarang	Cikarang	Berhasil
4	Kondisi Palembang sekarang gimana?	SINGLE_CITY	SINGLE_CITY	0.5593	Palembang	Palembang	Berhasil
5	Kuala Lumpur sekarang hujan tidak?	SINGLE_CITY	SINGLE_CITY	0.6765	Kuala Lumpur	Kuala Lumpur	Berhasil
6	Cuaca di Jakarta besok bagaimana?	FORECAST_CITY	FORECAST_CITY	0.7486	Jakarta	Jakarta	Berhasil
7	Bandung hujan nggak beberapa hari ke depan?	FORECAST_CITY	FORECAST_CITY	0.8590	Bandung	Bandung	Berhasil
8	Prakiraan cuaca Surabaya untuk 5 hari ke depan gimana?	FORECAST_CITY	FORECAST_CITY	0.6600	Surabaya	Surabaya	Berhasil
9	Besok pagi di Yogyakarta kemungkinan hujan nggak sih?	FORECAST_CITY	FORECAST_CITY	0.7701	Yogyakarta	Yogyakarta	Berhasil
10	Tokyo minggu ini akan hujan gak ya?	FORECAST_CITY	FORECAST_CITY	0.7063	Tokyo	Tokyo	Berhasil
11	Lebih panas cuaca Jakarta atau Surabaya?	COMPARE_CITIES	COMPARE_CITIES	0.7092	Jakarta, Surabaya	Jakarta, Surabaya	Berhasil
12	Cuaca Surabaya dan Malang sekarang beda nggak?	SINGLE_CITY*	COMPARE_CITIES	0.5723	Surabaya, Malang	Surabaya, Malang	Berhasil
13	Lebih dingin di Palembang atau Manokwari?	COMPARE_CITIES	COMPARE_CITIES	0.6866	Palembang, Manokwari	Palembang, Manokwari	Berhasil
14	Bandingkan cuaca Kuala Lumpur dengan Bandung	SINGLE_CITY*	COMPARE_CITIES	0.6077	Kuala Lumpur, Bandung	Kuala Lumpur, Bandung	Berhasil
15	Denpasar dan Mataram langitnya mana yang lebih cerah?	COMPARE_CITIES	COMPARE_CITIES	0.5727	Denpasar, Mataram	Denpasar, Mataram	Berhasil
16	Siapa presiden Indonesia?	UNKNOWN	UNKNOWN	0.8978	[]	[]	Berhasil
17	Ceritain lelucon dong	UNKNOWN	UNKNOWN	0.8436	[]	[]	Berhasil

18	Perlu bawa jaket nggak?	UNKNO WN	UNKNOWN	0.4532*	☐	☐	Berhasil
19	Kamu siapa?	UNKNO WN	UNKNOWN	0.9925	☐	☐	Berhasil
20	Halo, gimana kabarnya hari ini?	UNKNO WN	UNKNOWN	0.8268	☐	☐	Berhasil

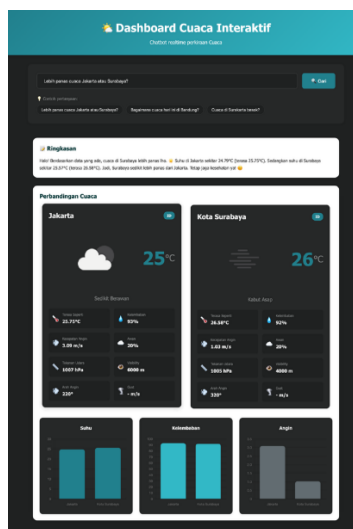
Catatan:

- 1) *Rule-based auto-adjust aktif*: ≥ 2 kota \rightarrow COMPARE_CITIES
- 2) *Score < 0,5, rule-based fallback aktif*: kueri diklasifikasikan sebagai UNKNOWN

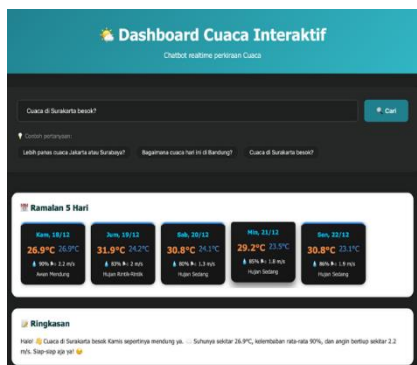
Pengujian fungsional juga menunjukkan bahwa sistem mampu: (1) menampilkan cuaca saat ini dan prakiraan lima hari untuk satu kota, (2) menyajikan perbandingan cuaca antara dua kota secara visual, dan (3) menangani kueri di luar ranah cuaca dengan mengalihkan ke respons generatif Gemini tanpa memanggil *OpenWeather API*. Integrasi Gemini 2.5-flash menghasilkan ringkasan cuaca yang alami dan mudah dipahami, sementara mekanisme *fallback* memastikan sistem tetap memberikan respons meskipun layanan LLM tidak tersedia. Kombinasi ini menunjukkan bahwa pendekatan hibrida SBERT *zero-shot*, *rule-based*, dan LLM generatif efektif dalam membangun *chatbot* cuaca edukatif berbasis *dashboard*.



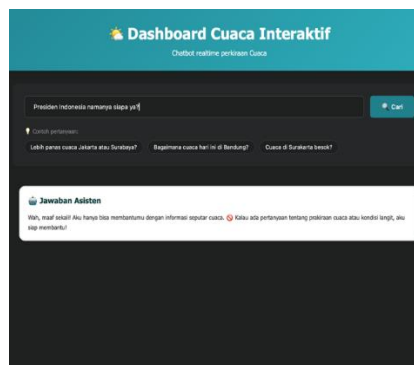
Gambar 6. Dashboard SINGLE_CITY



Gambar 7. Dashboard COMPARE_CITIES



Gambar 8. Dashboard FORECAST_CITY



Gambar 9. Dashboard UNKNOWN

3.2 Pembahasan

Dalam penelitian ini, pengembangan sistem chatbot cuaca berbasis *SBERT* dan *LLM Gemini* yang memanfaatkan *OpenWeather API* memberikan kontribusi penting terhadap kemajuan pemrosesan bahasa alami, khususnya dalam konteks interaksi cuaca berbasis bahasa Indonesia. Pendekatan yang digunakan mengintegrasikan berbagai teknologi terkini, seperti *Sentence-BERT* untuk kesamaan teks, *LLM Gemini* untuk pembuatan ringkasan bahasa alami, dan *OpenWeather API* untuk data cuaca real-time. Hasil yang diperoleh menunjukkan bahwa sistem ini mampu mengklasifikasikan intent dengan akurasi yang sangat baik, yaitu 90%, yang meningkat menjadi 100% setelah penerapan mekanisme rule-based dan fallback. Hal ini sejalan dengan temuan dalam penelitian oleh Abdul *et al.* (2025), yang memanfaatkan *Sentence-BERT* dan *cosine similarity* untuk pencarian semantik dokumen, serta Khadija *et al.* (2025) yang mengembangkan chatbot generatif berbasis *transformer* untuk layanan akademik. Sebagai tambahan, penggunaan model zero-shot *SBERT* tanpa *fine-tuning* terbukti efektif dalam mengklasifikasikan intent berbahasa Indonesia, yang mengurangi beban komputasi, sesuai dengan hasil penelitian Peyton dan Unnikrishnan (2023) yang menyoroti efisiensi model pre-trained dalam aplikasi chatbot. Integrasi antara *SBERT*, *Gemini*, dan *OpenWeather API* ini juga menunjukkan hasil yang menjanjikan dalam penyajian cuaca secara natural dan informatif, selaras dengan pendekatan yang diusulkan oleh Hardiman *et al.* (2024) dalam memanfaatkan model berbasis AI untuk interaksi dialog dalam konteks yang lebih luas. Keberhasilan pengujian ini menunjukkan bahwa pendekatan hybrid berbasis *NLP* ini dapat diterapkan secara luas dalam pengembangan sistem chatbot yang lebih canggih dan informatif.

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan sistem dashboard chatbot cuaca berbasis *NLP* yang mengintegrasikan *SBERT*-based text similarity, *rule-based auto-adjustment*, *LLM Gemini*, dan *OpenWeather API* dalam satu arsitektur end-to-end. Dengan threshold *cosine similarity* 0,5, modul *SBERT* mencapai akurasi klasifikasi intent 90% yang kemudian meningkat menjadi 100% setelah dikombinasikan dengan mekanisme *rule-based* dan *fallback*, sementara ekstraksi entitas kota mencapai akurasi 100% pada dataset uji. Dari sisi pengalaman pengguna, integrasi *React dashboard*, *OpenWeather API*, dan *Gemini* memungkinkan penyajian informasi cuaca real-time secara visual dan naratif yang mudah dipahami, sehingga mendukung tujuan edukatif dan demonstratif di bidang web interaktif dan *NLP*. Pendekatan *zero-shot* tanpa *fine-tuning* yang digunakan menunjukkan bahwa model *SBERT* pre-trained dapat dimanfaatkan secara efektif untuk *intent classification* bahasa Indonesia dengan sumber daya komputasi yang lebih ringan. Pengujian dengan jumlah skenario yang lebih besar serta *user acceptance test* belum dilakukan karena keterbatasan kuota *LLM*, dan direncanakan sebagai pengembangan lanjutan.

5. Daftar Pustaka

- Abdul, M., Fathuddin, H., Prakarsa Mandyartha, E., & Nurlaili, A. L. (2025). Penerapan *Sentence-BERT* dan *cosine similarity* untuk pencarian semantik dokumen skripsi dalam format PDF. *R2J*, 8(1). <https://doi.org/10.38035/rrj.v8i1>.
- Amur, Z. H., Kwang Hooi, Y., Bhanbhro, H., Dahri, K., & Soomro, G. M. (2023). Short-text semantic similarity (STSS): Techniques, challenges and future perspectives. *Applied Sciences (Switzerland)*, 13(6). MDPI. <https://doi.org/10.3390/app13063911>.

- Bharath, D., Dolok Lauro Sitorus, M., Teknologi Informasi, F., & Tarumanagara, U. (2024). Perancangan dashboard sebagai perbandingan visualisasi data iklim BMKG di Jakarta. *Journal of Information Technology and Computer Science (INTECOMS)*, 7(6).
- Hardiman, J. P. W., Thio, D. C., Zakiyyah, A. Y., & Meiliana. (2024). AI-powered dialogues and quests generation in role-playing games using Google's *Gemini* and *Sentence-BERT* framework. *Procedia Computer Science*, 245(C), 1111–1119. <https://doi.org/10.1016/j.procs.2024.10.340>.
- Holis, R. M., Utomo, P. E. P., & Hutabarat, B. F. (2025). Semantic FAQ chatbot using *SBERT* (*Sentence-BERT*) and *cosine similarity* for academic services. *Brilliance: Research of Artificial Intelligence*, 5(2), 915–922. <https://doi.org/10.47709/brilliance.v5i2.7027>.
- Khadija, M. A., Harjito, B., Saberi, M., Paradhita, A. N., & Nurharjadmo, W. (2025). Generative Indonesian chatbot for university major selection using *transformers* embedding. *IAES International Journal of Artificial Intelligence*, 14(4), 3474–3482. <https://doi.org/10.11591/ijai.v14.i4.pp3474-3482>.
- Lee, J. S., & Hsiang, J. (2020). Patent classification by fine-tuning *BERT* language model. *World Patent Information*, 61. <https://doi.org/10.1016/j.wpi.2020.101965>.
- Owen, C., & Suhartono, D. (2025). Intent classification in artificial intelligence-based customer service chatbot for e-wallet service providers. *International Journal of Computing and Digital Systems*, 17(1). <https://doi.org/10.12785/ijcds/1571046212>.
- Peyton, K., & Unnikrishnan, S. (2023). A comparison of chatbot platforms with the state-of-the-art *Sentence-BERT* for answering online student FAQs. *Results in Engineering*, 17. <https://doi.org/10.1016/j.rineng.2022.100856>.
- Rashid, H. W., & Ahmed, S. H. (2025). Fine-tuning *SBERT* for semantic research title classification in trilingual university repository. *Kurdistan Journal of Applied Research*, 10(2), 119–135. <https://doi.org/10.24017/science.2025.2.9>.
- Rusydi, A. S., Win Afgani, Fatimah, Dian Septaria, Salsabila, & Gebriella Zahira. (2024). Metode penelitian kuantitatif pendekatan ilmiah untuk analisis data. *Jurnal Review Pendidikan dan Pengajaran*, 7(3), 11279–11289.
- Safitri, L., & Erfina, A. (2025). Analisis sentimen perbandingan *Lange Language Model* study kasus: *GPT*, *Gemini* dan *Llama* menggunakan *BERT* sentiment analysis. *Jurnalnya Orang Pintar Komputer*, 14(3). <https://doi.org/10.30591/smartcomp.v13i1.8373>.
- Setiawan, E. I., Purnomo, M. H., & Saputra, D. G. (2024). Pemrograman multiplatform dengan *Flutter* dilengkapi pengembangan AI chatbot dan retrieval augmented generation dengan *Gemini*. Deepublish.
- Siddik Hasibuan, M., Rifqi Al Fauzan, M., Komputer, I., dan Teknologi, S., & Islam Negeri Sumatera Utara, U. (2024). Kombinasi *TF-IDF* dan *Neural Network* untuk pelayanan informasi Al-Qur'an dalam bentuk chatbot. *Jurnal Fasilkom*, 14(2), 318–324.
- Sopyandi, D., Tria Sati, A., Prabowo, A., Maulana Putra Setya, D., Agung Nugroho, F., Studi Teknik Informatika Universitas Pamulang, P., Raya Puspitek, J., Pamulang, K., & Tangerang Selatan, K. (2024). Bangun aplikasi tracking cuaca (*Weather App*) menggunakan public API berbasis

website. *Jurnal Ilmiah Wahana Pendidikan*, 2024(20), 209–215.
<https://doi.org/10.5281/zenodo.14274718>.

Tat, O., & Aydoğan, İ. (2024). Discovering hidden patterns: Applying topic modeling in qualitative research. *Journal of Measurement and Evaluation in Education and Psychology*, 15(3), 247–259.
<https://doi.org/10.21031/epod.1539694>.