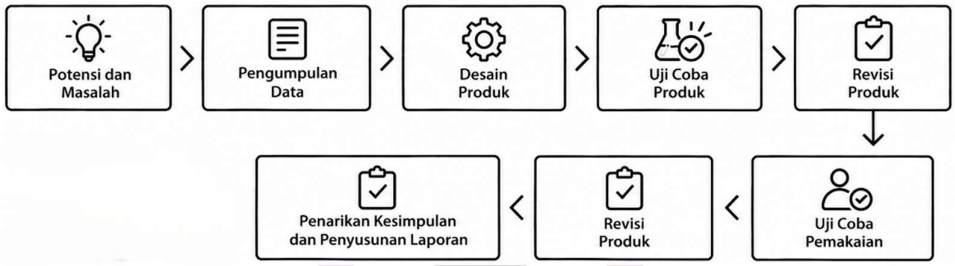
BAB III
METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini menggunakan metode *Research and Development* (R&D). Metode R&D dipilih karena penelitian ini bertujuan untuk menghasilkan dan menguji sebuah produk berupa sistem rekomendasi mobil Toyota berbasis kecerdasan buatan. Dalam penelitian R&D, proses penelitian tidak hanya berfokus pada kajian teori, tetapi juga mencakup proses pengembangan produk, pengujian, dan evaluasi untuk mengetahui kualitas produk yang dihasilkan. Secara umum, penelitian dan pengembangan dilakukan melalui tahapan yang mencakup identifikasi kebutuhan, perancangan, pengembangan, pengujian, dan evaluasi produk (Ade Rahayu, 2025).



Gambar 3. 1 Kerangka Alur Pelaksanaan Penelitian

Secara konseptual, alur penelitian ini mengacu pada sepuluh langkah model *Research and Development* (R&D) Borg & Gall yang diadaptasi oleh Sugiyono (2023) sebagaimana ditunjukkan pada Gambar 2.3, yaitu: (1) potensi dan masalah, (2) pengumpulan data, (3) desain produk, (4) validasi desain, (5) revisi desain, (6) uji coba produk, (7) revisi produk, (8) uji coba pemakaian, (9) revisi produk, dan (10) produksi masal.

Namun, untuk menyesuaikan kebutuhan dan batasan ruang lingkup penelitian, model tersebut dimodifikasi sebagaimana digambarkan pada Gambar 3.2. Tahap validasi desain (4) dan revisi desain (5) ditiadakan sehingga dari tahap desain produk langsung berlanjut ke uji coba produk, dilanjutkan dengan revisi produk, uji coba pemakaian, revisi produk akhir, serta ditutup dengan penarikan kesimpulan dan penyusunan laporan. Tahap produksi masal (10) tidak dilaksanakan karena penelitian ini berorientasi pada pembuktian

UIN IMAM BONJOL PADANG

konsep (*proof of concept*) dalam skala terbatas. Luaran yang dihasilkan berupa prototipe fungsional sistem rekomendasi mobil berbasis *Hybrid Retrieval-Augmented Generation* (Hybrid RAG) yang dikembangkan dan diuji dalam skenario *single-user*, sehingga penelitian ini tidak ditujukan hingga tahap komersialisasi masal ataupun kesiapan sistem di lingkungan *production-ready*.

Pemetaan tahapan modifikasi alur R&D pada penelitian ini dijabarkan sebagai berikut:

1. Potensi dan Masalah

Tahap inisiasi penelitian diawali dengan melakukan identifikasi masalah dan observasi terhadap tantangan operasional yang dihadapi oleh calon konsumen dalam memilih varian kendaraan Toyota pada *dealer* Auto2000 Rantauprapat. Observasi ini berfokus pada keterbatasan sistem pencarian konvensional yang kaku serta kelemahan model *Large Language Model* (LLM) standar yang rentan mengalami halusinasi data harga dan spesifikasi teknis. Guna mengatasi permasalahan tersebut, dilakukan studi literatur secara komprehensif terhadap literatur ilmiah terkini mengenai arsitektur *Hybrid Retrieval-Augmented Generation* (Hybrid RAG), representasi vektor semantik menggunakan model *embedding* BGE-M3, basis data terdistribusi TiDB Serverless, algoritma pemeringkatan *Reciprocal Rank Fusion* (RRF) dan BM25, serta kerangka kerja evaluasi kuantitatif DeepEval dan *System Usability Scale* (SUS) sebagai landasan teori dalam merumuskan celah penelitian (*research gap*).

2. Pengumpulan Data

Tahap pengumpulan data berfokus pada penghimpunan informasi resmi seluruh lini kendaraan Toyota yang relevan dengan ketersediaan unit pada *dealer* Auto2000 Rantauprapat. Data primer yang dikumpulkan mencakup dokumen brosur digital (*e-brochure*) spesifikasi teknis resmi berformat PDF serta gambar pendukung yang diunduh langsung dari laman resmi Toyota Astra Motor (*www.toyota.astra.co.id*). Dokumen tersebut mencakup berbagai segmen kendaraan, mulai dari kendaraan penumpang perkotaan (*city car* dan *passenger car*), kendaraan komersial, hingga varian

UIN IMAM BONJOL PADANG

elektrifikasi ramah lingkungan berbasis *Hybrid Electric Vehicle* (HEV) dan *Battery Electric Vehicle* (BEV). Selain data spesifikasi teknis, dihimpun pula data harga resmi *On-The-Road* (OTR) bulan Mei yang berlaku untuk wilayah Kabupaten Labuhanbatu. Data daftar harga tersebut diperoleh secara langsung dari *dealer* Auto2000 Rantauprapat melalui tenaga penjual (*sales executive*) terkait. Mengingat data spesifikasi dan harga diperoleh langsung dari sumber resmi yang valid, seluruh kumpulan informasi tersebut siap digunakan sebagai data rujukan pada tahap *data ingestion*.

3. Desain Produk

Tahap desain produk terdiri atas dua bagian utama, yaitu *data ingestion* dan perancangan arsitektur Hybrid RAG. Proses *data ingestion* merupakan pengolahan dokumen brosur mentah menjadi struktur data terorganisasi yang siap diakses oleh mesin pencari. Proses ini memanfaatkan kapabilitas multimodal dari model Google Gemini 3.1 Flash Lite yang dikombinasikan dengan pustaka PyMuPDF untuk memindai halaman dokumen PDF dan mengekstrak tabel spesifikasi teknis kendaraan secara terstruktur ke dalam format JSON. Selanjutnya, teks identitas varian dan rincian spesifikasi kendaraan dikonversi menjadi representasi vektor numerik berdimensi 1.024 menggunakan model *embedding* BGE-M3. Data atribut relasional (seperti nama model, varian, harga OTR Labuhanbatu, dan konsumsi bahan bakar) beserta kolom vektor *embedding* tersebut kemudian disimpan secara terintegrasi ke dalam tabel data_mobil_hybrid pada basis data TiDB Serverless.

UIN IMAM BONJOL PADANG

Setelah data siap diakses, dilakukan perancangan dan implementasi mesin pencari terpadu (*retrieval engine*) berbasis arsitektur Hybrid RAG yang mengombinasikan pencarian semantik dan penyaringan relasional. Alur pemrosesan diawali dengan modul *Self-Querying Retriever* yang memanfaatkan model Gemini di gerbang awal untuk membedah kueri bahasa alami pengguna menjadi parameter JSON terstruktur (batas anggaran, jenis mesin, kapasitas kursi, dan kata kunci fitur). Parameter tersebut dieksekusi melalui dua jalur paralel pada TiDB Serverless, yaitu

Dense Vector Search (menggunakan fungsi vec_cosine_distance) dan *Sparse SQL Search*. Seluruh kandidat hasil pencarian dari kedua jalur digabungkan ke dalam *candidate pool*, dideduplikasi, dinilai berdasarkan kemiripan leksikal dengan algoritma BM25, dan diperingkat ulang menggunakan *Reciprocal Rank Fusion* (RRF) untuk menghasilkan rekomendasi unit yang paling relevan.

4. Uji Coba Produk

Pada tahap uji coba produk, logika inferensi dan layanan *backend* dibangun menggunakan *Route API Handler* berbasis Next.js yang dieksekusi di sisi server (*server-side*) untuk menjamin keamanan kunci API. Sistem dilengkapi manajemen konkurensi dua lapis menggunakan Upstash Redis dan *in-memory semaphore*. Pada tahap ini, disusun pula instruksi sistem (*system prompt*) berpersona *AI Sales Executive* Auto2000 Rantauprapat pada model Google Gemini 3.1 Flash Lite agar menghasilkan rekomendasi yang komunikatif, ramah, persuasif, serta murni bersumber dari data faktual hasil *retrieval*.

Selanjutnya, dilakukan pengujian kinerja kuantitatif arsitektur RAG secara objektif menggunakan kerangka kerja DeepEval dengan pendekatan *LLM-as-a-Judge* terhadap 25 skenario kueri uji. Evaluasi ini mengukur metrik *Retrieval* (*Contextual Precision*, *Contextual Recall*, *Contextual Relevancy*) dan metrik *Generation* (*Faithfulness*, *Answer Relevancy*) dengan ambang batas kelulusan skor ≥ 0,60.

Setelah kinerja mesin *retrieval* teruji dan memenuhi kriteria tersebut, tahapan dilanjutkan dengan pembangunan antarmuka sistem. Antarmuka *frontend* web dibangun menggunakan Next.js, TypeScript, dan Tailwind CSS untuk menyediakan antarmuka *chatbot* interaktif yang responsif bagi pengguna akhir, serta panel admin berbasis Flask (admin_server.py) yang dilengkapi sistem autentikasi aman, operasi CRUD data kendaraan, dan fitur otomasi ekstraksi brosur langsung ke TiDB Serverless. Terakhir, pengujian fungsionalitas perangkat lunak dilakukan

UIN IMAM BONJOL PADANG

menggunakan metode *Black Box Testing* pada modul pengguna (*chatbot*) dan modul admin dengan kriteria kelulusan 100% *test cases Pass*

5. Revisi Produk (Tahap I)

Berdasarkan hasil evaluasi DeepEval Framework dan pengujian Black Box pada tahap uji coba produk, dilakukan perbaikan teknis terhadap sistem. Perbaikan mencakup penyempurnaan rekayasa *prompt*, kalibrasi bobot parameter pencarian hibrida (BM25 dan RRF), serta penanganan kendala fungsionalitas maupun antarmuka yang ditemukan saat pengujian *black box*.

6. Uji Coba Pemakaian

Tahap uji coba pemakaian dilakukan untuk mengukur tingkat penerimaan, kepuasan, dan kemudahan penggunaan antarmuka sistem oleh pengguna nyata (calon konsumen Auto2000). Pengukuran dilakukan menggunakan kuesioner terstandar *System Usability Scale* (SUS) yang memuat 10 butir pernyataan dengan skala Likert 1–5 terhadap 15 responden. Kriteria keberhasilan sistem ditetapkan berdasarkan skor rata-rata SUS ≥ 68, yang menunjukkan bahwa tingkat kebergunaan sistem masuk ke dalam kategori *Acceptable* (Grade C / *Good*).

7. Revisi Produk (Tahap Akhir)

Merupakan tahap penyempurnaan akhir terhadap antarmuka dan pengalaman pengguna (*user experience*) berdasarkan masukan, catatan, dan temuan empiris selama proses uji coba pemakaian dan evaluasi SUS dengan 15 responden nyata.

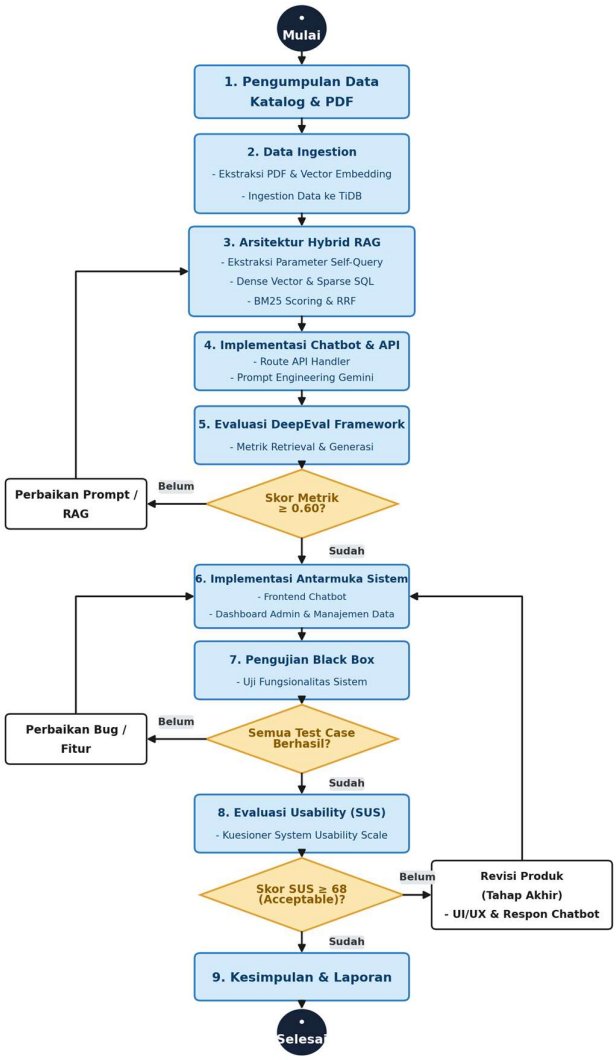
8. Penarikan Kesimpulan dan Penyusunan Laporan

Tahap akhir dari alur penelitian ini merupakan pengganti dari tahap produksi masal (yang ditiadakan karena fokus penelitian adalah *proof of concept* dalam skala terbatas). Pada tahap ini, dilakukan perangkuman dan integrasi terhadap seluruh data hasil capaian kinerja *Hybrid*-RAG via DeepEval, keandalan fungsional via Black Box, dan kegunaan sistem via SUS untuk ditarik menjadi kesimpulan ilmiah yang menjawab rumusan masalah penelitian. Selanjutnya, dirumuskan saran-saran teknis bagi

UIN IMAM BONJOL PADANG

pengembangan sistem di masa mendatang, serta disusun dokumen laporan tugas akhir/skripsi secara utuh dan sistematis sesuai kaidah akademik yang berlaku.

Untuk memberikan gambaran yang lebih rinci mengenai penerapan teknis dari tahapan Desain Produk dan Uji Coba Produk yang telah dijabarkan di atas, alur implementasi teknis sistem divisualisasikan pada Gambar 3.2 berikut, mencakup jalur iterasi perbaikan berdasarkan hasil evaluasi DeepEval, pengujian *Black Box*, dan evaluasi *System Usability Scale* (SUS).



Gambar 3. 2 Alur Teknis Implementasi Sistem

C. Setting Penelitian

1. Lokasi Penelitian

   Penelitian ini menjadikan jaringan *dealer* resmi otomotif Auto2000 Rantauprapat sebagai objek dan lokasi penelitian utama. Keterlibatan penulis secara langsung melalui program praktik kerja lapangan (magang) di lokasi tersebut memberikan akses untuk melakukan observasi. Observasi dilakukan untuk mendapatkan pemahaman empiris mengenai dinamika permasalahan riil yang dihadapi tenaga pemasar maupun konsumen dalam menavigasi kompleksitas variasi spesifikasi mobil. Pemilihan lokasi ini sangat krusial agar eksperimen sistem *Retrieval-Augmented Generation* (RAG) yang dikembangkan menggunakan *dataset* harga yang valid, sehingga sistem mampu merespons dan merepresentasikan skenario permasalahan di dunia nyata secara faktual.

2. Jadwal Penelitian

   Rangkaian penelitian ini dilaksanakan mulai bulan Februari 2026 dan diperkirakan akan diselesaikan secara komprehensif dalam kurun waktu 7 (tujuh) bulan, hingga Agustus 2026. Jadwal kegiatan disusun agar selaras dengan tahapan metode *Research and Development* yang telah dijabarkan sebelumnya. Rincian alur waktu pelaksanaan kegiatan penelitian terstruktur pada Tabel 3.1 berikut:

UIN IMAM BONJOL PADANG

Tabel 3. 1 Jadwal Penelitian

No	Tahapan Kegiatan / Waktu Pelaksanaan	Feb	Mar	Apr	Mei	Jun	Jul	Agu
1	Identifikasi Potensi & Masalah							
2	Pengumpulan Data							
3	Desain Produk							
4	Uji Coba Produk							
5	Revisi Produk (Tahap I)							
6	Uji Coba Pemakaian							
7	Revisi Produk (Tahap Akhir)							

No	Tahapan Kegiatan / Waktu Pelaksanaan	Feb	Mar	Apr	Mei	Jun	Jul	Agu
8	Penarikan Kesimpulan dan Penyusunan Laporan							

D. Bahan/Materi Penelitian

Bahan atau materi utama yang digunakan dalam penelitian ini adalah dataset spesifikasi kendaraan Toyota yang bersumber dari brosur dan daftar harga (*pricelist*) resmi OTR wilayah Labuhanbatu yang bersumber dari brosur penjualan Auto2000 Rantauprapat edisi Mei 2026. Data ini berperan sebagai basis pengetahuan (*knowledge base*) yang menjadi fondasi seluruh sistem rekomendasi.

1. Populasi dan Sampel

   Populasi dalam penelitian ini adalah keseluruhan varian kendaraan Toyota yang dipasarkan secara resmi di Indonesia, mencakup berbagai segmen seperti kendaraan penumpang, kendaraan niaga ringan, hingga kendaraan listrik. Penelitian ini menggunakan pendekatan census (*complete enumeration*) atau sampling jenuh, yaitu seluruh anggota populasi dijadikan sebagai unit analisis tanpa dilakukan proses pengambilan sampel (Sugiyono, 2023). Data yang digunakan mencakup seluruh model kendaraan Toyota yang terdaftar secara aktif dalam pricelist resmi Toyota Astra Motor berdasarkan data yang diakses pada Mei 2026. Secara keseluruhan, terdapat 32 (tiga puluh dua) tipe mobil dan 107 varian aktif yang diidentifikasi dan dikelompokkan ke dalam beberapa kategori segmen kendaraan.

2. Spesifikasi Bahan Kerja

   Karena bahan penelitian berupa dokumen digital (brosur PDF dan data harga), maka bahan kerja tersebut memiliki sifat dan spesifikasi sebagai berikut. Format dokumen sumber adalah PDF resmi yang diterbitkan oleh Toyota Astra Motor (*www.toyota.astra.co.id*). Setiap varian kendaraan yang diperoleh dari hasil ekstraksi brosur disusun dalam format JSON dengan 15 atribut data terstruktur. Data tersebut mencakup identitas kendaraan seperti

UIN IMAM BONJOL PADANG

nama model, tipe varian, segmentasi, dan pilihan warna, harga OTR wilayah Labuhanbatu, spesifikasi teknis seperti performa mesin, dimensi, sasis dan drivetrain, kapasitas tangki atau baterai, serta estimasi konsumsi BBM untuk penggunaan di dalam kota dan jalan tol. Selain itu, data juga mencakup deskripsi fitur kendaraan yang terdiri dari fitur keselamatan, kenyamanan interior, eksterior, dan teknologi *infotainment*.

Seluruh data tersebut distrukturkan ke dalam format JSON (*JavaScript Object Notation*) melalui ekstraksi otomatis berbasis AI agar dapat diproses oleh basis data TiDB Serverless. Apabila terdapat atribut yang bernilai kosong (*null*), data tersebut akan dilengkapi dan diisi secara manual berdasarkan data resmi dealer. Standar kualitas data yang disyaratkan adalah tidak terdapat nilai kosong (*null*) pada atribut harga OTR maupun spesifikasi teknis inti pada setiap baris data yang dimasukkan ke dalam sistem.

E. Alat/Instrumen Penelitian

Penelitian ini menggunakan sejumlah instrumen komputasi yang mendukung secara penuh dalam tahapan pra-pemrosesan data korpus, implementasi model kecerdasan buatan, hingga komputasi evaluasi metrik pada sistem berbasis *Retrieval-Augmented Generation* (RAG). Instrumen tersebut diklasifikasikan menjadi dua kategori utama, yaitu:

1. Perangkat Keras (*Hardware*)

   Perangkat keras adalah komponen fisik dari sistem komputer yang memiliki spesifikasi tertentu untuk memengaruhi kelancaran komputasi vektor, pengeksekusian skrip algoritma, dan interaksi jaringan API model kecerdasan buatan. Rincian perangkat keras yang digunakan dalam penelitian eksperimental ini diuraikan pada Tabel 3.3 berikut:

UIN IMAM BONJOL PADANG

Tabel 3. 2 Perangkat Keras

NO	Keterangan	Perangkat Keras (Hardware)	Deskripsi
1.	**Laptop**	ASUS TUF A15 (*FA507RE*)	Perangkat komputasi utama yang memberikan daya pemrosesan stabil untuk menjalankan peladen lokal (*localhost*) dan mengeksekusi skrip *machine learning*.
2.	**Processor**	AMD Ryzen 7 6800H	Spesifikasi prosesor yang mumpuni untuk menangani komputasi asinkron dari *framework* peladen dan kompilasi pemecahan teks (*chunking*).
3.	**GPU (Dual GPU)**	AMD Radeon 680M & RTX 3050Ti 4GB	Konfigurasi *Dual Graphic Processing Unit* memastikan komputasi grafis dan *rendering* antarmuka berjalan optimal tanpa membebani memori sistem utama.
4.	**RAM**	16GB DDR5 (4800MHz)	Kapasitas memori utama yang memadai untuk melakukan *multitasking* saat menguji API, membuka basis data TiDB, dan menjalankan kerangka pengujian secara bersamaan.
5.	**Storage**	512GB M.2 NVMe PCIe Gen4 SSD	Penyimpanan data (*Solid State Drive*) yang memberikan kecepatan *Input/Output* (I/O)

UIN IMAM BONJOL PADANG

NO	Keterangan	Perangkat Keras (Hardware)	Deskripsi
			tinggi untuk memproses pembacaan *dataset* korpus berformat JSON dengan latensi rendah.

2. Perangkat Lunak (*Software*)

Perangkat lunak merupakan instrumen program komputer yang dirancang untuk melakukan tugas komputasi spesifik, memfasilitasi penulisan kode analitik, pengelolaan basis data vektor, hingga pengorkestrasian dan pengujian model bahasa AI. Rincian perangkat lunak yang digunakan diuraikan pada Tabel 3.4 berikut:

Tabel 3. 3 Perangkat Lunak

NO	Perangkat Lunak (Software)	Fungsi	Deskripsi
1.	Windows 11 (Ver.25H2)	Sistem Operasi	Platform utama untuk menjalankan1. lingkungan pengembangan dan eksekusi kode.
2.	Google Antigravity	*Code Editor* / IDE	*Integrated Development Environment* (IDE) mutakhir berbasis *agent-first* dari Google. Digunakan untuk menulis kode eksperimen Python dan memfasilitasi integrasi cerdas dalam rekayasa AI.
3.	Google Chrome	Peramban (*Browser*)	Peramban *web* yang digunakan untuk mengakses antarmuka evaluasi, sistem *chat*, serta

UIN IMAM BONJOL PADANG

NO	Perangkat Lunak (Software)	Fungsi	Deskripsi
			menelusuri dokumentasi teknis pustaka secara *real-time*.
4.	Node.js	Runtime Environment	Lingkungan eksekusi JavaScript/TypeScript server-side untuk menjalankan aplikasi Next.js.
5.	Next.js & React	Framework Web Utama	Framework full-stack untuk antarmuka chat kustomer dan Route API Handler utama (/api/ai-chat).
6.	Flask Framework	Web Server Backend Admin	Framework web Python untuk menjalankan portal back-office manajemen pangkalan data.
7.	TiDB Cloud (*Serverless*)	Basis Data (*Database*)	Pangkalan data *hybrid* untuk penyimpanan data relasional dan *vector embedding* spesifikasi mobil.
8.	TypeScript & Python	Bahasa Pemrograman	Bahasa utama untuk pengembangan antarmuka/API route (TypeScript) serta pemrosesan data/evaluasi (Python).
9.	Upstash Redis	In-Memory Database & Semaphore	Digunakan untuk manajemen concurrency lock guna mencegah rate-limit API Gemini pada trafik tinggi.
10.	Google Gemini API	*Model AI (LLM)*	Digunakan sebagai mesin utama untuk generasi jawaban

UIN IMAM BONJOL PADANG

NO	Perangkat Lunak (Software)	Fungsi	Deskripsi
			(*inference*) menggunakan model Gemini 3.1 Flash Lite.
11.	Google GenAI SDK	Official SDK Client	Pustaka resmi Google (@google/generative-ai & google-genai) untuk interaksi native dengan Gemini API.
12.	BGE-M3	*Embedding Model*	Digunakan untuk proses vektorisasi data spesifikasi kendaraan menjadi representasi numerik 1024-dimensi guna mendukung pencarian semantik.
13.	HuggingFace Space	*Cloud Platform / MLOps*	Platform infrastruktur *cloud* yang digunakan untuk men-*deploy* dan menjalankan model BGE-M3 agar dapat diakses oleh sistem RAG melalui *Gradio Client*.
14.	DeepEval Framework	*Evaluation Tool*	Framework untuk melakukan pengujian otomatis terhadap metrik *faithfulness*, *relevancy*, *precision*, dan *recall*.
15.	Gradio Client	*Library / API Client*	Pustaka Python yang berfungsi sebagai jembatan komunikasi untuk mengirim data teks ke HuggingFace Space dan menerima hasil vektor dari model BGE-M3 secara *real-time*.

UIN IMAM BONJOL PADANG

NO	Perangkat Lunak (Software)	Fungsi	Deskripsi
16.	PyMuPDF	PDF Processing Library	Pustaka Python untuk memindai dan merender halaman spesifikasi PDF brosur ke gambar PNG 150 DPI.
17.	mysql2 & PyMySQL	*Library* Konektivitas	Menghubungkan aplikasi Python dengan TiDB melalui koneksi SSL yang aman.
18.	PlantUML & Eraser	Alat Visualisasi	Digunakan untuk merancang diagram alur penelitian dan arsitektur sistem secara profesional.
19.	Tailwind CSS	UI Styling Framework	Framework CSS berbasis utilitas untuk perancangan antarmuka obrolan pengguna yang responsif.
20.	Git & GitHub	Version Control System	Sistem pengontrol versi untuk mengelola repositori kode proyek secara terstruktur.
21.	Microsoft Word	Dokumen Teknis	Media utama untuk penyusunan laporan akhir dan naskah skripsi.

UIN IMAM BONJOL PADANG

F. Teknik Pengumpulan Data

Teknik pengumpulan data digunakan untuk memperoleh informasi yang diperlukan dalam pengembangan dan evaluasi sistem rekomendasi mobil Toyota berbasis Hybrid RAG. Data yang dikumpulkan mencakup data kendaraan yang digunakan sebagai sumber informasi sistem, referensi yang mendukung pengembangan metode, serta penilaian pakar yang digunakan sebagai acuan dalam proses evaluasi. Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi:

1. Studi Dokumen (Dokumentasi)

   Studi dokumentasi digunakan untuk memperoleh data kendaraan yang menjadi sumber informasi dalam sistem. Data spesifikasi dan fitur kendaraan diperoleh dari dokumen resmi Toyota Indonesia, khususnya brosur kendaraan yang telah dipublikasikan. Sementara itu, data harga OTR wilayah Labuhanbatu diperoleh secara langsung dari *Sales* resmi Auto2000 Rantauprapat dalam bentuk tabel daftar harga OTR kendaraan Toyota pada bulan Mei 2026. Data yang diperoleh meliputi identitas kendaraan, harga OTR, spesifikasi teknis, serta fitur kendaraan. Seluruh data kemudian diperiksa dan disusun secara terstruktur dalam format JSON untuk digunakan sebagai sumber informasi pada sistem *Hybrid* RAG dan disimpan dalam TiDB *Serverless*.

G. Teknik Analisis Data

Teknik analisis data digunakan untuk mengolah dan menilai hasil pengujian sistem rekomendasi mobil Toyota berbasis *Hybrid* RAG. Analisis dilakukan secara kuantitatif untuk mengetahui kemampuan sistem dalam menjalankan fungsi yang dirancang, mengambil informasi yang relevan, menghasilkan jawaban yang sesuai dengan sumber data, serta memberikan kemudahan penggunaan bagi pengguna. Analisis data dilakukan melalui tiga metode, yaitu evaluasi DeepEval, Black Box Testing, dan *System Usability Scale* (SUS).

UIN IMAM BONJOL PADANG

1. Evaluasi DeepEval

   Evaluasi DeepEval digunakan untuk mengukur kualitas sistem Hybrid RAG, terkhusus pada proses pencarian informasi (*retrieval*) dan jawaban yang dihasilkan oleh Google Gemini. Evaluasi dilakukan menggunakan pendekatan *LLM-as-a-Judge*, yaitu menggunakan model bahasa sebagai penilai untuk memberikan skor terhadap hasil sistem berdasarkan kriteria tertentu. Penilaian dilakukan menggunakan lima metrik, yaitu:

a. *Contextual Precision*, digunakan untuk menilai ketepatan informasi yang ditemukan dan urutan relevansinya terhadap pertanyaan pengguna.
b. *Contextual Recall*, digunakan untuk menilai kemampuan sistem dalam menemukan informasi yang dibutuhkan untuk menjawab pertanyaan pengguna.
c. *Contextual Relevancy*, digunakan untuk menilai kesesuaian informasi yang diambil dengan pertanyaan pengguna.
d. *Faithfulness*, digunakan untuk menilai apakah jawaban yang dihasilkan sesuai dengan informasi yang digunakan sebagai konteks.
e. *Answer Relevancy*, digunakan untuk menilai kesesuaian jawaban yang dihasilkan dengan pertanyaan pengguna.

Setiap metrik menghasilkan nilai yang digunakan untuk menggambarkan kualitas sistem. Nilai tersebut kemudian dianalisis untuk mengetahui kemampuan *Hybrid* RAG dalam mengambil informasi yang relevan dan menghasilkan jawaban yang sesuai dengan sumber data kendaraan.

2. Black Box Testing

Black Box Testing digunakan untuk mengetahui apakah fungsi-fungsi pada sistem telah berjalan sesuai dengan kebutuhan yang telah ditentukan. Pengujian dilakukan dengan memberikan berbagai masukan kepada sistem dan mengamati keluaran yang dihasilkan tanpa melihat proses internal program.

Pada penelitian ini, pengujian mencakup fungsi utama chatbot, seperti menerima pertanyaan pengguna, memahami parameter pencarian, melakukan pencarian data kendaraan, memberikan rekomendasi, serta menghasilkan jawaban kepada pengguna. Hasil pengujian dicatat berdasarkan kesesuaian antara keluaran sistem dan keluaran yang diharapkan. Sistem dinyatakan berhasil pada suatu skenario apabila keluaran yang dihasilkan sesuai dengan hasil yang telah ditentukan.

UIN IMAM BONJOL PADANG

3. *System Usability Scale* (SUS)

System Usability Scale (SUS) digunakan untuk mengukur tingkat kemudahan penggunaan sistem dari sudut pandang pengguna. Pengujian dilakukan setelah pengguna mencoba menggunakan chatbot rekomendasi mobil Toyota.

Instrumen SUS terdiri atas 10 pernyataan dengan lima pilihan tingkat persetujuan. Pernyataan bernomor ganjil merupakan pernyataan positif, sedangkan pernyataan bernomor genap merupakan pernyataan negatif. Perhitungan skor dilakukan dengan mengubah nilai jawaban sesuai aturan SUS, kemudian menjumlahkan seluruh nilai dan mengalikannya dengan 2,5 sehingga diperoleh skor akhir dalam rentang 0–100.

Skor SUS selanjutnya digunakan untuk menentukan tingkat *usability* sistem. Semakin tinggi skor yang diperoleh, semakin baik tingkat penerimaan dan kemudahan penggunaan sistem berdasarkan penilaian pengguna.

4. Analisis Hasil

Hasil dari ketiga metode pengujian tersebut kemudian dianalisis untuk memperoleh gambaran menyeluruh mengenai kualitas sistem yang dikembangkan. DeepEval digunakan untuk mengukur kualitas proses *retrieval* dan jawaban yang dihasilkan, Black Box Testing digunakan untuk mengetahui keberhasilan fungsi sistem, sedangkan SUS digunakan untuk mengetahui tingkat kemudahan penggunaan sistem.

Hasil analisis tersebut menjadi dasar untuk menilai apakah sistem rekomendasi mobil Toyota berbasis *Hybrid* RAG telah memenuhi tujuan pengembangan yang ditetapkan dalam penelitian. Dengan demikian, keberhasilan sistem tidak hanya dilihat dari apakah sistem dapat berjalan, tetapi juga dari ketepatan fungsi, kualitas informasi dan jawaban, serta kemudahan penggunaannya.

.

UIN IMAM BONJOL PADANG