Copyright © The Author(s) 2026
Jurnal Artificial, Vol 4 No 2, Oktober 2026

Pengembangan Chatbot RAG dengan Hybrid Retrieval untuk Informasi Penyakit Menular di Indonesia

Dian Assarini [1*], Cindi Wulandari [2], Rusdiyanto [3], Tri Hasannah Bimastari Aviani

[1, 2, 3, 4] Universitas Bina Insan, Indonesia

* dianassarini117@gmail.com

Abstrak

Ketersediaan informasi mengenai penyakit menular yang mudah diakses dan relevan menjadi salah satu kebutuhan dalam pemanfaatan teknologi digital di bidang kesehatan. Namun, chatbot berbasis *Large Language Model* (LLM) dapat menghasilkan respons yang kurang sesuai dengan sumber pengetahuan apabila tidak didukung mekanisme pengambilan informasi yang relevan. Penelitian ini bertujuan mengembangkan chatbot berbasis *Retrieval-Augmented Generation* (RAG) dengan pendekatan *hybrid retrieval* untuk menyediakan informasi penyakit menular di Indonesia secara kontekstual. Penelitian menggunakan pendekatan pengembangan sistem dengan basis pengetahuan yang bersumber dari Halodoc dan Alodokter, mencakup delapan penyakit menular, yaitu TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Proses retrieval menggabungkan pencarian semantik menggunakan *Sentence Transformer* dan pencarian berbasis kata kunci menggunakan TF-IDF, kemudian konteks yang diperoleh digunakan sebagai masukan dalam proses pembentukan respons oleh LLM. Hasil implementasi menunjukkan bahwa chatbot berhasil dikembangkan dalam bentuk aplikasi berbasis web dan mampu menerima pertanyaan serta menghasilkan respons sesuai dengan konteks informasi yang tersedia. Pengujian terhadap pertanyaan mengenai delapan penyakit menunjukkan bahwa sistem dapat memproses pertanyaan dan menampilkan respons, sedangkan pengujian *backend* menunjukkan proses sistem berjalan tanpa kesalahan. Dengan demikian, integrasi RAG dan *hybrid retrieval* dapat digunakan sebagai pendekatan untuk membangun chatbot informasi penyakit menular yang lebih terarah berdasarkan basis pengetahuan tertentu. Pengembangan selanjutnya diperlukan dengan memperluas cakupan sumber dan penyakit serta memperkuat evaluasi kualitas retrieval dan respons secara kuantitatif.

Kata Kunci: *Chatbot Kesehatan, Hybrid Retrieval, Large Language Model, Penyakit Menular, Retrieval-Augmented Generation*

Pendahuluan

Perkembangan kecerdasan buatan (*Artificial Intelligence*/AI) telah mendorong perubahan dalam cara masyarakat memperoleh dan berinteraksi dengan informasi, termasuk dalam bidang kesehatan. Salah satu penerapan AI yang semakin berkembang adalah *chatbot*, yaitu sistem yang memungkinkan pengguna berinteraksi dengan komputer melalui bahasa alami untuk memperoleh informasi atau bantuan tertentu. Dalam konteks layanan kesehatan, chatbot memiliki potensi untuk memperluas akses informasi karena dapat memberikan respons secara cepat, interaktif, dan tersedia tanpa terbatas pada waktu pelayanan konvensional (Xu et al., 2021). Pemanfaatan chatbot juga dapat membantu menyederhanakan informasi kesehatan yang kompleks menjadi bentuk

https://doi.org/10.54065/artificial.4.2.2026.1992

yang lebih mudah dipahami oleh masyarakat (Laymouna et al., 2024). Namun, penerapan chatbot pada domain kesehatan memiliki tuntutan yang lebih tinggi dibandingkan penggunaan chatbot pada domain umum karena informasi yang disampaikan berkaitan dengan kondisi kesehatan dan berpotensi memengaruhi pemahaman maupun keputusan pengguna (Septian et al., 2026). Oleh karena itu, sistem tidak hanya dituntut mampu menghasilkan respons yang komunikatif, tetapi juga harus mampu memberikan informasi yang relevan, konsisten, dan didukung oleh sumber pengetahuan yang dapat dipertanggungjawabkan. Penelitian mengenai chatbot kesehatan juga menunjukkan bahwa kualitas informasi dan algoritma yang digunakan merupakan faktor penting dalam menentukan efektivitas sistem (Li et al., 2024; Hayat et al., 2024).

Permasalahan tersebut menjadi semakin penting dalam penyediaan informasi mengenai penyakit menular di Indonesia. Penyakit menular masih menjadi salah satu persoalan kesehatan masyarakat karena dapat menyebar melalui berbagai mekanisme dan berpotensi menimbulkan peningkatan kasus pada kelompok maupun wilayah tertentu. Tuberkulosis (TBC), demam berdarah dengue (DBD), HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19 merupakan beberapa penyakit yang membutuhkan pemahaman masyarakat mengenai gejala, penyebab, cara penularan, pencegahan, serta penanganan awal. TBC, misalnya, masih menjadi masalah kesehatan yang serius di Indonesia dengan beban kasus yang tinggi. Kondisi tersebut menunjukkan bahwa kebutuhan terhadap informasi kesehatan yang mudah diakses tidak hanya berkaitan dengan penyediaan informasi, tetapi juga dengan kemampuan sistem dalam menyajikan informasi yang sesuai dengan pertanyaan pengguna (Kemenkes, 2024).

Di sisi lain, perkembangan teknologi informasi telah menyebabkan masyarakat memperoleh informasi kesehatan dari berbagai saluran digital dengan tingkat kualitas yang tidak selalu sama. Informasi mengenai penyakit dapat ditemukan melalui mesin pencari, media sosial, forum daring, maupun berbagai situs kesehatan. Banyaknya sumber tersebut memang meningkatkan ketersediaan informasi, tetapi sekaligus dapat menimbulkan kesulitan dalam menentukan informasi yang relevan dan dapat dipercaya (Aggarwal et al., 2023; Kurniawan et al., 2024). Dalam konteks penyakit menular, informasi yang tidak tepat dapat menyebabkan kesalahpahaman mengenai gejala, mekanisme penularan, pencegahan, maupun tindakan yang seharusnya dilakukan. Karena itu, diperlukan pendekatan sistem informasi yang mampu menggabungkan kemudahan interaksi dengan kemampuan memperoleh informasi dari sumber pengetahuan yang telah ditentukan. Salah satu pendekatan yang dapat digunakan adalah *Large Language Model* (LLM) yang diintegrasikan dengan *Retrieval-Augmented Generation* (RAG) (Cahyanti & Raya, 2025).

LLM memiliki kemampuan untuk memahami pertanyaan dalam bahasa alami dan menghasilkan respons yang menyerupai komunikasi manusia. Kemampuan tersebut membuat LLM sesuai untuk dikembangkan sebagai komponen pembangkit jawaban pada chatbot (Yang et al., 2023). Meskipun demikian, LLM memiliki keterbatasan karena pengetahuan yang digunakan dalam menghasilkan respons bergantung pada data dan proses pelatihannya. Model juga dapat menghasilkan informasi yang terdengar meyakinkan tetapi tidak sepenuhnya sesuai dengan fakta atau konteks yang dibutuhkan. Permasalahan tersebut dikenal sebagai *hallucination*, sehingga penggunaan LLM secara mandiri menjadi kurang ideal untuk sistem yang membutuhkan informasi berbasis pengetahuan tertentu (Ji et al., 2023). Dalam domain kesehatan, keterbatasan tersebut

menjadi perhatian karena jawaban yang tidak didukung konteks sumber dapat menurunkan keandalan sistem (Hayat et al., 2024; Devine, 2025).

RAG dikembangkan sebagai salah satu pendekatan untuk mengatasi keterbatasan tersebut dengan menggabungkan proses *retrieval* dan *generation*. Pada tahap *retrieval*, sistem mencari informasi yang relevan dari basis pengetahuan eksternal berdasarkan pertanyaan pengguna. Informasi yang ditemukan kemudian diberikan kepada LLM sebagai konteks dalam proses *generation*. Dengan mekanisme tersebut, model tidak hanya mengandalkan pengetahuan internalnya, tetapi memperoleh informasi tambahan dari dokumen yang telah disiapkan. Pendekatan ini memungkinkan sistem menghasilkan respons yang lebih sesuai dengan domain pengetahuan yang digunakan dan dapat membantu mengurangi risiko munculnya informasi yang tidak didukung oleh sumber. Sejumlah penelitian menunjukkan bahwa RAG memiliki potensi dalam meningkatkan relevansi dan kualitas respons pada sistem tanya jawab berbasis LLM, termasuk pada domain kesehatan (Tohir et al., 2024; Murtiyoso et al., 2025).

Walaupun RAG memberikan mekanisme untuk menghubungkan LLM dengan sumber pengetahuan eksternal, kualitas sistem RAG sangat bergantung pada kemampuan komponen *retriever* dalam menemukan informasi yang tepat. Jika dokumen yang relevan tidak berhasil ditemukan, LLM akan menerima konteks yang kurang sesuai sehingga kualitas jawaban juga dapat menurun. Permasalahan ini menjadi penting pada informasi penyakit menular karena pertanyaan pengguna dapat menggunakan variasi istilah, sinonim, istilah awam, maupun istilah medis. Pencarian berdasarkan kemiripan semantik dapat membantu sistem memahami makna pertanyaan meskipun kata yang digunakan berbeda dengan dokumen. Sebaliknya, pencarian berbasis kata kunci dapat memberikan keuntungan ketika pertanyaan mengandung istilah penyakit, nama gejala, atau terminologi tertentu yang perlu dicocokkan secara langsung. Oleh sebab itu, penggunaan satu metode retrieval saja berpotensi belum optimal untuk menangani variasi pertanyaan pengguna (Li et al., 2023).

Pendekatan *hybrid retrieval* dapat digunakan untuk mengatasi keterbatasan tersebut dengan menggabungkan lebih dari satu mekanisme pencarian. Dalam penelitian ini, *hybrid retrieval* mengombinasikan *semantic search* berbasis representasi embedding dengan *keyword search* berbasis TF-IDF. *Semantic search* digunakan untuk menemukan dokumen berdasarkan kedekatan makna antara pertanyaan dan isi dokumen, sedangkan *keyword search* digunakan untuk mempertahankan kemampuan pencarian berdasarkan kecocokan istilah secara langsung. Hasil dari kedua mekanisme tersebut selanjutnya dikombinasikan untuk memperoleh konteks yang lebih relevan sebelum diberikan kepada LLM. Rancangan tersebut sesuai dengan karakteristik informasi kesehatan yang mengandung istilah teknis sekaligus dapat ditanyakan oleh pengguna menggunakan bahasa sehari-hari. Implementasi sistem pada penelitian ini menggunakan proses tersebut dalam rangka meningkatkan ketepatan pengambilan informasi sebelum tahap pembangkitan jawaban (Jayarana et al., 2025).

Beberapa penelitian terdahulu menunjukkan bahwa *Retrieval-Augmented Generation* (RAG) memiliki potensi dalam meningkatkan kualitas chatbot pada domain kesehatan. Pengembangan chatbot kesehatan untuk penyediaan informasi mengenai sistem pencernaan dengan memanfaatkan dokumen kesehatan sebagai sumber pengetahuan, pemrosesan teks, basis data vektor, serta mekanisme *retrieval* yang diintegrasikan dengan model bahasa menunjukkan bahwa pendekatan RAG mampu menghasilkan

respons yang lebih relevan dan kontekstual serta membantu mengurangi kesalahan informasi dibandingkan chatbot tanpa mekanisme *retrieval* (Samudra et al., 2025). Namun, penelitian tersebut masih berfokus pada informasi kesehatan mengenai sistem pencernaan sehingga belum secara khusus membahas penyediaan informasi penyakit menular di Indonesia dengan pendekatan *hybrid retrieval*. Penelitian lain mengembangkan chatbot berbahasa Indonesia untuk diagnosis penyakit menggunakan RAG dengan pendekatan *hybrid retrieval* yang mengombinasikan BM25 dan *embedding-based retrieval*, kemudian mengintegrasikannya dengan LLM GPT-OSS-20B serta artikel medis berbahasa Indonesia, dengan hasil berupa respons yang relevan dan kontekstual serta nilai *Top-30 Accuracy* sebesar 0,699 (Abdurrazzaq et al., 2025). Meskipun memiliki kesamaan dalam penerapan RAG dan *hybrid retrieval* pada domain kesehatan berbahasa Indonesia, penelitian tersebut berfokus pada diagnosis penyakit, sedangkan penelitian ini diarahkan pada penyediaan informasi edukatif mengenai penyakit menular di Indonesia dengan cakupan delapan jenis penyakit.

Penelitian mengenai penerapan *prompt engineering* pada LLM yang diintegrasikan dengan RAG untuk penyediaan informasi obat dan vitamin melalui perbandingan teknik *zero-shot prompting* dan *few-shot prompting* menunjukkan bahwa strategi *prompt engineering* dapat meningkatkan relevansi dan kualitas respons pada domain informasi kesehatan (Haromain et al., 2025). Temuan tersebut menunjukkan bahwa kualitas keluaran sistem berbasis RAG tidak hanya dipengaruhi oleh mekanisme *retrieval*, tetapi juga oleh cara konteks hasil *retrieval* diberikan kepada model bahasa. Selain itu, penerapan LLM berbasis RAG untuk *fact-checking* informasi COVID-19 menunjukkan bahwa pendekatan tersebut dapat mendukung penyediaan informasi kesehatan yang lebih andal, sekaligus menegaskan pentingnya kualitas sumber pengetahuan dan mekanisme pengolahan informasi dalam menghasilkan respons yang akurat (Li et al., 2025). Berdasarkan penelitian-penelitian tersebut, dapat diketahui bahwa efektivitas chatbot kesehatan dipengaruhi oleh keterpaduan antara sumber pengetahuan, mekanisme *retrieval*, strategi pemberian konteks, dan model generatif, sedangkan aspek yang belum banyak dibahas adalah penerapan *hybrid retrieval* untuk chatbot yang secara khusus menyediakan informasi edukatif mengenai berbagai penyakit menular di Indonesia. Oleh karena itu, penelitian ini berfokus pada pengembangan chatbot edukasi penyakit menular dengan menerapkan *hybrid retrieval* untuk meningkatkan relevansi dan ketepatan informasi yang diberikan kepada pengguna.

Berdasarkan penelitian-penelitian tersebut, dapat diketahui bahwa penggunaan RAG telah banyak menunjukkan manfaat dalam meningkatkan relevansi dan kualitas respons chatbot kesehatan. Penelitian terdahulu juga memperlihatkan bahwa penggunaan *hybrid retrieval* dapat menjadi strategi untuk menggabungkan kelebihan pencarian semantik dan pencarian berbasis kata kunci, khususnya pada informasi medis berbahasa Indonesia. Namun, masih terdapat ruang pengembangan pada penerapan pendekatan tersebut untuk sistem tanya jawab yang secara khusus menyediakan informasi penyakit menular di Indonesia. Selain itu, penelitian yang menggunakan RAG pada domain kesehatan tidak selalu menggunakan ruang lingkup penyakit yang terstruktur dan sumber pengetahuan yang dibatasi secara jelas. Padahal, pada sistem informasi kesehatan untuk masyarakat, batasan sumber pengetahuan penting untuk menjaga konsistensi informasi yang menjadi dasar pembentukan jawaban.

Kesenjangan penelitian tersebut penting karena informasi mengenai penyakit menular memiliki karakteristik yang beragam dan membutuhkan kemampuan retrieval yang dapat menangani istilah medis maupun pertanyaan dalam bahasa pengguna. Sistem yang hanya mengandalkan pencarian semantik dapat menghadapi kendala ketika istilah spesifik perlu dicocokkan secara langsung, sedangkan pencarian berbasis kata kunci dapat kehilangan hubungan makna ketika pengguna menggunakan istilah yang berbeda dari dokumen. Penggabungan kedua pendekatan diharapkan dapat memberikan konteks yang lebih sesuai bagi LLM sehingga jawaban yang dihasilkan lebih relevan terhadap pertanyaan. Dengan demikian, *hybrid retrieval* tidak hanya diposisikan sebagai fitur teknis, tetapi sebagai bagian penting dari strategi untuk meningkatkan kualitas penyediaan informasi kesehatan berbasis RAG.

Berdasarkan kesenjangan tersebut, penelitian ini mengembangkan chatbot RAG dengan *hybrid retrieval* untuk menyediakan informasi mengenai penyakit menular di Indonesia. Sistem dibangun menggunakan basis pengetahuan yang terdiri atas dokumen informasi untuk delapan penyakit menular, yaitu TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Data pengetahuan dikumpulkan dari sumber informasi kesehatan yang digunakan dalam penelitian, yaitu Halodoc dan Alodokter, kemudian melalui tahapan pengolahan dokumen, *chunking*, embedding, dan penyimpanan untuk mendukung proses retrieval. Sistem selanjutnya menggabungkan *semantic search* dan *keyword search* untuk memperoleh konteks yang relevan sebelum konteks tersebut diberikan kepada LLM untuk menghasilkan respons. Ruang lingkup delapan penyakit dan sumber pengetahuan tersebut merupakan batasan yang ditetapkan dalam penelitian sehingga sistem tidak diarahkan untuk menggantikan konsultasi medis, melainkan sebagai media penyediaan informasi kesehatan.

Kontribusi penelitian ini adalah menghasilkan rancangan dan implementasi chatbot berbasis RAG yang memanfaatkan *hybrid retrieval* untuk informasi penyakit menular dalam konteks bahasa Indonesia. Kontribusi berikutnya adalah menerapkan kombinasi *semantic search* dan *keyword search* untuk memperoleh konteks informasi yang lebih relevan sebelum proses generasi jawaban. Penelitian ini juga memberikan evaluasi terhadap kemampuan sistem dalam menghasilkan respons berdasarkan konteks pengetahuan yang tersedia. Dengan demikian, penelitian tidak hanya berfokus pada pembangunan antarmuka chatbot, tetapi juga pada integrasi antara basis pengetahuan, proses retrieval, dan LLM sebagai satu kesatuan sistem tanya jawab otomatis. Evaluasi sistem dalam penelitian dilakukan dengan memperhatikan aspek *faithfulness*, *answer relevance*, dan *context relevance*.

Metode

Desain Penelitian

Penelitian ini merupakan penelitian terapan (*applied research*) dengan pendekatan kuantitatif yang berfokus pada pengembangan chatbot berbasis *Large Language Model* (LLM) dan *Retrieval-Augmented Generation* (RAG) untuk menyediakan informasi mengenai penyakit menular di Indonesia. Penelitian terapan dipilih karena penelitian diarahkan untuk menghasilkan solusi terhadap kebutuhan nyata, yaitu penyediaan sistem informasi kesehatan yang dapat memberikan respons secara interaktif berdasarkan

sumber informasi yang telah ditentukan. Pendekatan kuantitatif digunakan dalam proses pengujian dan evaluasi sistem berdasarkan indikator yang telah ditetapkan.

Tahapan penelitian mengacu pada *AI Project Life Cycle* yang terdiri atas *problem scoping*, *data acquisition*, *data exploration*, *modelling*, dan *deployment*. Pada tahap *problem scoping*, permasalahan difokuskan pada kebutuhan terhadap sistem tanya jawab otomatis yang mampu menyediakan informasi penyakit menular secara relevan dan kontekstual. Tahap *data acquisition* dilakukan dengan mengumpulkan dokumen informasi penyakit menular dari sumber yang digunakan sebagai basis pengetahuan. Selanjutnya, *data exploration* dilakukan untuk memahami karakteristik, cakupan, dan relevansi informasi yang dikumpulkan. Tahap *modelling* meliputi pengolahan dokumen, pembentukan *embedding*, penerapan mekanisme *retrieval*, serta integrasi RAG dengan LLM. Tahap *deployment* merupakan tahap penerapan sistem chatbot berbasis web dan pengujian fungsional sistem. Secara keseluruhan, alur penelitian dimulai dari identifikasi permasalahan, pengumpulan dan eksplorasi data, persiapan dokumen, pembentukan basis pengetahuan, penerapan *hybrid retrieval*, integrasi konteks dengan LLM, implementasi chatbot, hingga evaluasi kualitas respons dan pengujian fungsional sistem.

Objek dan Sumber Data

Objek penelitian adalah chatbot berbasis LLM dan RAG yang digunakan untuk menyediakan informasi mengenai penyakit menular di Indonesia. Sistem dikembangkan dalam bentuk aplikasi berbasis web yang memungkinkan pengguna mengajukan pertanyaan melalui antarmuka percakapan dan memperoleh respons berdasarkan informasi yang tersedia pada basis pengetahuan.

Data penelitian berupa dokumen teks yang memuat informasi mengenai delapan jenis penyakit menular, yaitu Tuberkulosis (TBC), Demam Berdarah Dengue (DBD), HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Dokumen yang digunakan sebagai sumber informasi diperoleh dari Halodoc dan Alodokter. Pemilihan sumber dilakukan untuk menyediakan informasi kesehatan yang relevan dengan ruang lingkup penelitian dan selanjutnya digunakan sebagai basis pengetahuan sistem.

Pengumpulan data dilakukan melalui eksplorasi dokumen dan studi pustaka. Eksplorasi dokumen dilakukan untuk mengidentifikasi informasi yang tersedia dan menentukan kesesuaiannya dengan cakupan penyakit yang diteliti. Studi pustaka digunakan untuk mendukung pemahaman terhadap konsep dan teknologi yang diterapkan, khususnya LLM, RAG, *information retrieval*, *embedding*, dan pemrosesan bahasa alami. Dokumen yang telah dikumpulkan kemudian diseleksi berdasarkan relevansi terhadap ruang lingkup penelitian. Data yang sesuai selanjutnya dipersiapkan untuk digunakan dalam proses pembentukan basis pengetahuan dan *retrieval*.

Persiapan dan Pengolahan Data

Persiapan data dilakukan sebelum dokumen digunakan dalam proses *retrieval*. Tahap awal berupa pembersihan teks (*text cleaning*) untuk menghilangkan karakter atau bagian dokumen yang tidak diperlukan. Selanjutnya dilakukan normalisasi bahasa untuk menghasilkan teks yang lebih konsisten dan dapat diproses oleh sistem. Dokumen yang telah melalui proses pembersihan kemudian dibagi menjadi beberapa potongan teks (*chunking*). Proses *chunking* dilakukan agar informasi dalam dokumen dapat dikelola dalam unit yang lebih kecil dan memudahkan proses pencarian informasi yang relevan.

Setiap *chunk* selanjutnya diubah menjadi representasi vektor menggunakan teknik *embedding* dengan model *Sentence Transformer*. Representasi vektor tersebut digunakan untuk menggambarkan informasi berdasarkan karakteristik semantiknya sehingga sistem dapat membandingkan kemiripan antara pertanyaan pengguna dan potongan informasi dalam basis pengetahuan. Hasil pengolahan dokumen dan *embedding* selanjutnya digunakan sebagai basis pengetahuan untuk proses *retrieval*. Dengan tahapan tersebut, sistem tidak melakukan pencarian secara langsung terhadap dokumen mentah, tetapi menggunakan dokumen yang telah diproses dan diorganisasikan menjadi potongan informasi yang dapat dicari berdasarkan kesamaan makna maupun kata kunci.

Penerapan RAG dengan Hybrid Retrieval

Metode utama yang diterapkan dalam penelitian ini adalah *Retrieval-Augmented Generation* (RAG). RAG menggabungkan proses pengambilan informasi dari basis pengetahuan dengan kemampuan LLM dalam menghasilkan respons menggunakan bahasa alami. Dalam penelitian ini, RAG digunakan agar jawaban chatbot disusun berdasarkan konteks informasi yang diperoleh dari basis pengetahuan penyakit menular.

Mekanisme *retrieval* menggunakan pendekatan *hybrid retrieval* yang menggabungkan *semantic search* dan *keyword search*. Penggunaan kedua pendekatan tersebut dimaksudkan untuk memanfaatkan kemampuan pencarian berdasarkan makna sekaligus pencocokan istilah secara langsung.

Pada *semantic search*, pertanyaan pengguna diubah menjadi representasi vektor menggunakan *Sentence Transformer*. Sistem kemudian membandingkan representasi vektor pertanyaan dengan representasi vektor *chunk* pada basis pengetahuan untuk menemukan informasi yang memiliki kemiripan semantik. Pendekatan ini memungkinkan sistem menemukan informasi yang relevan meskipun pertanyaan dan dokumen tidak menggunakan kata yang identik.

Selain pencarian semantik, sistem menerapkan *keyword search* menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF memberikan bobot terhadap suatu istilah berdasarkan tingkat kemunculannya pada dokumen dan distribusinya dalam kumpulan dokumen. Secara umum, nilai TF-IDF dapat dinyatakan sebagai:

$$TFIDF(t, d) = TF(t, d) \times IDF(t)$$

dengan $TF(t,d)$ menunjukkan frekuensi istilah t pada dokumen d, sedangkan $IDF(t)$ menunjukkan tingkat kepentingan istilah berdasarkan distribusinya pada keseluruhan dokumen.

Hasil dari *semantic search* dan *keyword search* kemudian digabungkan menggunakan mekanisme *hybrid search* dengan pembobotan tertentu. Secara konseptual, kombinasi skor dapat dituliskan sebagai:

$$S_{hybrid} = \alpha\, S_{semantic} + \beta\, S_{keyword}$$

dengan $S_{semantic}$ merupakan skor pencarian semantik, $S_{keyword}$ merupakan skor pencarian berbasis kata kunci, sedangkan α dan β merupakan bobot yang digunakan dalam penggabungan kedua hasil pencarian. Nilai bobot mengikuti konfigurasi yang diterapkan pada implementasi sistem.

Hasil *hybrid retrieval* digunakan untuk menentukan informasi yang paling relevan sebagai konteks. Konteks tersebut kemudian digabungkan dengan pertanyaan pengguna dan diberikan kepada LLM sebagai masukan untuk menghasilkan jawaban. Dengan mekanisme ini, jawaban yang dihasilkan didasarkan pada informasi yang berhasil diambil dari basis pengetahuan, bukan hanya pada pengetahuan internal LLM.

Implementasi dan Alur Kerja Sistem

Implementasi sistem dilakukan dengan membangun chatbot berbasis web menggunakan Python sebagai bahasa pemrograman utama. Sistem dirancang sebagai satu kesatuan aplikasi yang mengintegrasikan antarmuka chatbot dengan komponen pemrosesan bahasa alami, *retrieval*, basis pengetahuan, dan LLM.

Alur kerja sistem dimulai ketika pengguna memasukkan pertanyaan dalam bahasa Indonesia melalui antarmuka chatbot. Pertanyaan tersebut kemudian diproses untuk menghasilkan representasi yang dapat digunakan dalam proses pencarian. Selanjutnya, sistem menjalankan *semantic search* dan *keyword search* untuk menemukan informasi yang relevan dari basis pengetahuan. Hasil kedua proses tersebut digabungkan melalui *hybrid retrieval* untuk memperoleh konteks yang paling sesuai dengan pertanyaan.

Konteks hasil *retrieval* kemudian dikombinasikan dengan pertanyaan pengguna dan diberikan kepada LLM. LLM menggunakan pertanyaan dan konteks tersebut untuk menghasilkan jawaban dalam bentuk bahasa alami. Jawaban akhir kemudian ditampilkan kepada pengguna melalui antarmuka chatbot berbasis web.

Pengujian dan Evaluasi Sistem

Pengujian dilakukan melalui dua pendekatan, yaitu pengujian kualitas respons menggunakan RAGAs dan pengujian fungsional menggunakan *Black-Box Testing*. Pengujian kualitas respons digunakan untuk mengetahui hubungan antara pertanyaan, konteks yang berhasil diambil, dan jawaban yang dihasilkan. Sementara itu, *Black-Box Testing* digunakan untuk memastikan fungsi sistem berjalan sesuai kebutuhan tanpa memeriksa struktur internal kode program.

Evaluasi kualitas respons menggunakan *Retrieval Augmented Generation Assessment* (RAGAs). Evaluasi ini digunakan untuk menilai tiga komponen utama pada sistem RAG, yaitu *query*, *context*, dan *answer*. Tiga aspek yang digunakan dalam evaluasi adalah *Faithfulness*, *Answer Relevance*, dan *Context Relevance*.

Faithfulness mengukur sejauh mana pernyataan dalam jawaban didukung oleh konteks yang diperoleh dari proses *retrieval*. Jika $S=\{S_1, S_2,... S_3\}$ merupakan kumpulan pernyataan pada jawaban dan (V) merupakan jumlah pernyataan yang dinilai didukung oleh konteks, maka skor *Faithfulness* dihitung dengan:

$$F = \frac{|V|}{|S|}$$

Nilai yang semakin tinggi menunjukkan bahwa semakin banyak pernyataan pada jawaban yang memiliki dukungan dari konteks yang diberikan.

Answer Relevance digunakan untuk mengukur kesesuaian jawaban terhadap pertanyaan pengguna. Penilaian dilakukan dengan membangkitkan sejumlah pertanyaan potensial berdasarkan jawaban yang dihasilkan, kemudian membandingkannya dengan

pertanyaan asli menggunakan fungsi kemiripan. Skor *Answer Relevance* dinyatakan sebagai:

$$AR = \frac{1}{n}\sum_{i=1}^{n} sim(q, q_i)$$

dengan (q) merupakan pertanyaan asli, (q_i) merupakan pertanyaan potensial yang dihasilkan dari jawaban, (sim(q,q_i) merupakan nilai kemiripan antara pertanyaan asli dan pertanyaan potensial, sedangkan (n) merupakan jumlah pertanyaan potensial.

Context Relevance digunakan untuk mengukur tingkat relevansi konteks yang diperoleh terhadap pertanyaan pengguna. Konteks dinilai relevan apabila informasi yang terdapat di dalamnya benar-benar diperlukan untuk menjawab pertanyaan dan meminimalkan informasi yang tidak diperlukan. Skor *Context Relevance* dihitung berdasarkan perbandingan jumlah kalimat relevan dengan jumlah keseluruhan kalimat dalam konteks:

$$CR = \frac{Total\ kalimat\ diekstrak}{Total\ Kalimat\ dalam\ konteks\ c(q)}$$

Ketiga indikator tersebut digunakan secara bersama-sama untuk menilai kualitas sistem RAG dari sisi kesesuaian jawaban dengan konteks, relevansi jawaban terhadap pertanyaan, serta relevansi konteks yang berhasil diperoleh.

Black-Box Testing digunakan untuk menguji fungsi sistem berdasarkan masukan dan keluaran yang dihasilkan melalui antarmuka pengguna. Pengujian tidak berfokus pada struktur kode program, tetapi pada kemampuan sistem dalam menjalankan fungsi sesuai kebutuhan.

Skenario pengujian meliputi pemberian pertanyaan yang sesuai maupun tidak sesuai dengan korpus data, pengujian pertanyaan di luar cakupan basis pengetahuan, pengujian kemampuan sistem memberikan respons ketika informasi yang diminta tidak tersedia, pengujian relevansi jawaban terhadap pertanyaan, pemeriksaan keberadaan sumber informasi pada respons, serta pengujian waktu respons sistem.

Pengujian respons dilakukan terhadap pertanyaan mengenai delapan penyakit yang menjadi cakupan basis pengetahuan, yaitu TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Sistem dinyatakan berfungsi apabila fitur utama dapat berjalan sesuai spesifikasi dan menghasilkan respons yang sesuai dengan masukan pengguna.

Lingkungan Implementasi

Pengembangan dan pengujian sistem dilakukan pada lingkungan pengembangan berbasis komputer dengan memanfaatkan sumber informasi digital. Python digunakan sebagai bahasa pemrograman utama untuk membangun dan mengintegrasikan komponen sistem. Visual Studio Code digunakan sebagai *code editor* dalam proses pengembangan.

Lingkungan pengembangan juga memanfaatkan pustaka pendukung untuk pemrosesan bahasa alami, pembentukan *embedding*, pengelolaan basis pengetahuan, proses *retrieval*, serta integrasi LLM dan RAG. Komponen sistem diintegrasikan sehingga proses penerimaan pertanyaan, pemrosesan teks, pencarian informasi, penyediaan

konteks, pembangkitan jawaban, dan penyajian respons dapat berjalan dalam satu alur sistem.

Hasil

Hasil Pengumpulan dan Pembentukan Basis Pengetahuan

Hasil pengumpulan data menghasilkan basis pengetahuan yang digunakan sebagai sumber informasi chatbot mengenai penyakit menular. Data berbentuk dokumen informasi kesehatan yang diperoleh dari dua sumber, yaitu Halodoc dan Alodokter. Berdasarkan hasil penelitian, basis pengetahuan mencakup delapan jenis penyakit menular, yaitu Tuberkulosis (TBC), Demam Berdarah Dengue (DBD), HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19.

Tabel 1. *Cakupan Basis Pengetahuan Penyakit Menular*

Jenis Penyakit	Sumber Informasi
Tuberkulosis (TBC)	Halodoc dan Alodokter
Demam Berdarah Dengue (DBD)	Halodoc dan Alodokter
HIV/AIDS	Halodoc dan Alodokter
Influenza	Halodoc dan Alodokter
Campak	Halodoc dan Alodokter
Cacar air	Halodoc dan Alodokter
Malaria	Halodoc dan Alodokter
COVID-19	Halodoc dan Alodokter

Sumber data tersebut selanjutnya diproses untuk membentuk basis pengetahuan yang dapat digunakan oleh mekanisme RAG. Dokumen melalui proses pembersihan, normalisasi, dan pemecahan menjadi *chunk*. Setiap *chunk* kemudian direpresentasikan dalam bentuk *embedding* sehingga dapat digunakan dalam proses pencarian berdasarkan kemiripan semantik. Hasil pengolahan tersebut menjadi sumber konteks yang digunakan ketika chatbot menerima pertanyaan pengguna.

Hasil penelitian pada tahap ini menunjukkan bahwa informasi dari berbagai dokumen dapat disusun menjadi basis pengetahuan yang terarah berdasarkan cakupan penyakit yang telah ditentukan. Pembatasan cakupan tersebut juga menjadi dasar bagi sistem dalam memberikan respons terhadap pertanyaan pengguna.

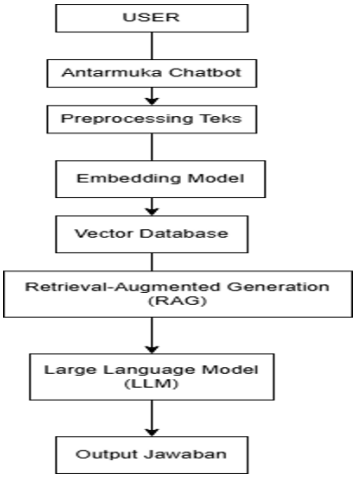
Hasil Implementasi Hybrid Retrieval dan RAG

Sistem berhasil mengimplementasikan mekanisme *Retrieval-Augmented Generation* (RAG) dengan pendekatan *hybrid retrieval*. Proses dimulai ketika pengguna memasukkan pertanyaan melalui antarmuka chatbot. Pertanyaan kemudian diproses dan direpresentasikan dalam bentuk *embedding* menggunakan *Sentence Transformer*.

Proses pencarian dilakukan menggunakan dua mekanisme, yaitu *semantic search* dan *keyword search*. *Semantic search* digunakan untuk menemukan informasi berdasarkan kemiripan makna antara pertanyaan pengguna dan dokumen pada basis pengetahuan. Sementara itu, *keyword search* menggunakan TF-IDF untuk melakukan pencocokan berdasarkan kata atau istilah yang terdapat pada pertanyaan.

Hasil dari kedua mekanisme pencarian kemudian digabungkan menggunakan pendekatan *hybrid search* dengan pembobotan tertentu. Hasil penggabungan digunakan untuk memperoleh konteks yang relevan. Konteks tersebut selanjutnya diberikan kepada LLM bersama dengan pertanyaan pengguna untuk menghasilkan jawaban.

Hasil implementasi menunjukkan bahwa keseluruhan proses tersebut dapat berjalan dalam satu alur. Sistem mampu menerima pertanyaan, melakukan proses pencarian informasi, meneruskan konteks hasil *retrieval* kepada LLM, dan menampilkan jawaban kepada pengguna melalui antarmuka chatbot.



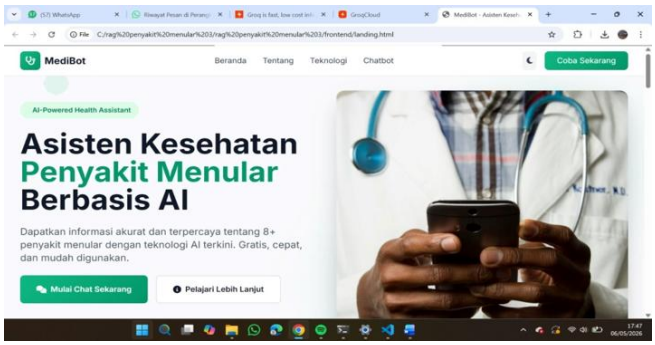
Gambar 1. *Alur Implementasi RAG dan Proses Retrieval pada Sistem*

Hasil tersebut memperlihatkan bahwa komponen *retrieval* dan *generation* telah berhasil diintegrasikan. Dengan demikian, chatbot tidak hanya berfungsi sebagai antarmuka percakapan, tetapi juga memiliki mekanisme pencarian informasi pada basis pengetahuan sebelum jawaban dibangkitkan oleh LLM.

Hasil Implementasi Antarmuka Chatbot

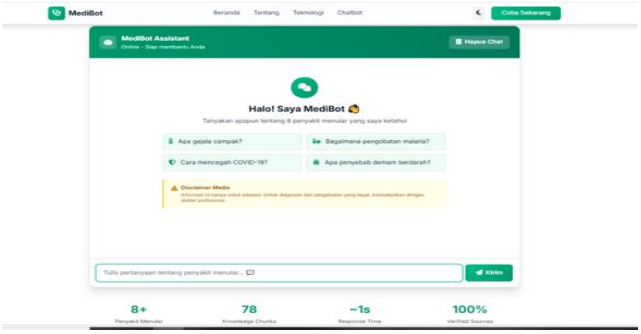
Sistem chatbot berhasil diimplementasikan dalam bentuk aplikasi berbasis web. Antarmuka sistem dirancang dengan tampilan sederhana yang berfokus pada interaksi pengguna dengan chatbot. Halaman utama menyediakan informasi mengenai fungsi sistem serta navigasi menuju halaman chatbot.

Pada halaman chatbot, pengguna dapat memasukkan pertanyaan mengenai penyakit menular melalui kolom percakapan. Sistem juga menyediakan contoh pertanyaan untuk membantu pengguna memahami jenis informasi yang dapat ditanyakan.



Gambar 2. *Halaman Utama Sistem MediBot*

Hasil implementasi menunjukkan bahwa halaman utama dapat menampilkan informasi dan navigasi sistem dengan baik. Antarmuka dirancang agar pengguna dapat memahami fungsi chatbot sebelum melakukan interaksi.



Gambar 3. *Halaman Chatbot*

Halaman chatbot menyediakan ruang percakapan untuk memasukkan pertanyaan. Setelah pengguna mengirimkan pertanyaan, sistem meneruskan masukan tersebut ke proses pemrosesan dan *retrieval* sebelum menghasilkan respons.

Hasil Pengujian Pertanyaan dan Respons Chatbot

Pengujian dilakukan dengan memberikan pertanyaan mengenai penyakit yang terdapat pada basis pengetahuan. Berdasarkan hasil pengujian, sistem mampu menerima pertanyaan melalui antarmuka chatbot dan menghasilkan respons berdasarkan informasi yang tersedia pada basis pengetahuan.

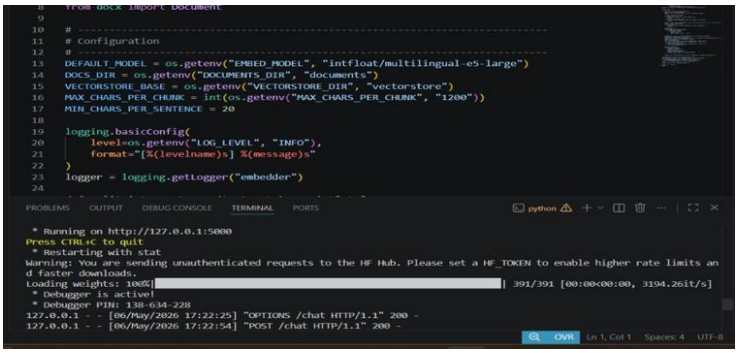
Tabel 2. *Hasil Pengujian Respons Berdasarkan Cakupan Penyakit*

Penyakit	Input Pertanyaan	Respon Chatbot	Status
TBC	Pertanyaan mengenai TBC	Sistem memberikan respons mengenai TBC	Berhasil
DBD	Pertanyaan mengenai DBD	Sistem memberikan respons mengenai DBD	Berhasil
HIV/AIDS	Pertanyaan mengenai HIV/AIDS	Sistem memberikan respons mengenai HIV/AIDS	Berhasil
Influenza	Pertanyaan mengenai influenza	Sistem memberikan respons mengenai influenza	Berhasil
Campak	Pertanyaan mengenai campak	Sistem memberikan respons mengenai campak	Berhasil
Cacar Air	Pertanyaan mengenai cacar air	Sistem memberikan respons mengenai cacar air	Berhasil
Malaria	Pertanyaan mengenai malaria	Sistem memberikan respons mengenai malaria	Berhasil
COVID-19	Pertanyaan mengenai COVID-19	Sistem memberikan respons mengenai COVID-19	Berhasil

Pengujian tersebut menunjukkan bahwa chatbot dapat memberikan respons terhadap pertanyaan yang berkaitan dengan seluruh penyakit yang termasuk dalam cakupan basis pengetahuan. Sistem menerima input pengguna tanpa kendala dan menampilkan jawaban secara *real-time* pada halaman percakapan. Secara keseluruhan, hasil pengujian terhadap delapan penyakit menunjukkan bahwa sistem dapat menjalankan fungsi utama chatbot untuk menerima pertanyaan dan menghasilkan respons pada topik yang termasuk dalam cakupan basis pengetahuan.

Hasil Pengujian Backend

Selain pengujian melalui antarmuka pengguna, dilakukan pengujian terhadap backend untuk memastikan proses pengolahan data dan komunikasi antarkomponen sistem berjalan dengan baik. Pengujian menunjukkan bahwa backend mampu menerima dan memproses permintaan yang berasal dari sistem chatbot serta mengembalikan respons tanpa mengalami *error*.



Gambar 4. *Hasil pengujian backend sistem*

Hasil tersebut menunjukkan bahwa komponen backend dapat menjalankan proses pengolahan data dan menghasilkan respons chatbot sesuai dengan alur sistem yang dirancang. Dengan demikian, komunikasi antara antarmuka dan komponen pemrosesan chatbot dapat berjalan dengan baik.

Rekapitulasi Hasil Pengujian Sistem

Berdasarkan keseluruhan pengujian yang dilakukan, sistem berhasil menjalankan fungsi utama yang dirancang. Rekapitulasi hasil pengujian disajikan pada tabel 3.

Tabel 3. *Rekapitulasi Hasil Pengujian Sistem*

Komponen yang Diuji	Hasil Pengujian	Support
Halaman utama	Dapat ditampilkan dan digunakan	Berhasil
Halaman chatbot	Dapat digunakan untuk memasukkan pertanyaan	Berhasil
Input pertanyaan	Sistem dapat menerima pertanyaan pengguna	Berhasil
Respon TBC	Sistem menghasilkan respons	Berhasil
Respon DBD	Sistem menghasilkan respons	Berhasil
Respons HIV/AIDS	Sistem menghasilkan respons	Berhasil
Respons influenza	Sistem menghasilkan respons	Berhasil
Respons campak	Sistem menghasilkan respons	Berhasil
Respons cacar air	Sistem menghasilkan respons	Berhasil
Respons malaria	Sistem menghasilkan respons	Berhasil
Respons COVID-19	Sistem menghasilkan respons	Berhasil
Backend	Pengolahan data dan komunikasi berjalan tanpa *error*	Berhasil

Hasil pengujian menunjukkan bahwa seluruh fungsi utama yang diuji dapat berjalan sesuai dengan rancangan. Chatbot mampu menerima masukan pengguna, memproses pertanyaan, menghasilkan respons, dan menampilkan respons melalui antarmuka web. Pengujian backend juga menunjukkan bahwa proses pengolahan data dan komunikasi sistem berjalan tanpa *error*.

Pembahasan

Implementasi RAG pada Chatbot Informasi Penyakit Menular

Hasil penelitian menunjukkan bahwa penerapan *Retrieval-Augmented Generation* (RAG) dapat digunakan untuk membangun chatbot yang memberikan informasi mengenai penyakit menular melalui pendekatan berbasis pengetahuan. Sistem yang dikembangkan tidak hanya mengandalkan kemampuan model bahasa dalam menghasilkan respons, tetapi terlebih dahulu melakukan proses retrieval terhadap informasi yang terdapat dalam basis pengetahuan. Informasi yang diperoleh dari proses retrieval kemudian digunakan sebagai konteks bagi model bahasa dalam menghasilkan

jawaban terhadap pertanyaan pengguna. Pendekatan tersebut sesuai dengan karakteristik RAG yang menggabungkan mekanisme pencarian informasi dengan kemampuan generatif model bahasa.

Penerapan RAG menjadi relevan pada domain kesehatan karena kualitas jawaban chatbot sangat dipengaruhi oleh sumber informasi yang digunakan. Chatbot dapat membantu meningkatkan akses terhadap komunikasi dan informasi kesehatan, tetapi kualitas informasi yang diberikan tetap bergantung pada sumber data dan algoritma yang digunakan (Noviandhani et al., 2024). Temuan tersebut sejalan dengan penelitian ini karena basis pengetahuan chatbot dibangun dari dokumen informasi penyakit yang berasal dari sumber kesehatan yang digunakan dalam penelitian, yaitu Halodoc dan Alodokter. Dengan demikian, proses *generation* tidak berdiri sendiri, tetapi diarahkan menggunakan informasi yang telah tersedia dalam basis pengetahuan.

Hasil implementasi juga menunjukkan bahwa chatbot dapat menerima pertanyaan pengguna, memprosesnya, kemudian menghasilkan respons secara otomatis melalui antarmuka web. Pengujian dilakukan terhadap delapan topik penyakit, yaitu TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Seluruh pertanyaan uji tersebut dapat diproses oleh sistem dan menghasilkan respons pada antarmuka chatbot. Hasil tersebut memperlihatkan bahwa integrasi antara *frontend*, *backend*, *retrieval*, dan *generation* telah berjalan sesuai dengan rancangan sistem.

Hasil penelitian ini sejalan dengan penelitian mengenai penerapan RAG pada chatbot kesehatan yang menunjukkan bahwa mekanisme retrieval memungkinkan chatbot menghasilkan jawaban yang lebih relevan dan kontekstual karena model memperoleh informasi dari dokumen yang telah disediakan (Samudra et al., 2025). Kesamaan tersebut memperlihatkan bahwa RAG dapat diterapkan pada berbagai domain informasi kesehatan dengan menyesuaikan basis pengetahuan terhadap kebutuhan informasi yang ingin diberikan.

Peran Hybrid Retrieval dalam Pencarian Informasi

Salah satu karakteristik utama sistem yang dikembangkan adalah penggunaan *hybrid retrieval* yang menggabungkan *semantic search* dan *keyword search*. Dalam penelitian ini, *semantic search* menggunakan representasi *embedding* yang dihasilkan melalui Sentence Transformer, sedangkan pencarian berbasis kata menggunakan TF-IDF. Kedua hasil pencarian kemudian digabungkan menggunakan pembobotan tertentu sebelum konteks yang diperoleh diberikan kepada model bahasa.

Penggunaan dua pendekatan retrieval tersebut memiliki fungsi yang saling melengkapi. *Semantic search* memungkinkan sistem menemukan informasi berdasarkan kedekatan makna antara pertanyaan pengguna dan dokumen, sehingga pencarian tidak sepenuhnya bergantung pada kesamaan kata secara literal. Sebaliknya, *keyword search* melalui TF-IDF memberikan kemampuan pencocokan berdasarkan kata atau istilah yang terdapat pada pertanyaan. Kombinasi keduanya menjadi relevan untuk pertanyaan kesehatan karena pengguna dapat menggunakan istilah yang berbeda ketika menanyakan penyakit yang sama.

Hasil implementasi menunjukkan bahwa mekanisme *hybrid retrieval* dapat digunakan untuk memperoleh konteks yang relevan sebelum proses pembangkitan jawaban. Namun, hasil penelitian ini belum memberikan pengukuran kuantitatif yang secara khusus

membandingkan performa *semantic search*, *keyword search*, dan *hybrid retrieval*. Oleh karena itu, pembahasan mengenai *hybrid retrieval* dalam penelitian ini dibatasi pada keberhasilan implementasi dan hasil pengujian sistem, bukan pada klaim peningkatan akurasi *retrieval* secara numerik.

Pendekatan tersebut memiliki kesamaan dengan penelitian mengenai chatbot diagnosis penyakit berbasis RAG yang menggunakan *hybrid retrieval* dengan mengombinasikan retrieval berbasis *embedding* dan metode pencarian berbasis kata (Abdurrazzaq et al., 2025). Penelitian tersebut menunjukkan bahwa kombinasi beberapa mekanisme retrieval dapat digunakan untuk memperoleh informasi yang lebih relevan bagi proses pembangkitan jawaban (Abdurrazzaq et al., 2025). Perbedaannya terletak pada teknik yang digunakan karena penelitian tersebut menggunakan kombinasi FAISS dan BM25, sedangkan penelitian ini menggunakan Sentence Transformer dan TF-IDF. Perbedaan tersebut menunjukkan bahwa konsep *hybrid retrieval* dapat diterapkan menggunakan teknik yang berbeda sesuai dengan karakteristik data dan kebutuhan sistem.

Kesesuaian Respons terhadap Berbagai Topik Penyakit

Pengujian terhadap delapan penyakit menunjukkan bahwa sistem mampu memproses pertanyaan yang memiliki topik berbeda dalam satu domain, yaitu informasi penyakit menular. Pertanyaan yang diberikan mencakup beberapa penyakit yang telah dimasukkan ke dalam basis pengetahuan, dan sistem menghasilkan respons berdasarkan informasi yang tersedia. Hal ini menunjukkan bahwa ruang lingkup basis pengetahuan memiliki peran penting dalam menentukan kemampuan sistem dalam menjawab pertanyaan.

Keberadaan basis pengetahuan yang terarah juga membantu membatasi informasi yang diproses oleh sistem. Pada halaman utama, sistem memberikan panduan kepada pengguna agar memasukkan pertanyaan yang berkaitan dengan informasi penyakit menular. Pembatasan tersebut sesuai dengan tujuan pengembangan chatbot sebagai sistem informasi pada domain tertentu, bukan sebagai chatbot umum yang menjawab seluruh jenis pertanyaan.

Hasil tersebut sejalan dengan penelitian mengenai penggunaan Retrieval-Augmented Large Language Model untuk pemeriksaan fakta COVID-19 yang menunjukkan bahwa pemanfaatan sumber informasi sebagai konteks dapat mendukung penggunaan model bahasa dalam domain kesehatan yang membutuhkan informasi berbasis sumber (Li et al., 2025). Penelitian lain juga menunjukkan bahwa RAG dapat digunakan untuk meningkatkan kualitas respons model pada konsultasi di bidang radiologi melalui pemanfaatan informasi eksternal sebagai konteks bagi model bahasa (Wada et al., 2025). Kesamaan dari penelitian-penelitian tersebut adalah penggunaan informasi eksternal sebagai dasar atau konteks bagi model bahasa sehingga respons tidak sepenuhnya bergantung pada pengetahuan internal model.

Meskipun demikian, kemampuan chatbot dalam penelitian ini tidak dapat diartikan sebagai kemampuan untuk melakukan diagnosis medis. Sistem hanya dirancang untuk memberikan informasi mengenai penyakit yang terdapat dalam basis pengetahuan. Batasan ini penting karena informasi kesehatan memiliki konsekuensi yang lebih tinggi dibandingkan penggunaan chatbot pada domain umum. Penggunaan chatbot berbasis AI

dalam bidang kesehatan memiliki manfaat sekaligus keterbatasan, sehingga kualitas dan ruang lingkup informasi perlu diperhatikan dalam penerapannya (Hayat et al., 2024).

Hasil Pengujian Fungsional Sistem

Hasil pengujian menunjukkan bahwa komponen utama sistem, meliputi halaman utama, halaman chatbot, input pertanyaan, proses pemberian respons, dan *backend* dapat berjalan tanpa mengalami *error*. Pengujian juga memperlihatkan bahwa pertanyaan mengenai TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19 dapat diproses oleh sistem. Dari sisi fungsional, hasil tersebut menunjukkan bahwa alur sistem dari input pengguna hingga keluaran chatbot telah dapat dijalankan sesuai rancangan.

Keberhasilan pengujian fungsional memiliki arti penting karena penerapan metode RAG tidak hanya ditentukan oleh mekanisme retrieval dan generation, tetapi juga oleh integrasi seluruh komponen sistem. Proses retrieval harus dapat menerima pertanyaan pengguna, memperoleh konteks dari basis pengetahuan, dan meneruskan konteks tersebut ke model bahasa. Pada sisi lain, hasil generation harus dapat dikembalikan melalui backend dan ditampilkan pada frontend. Dengan demikian, keberhasilan pengujian backend menjadi bagian penting untuk memastikan bahwa mekanisme RAG dapat digunakan dalam aplikasi chatbot berbasis web.

Hasil tersebut mendukung penelitian mengenai perancangan sistem chatbot RAG berbasis website yang menunjukkan bahwa pendekatan RAG dapat diimplementasikan dalam sistem chatbot berbasis web (Cahyanti & Raya, 2025). Penerapan RAG dalam pemrosesan bahasa Indonesia juga menunjukkan bahwa pendekatan *retrieval-augmented* dapat digunakan untuk mendukung pemanfaatan model bahasa dalam konteks bahasa Indonesia (Tohir et al., 2024). Penelitian ini memperluas penerapan tersebut pada domain informasi penyakit menular dengan basis pengetahuan yang secara khusus disusun untuk kebutuhan chatbot kesehatan.

Relevansi Sistem terhadap Penyediaan Informasi Kesehatan Digital

Dari sisi penerapan, sistem yang dikembangkan memiliki potensi sebagai media penyedia informasi kesehatan digital karena pengguna dapat mengajukan pertanyaan secara langsung melalui antarmuka percakapan. Dibandingkan dengan pencarian informasi secara manual pada beberapa halaman web, model interaksi berbentuk chatbot memungkinkan pengguna menyampaikan pertanyaan dalam bentuk bahasa alami dan memperoleh respons melalui satu antarmuka.

Temuan tersebut sejalan dengan kajian mengenai efektivitas chatbot dalam layanan kesehatan yang menunjukkan bahwa chatbot dapat berperan dalam meningkatkan akses terhadap layanan komunikasi kesehatan (Noviandhani et al., 2024). Dalam penelitian ini, manfaat tersebut diwujudkan melalui sistem yang menyediakan informasi mengenai beberapa penyakit menular dalam satu basis pengetahuan. Pengguna tidak perlu memilih mekanisme pencarian yang berbeda untuk setiap penyakit karena pertanyaan dapat disampaikan melalui antarmuka chatbot yang sama.

Namun, manfaat tersebut tetap bergantung pada kelengkapan dan kualitas basis pengetahuan. Penambahan dokumen kesehatan yang lebih banyak dan lebih mutakhir dapat memperluas cakupan pertanyaan yang mampu ditangani oleh sistem. Hal ini juga menjadi salah satu arah pengembangan yang tercantum dalam penelitian, yaitu

menambah data dan informasi kesehatan agar sistem dapat memberikan informasi yang lebih lengkap.

Keterbatasan Hasil Penelitian

Meskipun sistem berhasil diimplementasikan dan seluruh pengujian fungsional yang dilaporkan menunjukkan hasil yang baik, penelitian ini masih memiliki beberapa keterbatasan. Pertama, basis pengetahuan masih terbatas pada delapan penyakit menular dan dokumen yang digunakan berasal dari sumber yang telah ditentukan. Kondisi tersebut menyebabkan kemampuan chatbot sangat bergantung pada cakupan informasi yang tersedia dalam basis pengetahuan.

Kedua, penelitian ini belum menyajikan hasil kuantitatif dari evaluasi RAGAs yang meliputi *Faithfulness*, *Answer Relevance*, dan *Context Relevance*. Oleh karena itu, keberhasilan sistem dalam menghasilkan respons tidak dapat disimpulkan sebagai tingkat kualitas numerik tertentu berdasarkan ketiga metrik tersebut. Hasil yang tersedia dalam penelitian lebih menekankan pada keberhasilan implementasi dan pengujian fungsional sistem.

Ketiga, penelitian ini belum melakukan eksperimen komparatif yang mengukur performa *semantic search*, *keyword search*, dan *hybrid retrieval* secara terpisah. Dengan demikian, penggunaan *hybrid retrieval* dapat dinilai telah berhasil diimplementasikan, tetapi besarnya kontribusi masing-masing metode terhadap kualitas retrieval belum dapat ditentukan secara kuantitatif.

Keterbatasan tersebut sekaligus memberikan arah pengembangan penelitian berikutnya. Penelitian selanjutnya dapat memperluas basis pengetahuan, menggunakan model AI yang lebih mutakhir, melakukan evaluasi RAGAs secara kuantitatif, serta membandingkan performa setiap strategi retrieval. Pengujian dengan jumlah pertanyaan yang lebih besar dan melibatkan evaluasi pengguna juga dapat dilakukan untuk memperoleh gambaran yang lebih komprehensif mengenai kualitas dan kegunaan chatbot.

Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi RAG dan *hybrid retrieval* dapat diimplementasikan untuk membangun chatbot berbasis web yang menyediakan informasi mengenai penyakit menular. Penggunaan Sentence Transformer pada *semantic search* dan TF-IDF pada *keyword search* memungkinkan sistem memanfaatkan dua karakteristik pencarian yang berbeda sebelum informasi diberikan kepada model bahasa sebagai konteks. Integrasi tersebut menghasilkan sistem yang mampu menerima pertanyaan pengguna dan memberikan respons untuk delapan topik penyakit yang diuji.

Temuan penelitian memperkuat hasil penelitian terdahulu yang menunjukkan bahwa RAG dapat digunakan untuk mendukung chatbot pada domain kesehatan dan domain pengetahuan tertentu (Li et al., 2025; Wada et al., 2025). Perbedaan penelitian ini terletak pada fokusnya terhadap informasi penyakit menular di Indonesia serta penggunaan kombinasi *semantic search* dan TF-IDF sebagai mekanisme *hybrid retrieval*. Meskipun demikian, hasil penelitian belum cukup untuk menyatakan bahwa pendekatan tersebut memiliki performa retrieval atau kualitas respons yang lebih tinggi secara kuantitatif dibandingkan metode lainnya. Oleh karena itu, kontribusi utama penelitian ini adalah pada implementasi dan integrasi arsitektur RAG dengan *hybrid retrieval* dalam chatbot

informasi penyakit menular, sedangkan pengukuran performa yang lebih mendalam masih menjadi ruang untuk penelitian selanjutnya.

Kesimpulan

Penelitian ini bertujuan mengembangkan chatbot berbasis *Retrieval-Augmented Generation* (RAG) dengan pendekatan *hybrid retrieval* untuk menyediakan informasi mengenai penyakit menular di Indonesia secara lebih relevan dan kontekstual. Permasalahan yang menjadi dasar penelitian adalah kebutuhan akan sistem informasi kesehatan berbasis percakapan yang mampu memanfaatkan sumber pengetahuan tertentu sehingga respons yang diberikan tidak hanya bergantung pada pengetahuan internal *Large Language Model* (LLM).

Hasil penelitian menunjukkan bahwa chatbot berhasil diimplementasikan dalam bentuk aplikasi berbasis web dengan basis pengetahuan yang mencakup delapan penyakit menular, yaitu TBC, DBD, HIV/AIDS, influenza, campak, cacar air, malaria, dan COVID-19. Pendekatan *hybrid retrieval* yang menggabungkan pencarian semantik menggunakan *Sentence Transformer* dan pencarian berbasis kata kunci menggunakan TF-IDF dapat digunakan untuk memperoleh konteks yang relevan sebelum proses pembentukan respons oleh LLM. Pengujian terhadap pertanyaan terkait penyakit menunjukkan bahwa sistem dapat menerima pertanyaan dan menghasilkan respons sesuai dengan konteks informasi yang tersedia, sementara pengujian *backend* menunjukkan bahwa proses sistem berjalan tanpa kesalahan.

Secara praktis, penelitian ini menunjukkan bahwa integrasi RAG dan *hybrid retrieval* dapat menjadi pendekatan yang potensial dalam pengembangan layanan informasi kesehatan berbasis chatbot. Namun, basis pengetahuan masih terbatas pada delapan penyakit dan sumber informasi tertentu, serta evaluasi kuantitatif kualitas respons masih perlu diperkuat. Penelitian selanjutnya dapat memperluas cakupan penyakit dan sumber pengetahuan, menggunakan model AI yang lebih mutakhir, serta melakukan evaluasi kualitas retrieval dan respons secara kuantitatif.

Acknowledgment

-

Daftar Pustaka

Abdurrazzaq, M. A., Tjiong, E. L., Fasya, A., Hiu, M., & Tanuwidjaya, J. (2025). An Indonesian chatbot for disease diagnosis using retrieval-augmented generation. *INOVTEK Polbeng - Seri Informatika, 10*(3), 1877–1887. https://doi.org/10.35314/9nnkn955

Aggarwal, A., Tam, C. C., Wu, D., Li, X., & Qiao, S. (2023). Artificial intelligence–based chatbots for promoting health behavioral changes: Systematic review. *Journal of Medical Internet Research, 25*, e40789. https://doi.org/10.2196/40789

Cahyanti, F. L. D., & Raya, R. D. A. (2025). Perancangan sistem informasi chatbot retrieval augmented generation berbasis website pada PT. Revolusi Cita Edukasi. *Indonesian Journal of Computer Science, 4*(1), 15–21. https://doi.org/10.31294/m75d4782

Devine, P. (2025). *ALoFTRAG: Automatic local fine tuning for retrieval augmented generation* [Preprint]. arXiv

Haromain, I., Munir, S., & Rahmah, A. (2025). Analisa prompt engineering pada large language model dengan retrieval-augmented generation untuk informasi obat dan vitamin. *Indonesian Journal of Computer Science, 4*(2), 144–153. https://doi.org/10.31294/ijcs.v4i2.10005

Hayat, J., Lari, M., AlHerz, M., & Lari, A. (2024). The utility and limitations of artificial intelligence-powered chatbots in healthcare. *Cureus, 16*(11), e73127. https://doi.org/10.7759/cureus.73127

Jayarana, I. G. N. A., Darma, I. G. W., Juliantara, I. W. A., & Putra, I. M. A. W. (2025). Study literatur information retrieval model: Teknik dan aplikasi. *Jurnal Sutasoma, 3*(2), 61–69. https://doi.org/10.58878/sutasoma.v3i2.392

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38. https://doi.org/10.1145/3571730

Kementerian Kesehatan Republik Indonesia. (2024, November 12). *Menkes tegaskan Indonesia serius tangani TBC*. https://www.kemkes.go.id/id/menkes-tegaskan-indonesia-serius-tangani-tbc

Kurniawan, M. H., Handiyani, H., Nuraini, T., Hariyati, R. T. S., & Sutrisno, S. (2024). A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. *Annals of Medicine, 56*(1), 2302980. https://doi.org/10.1080/07853890.2024.2302980

Laymouna, M., Ma, Y., Lessard, D., Schuster, T., Engler, K., & Lebouché, B. (2024). Roles, users, benefits, and limitations of chatbots in health care: Rapid review. *Journal of Medical Internet Research, 26*, e56930. https://doi.org/10.2196/56930

Li, H., Huang, J., Ji, M., Yang, Y., & An, R. (2025). Use of retrieval-augmented large language model for COVID-19 fact-checking: Development and usability study. *Journal of Medical Internet Research, 27*, e66098. https://doi.org/10.2196/66098

Li, J., Dada, A., Puladi, B., Kleesiek, J., & Egger, J. (2024). ChatGPT in healthcare: A taxonomy and systematic review. *Computer Methods and Programs in Biomedicine, 245*, 108013. https://doi.org/10.1016/j.cmpb.2024.108013

Li, Y., Liang, S., Zhu, B., Liu, X., Li, J., Chen, D., Qin, J., & Bressington, D. (2023). Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: A systematic review of randomized controlled trials. *International Journal of Nursing Studies, 143*, 104494. https://doi.org/10.1016/j.ijnurstu.2023.104494

Murtiyoso, M., Tahyudin, I., & Berlilana, B. (2025). A systematic review of retrieval-augmented generation for enhancing domain-specific knowledge in large language models. *Sinkron: Jurnal dan Penelitian Teknik Informatika, 9*(2), 969–977. https://doi.org/10.33395/sinkron.v9i2.14824

Noviandhani, W., Mita, S. R., & Ratnawati, A. (2024). Efektivitas layanan kesehatan menggunakan chatbot dan faktor-faktor yang memengaruhi kepuasan

penggunanya: Tinjauan literatur. Indonesian Journal of Clinical Pharmacy, 13(2), 97–106. https://doi.org/10.15416/ijcp.2024.13.2.57243

Samudra, G., Turmudi. A, & Ermanto. (2025). Implementasi retrieval augmented generation (RAG) dalam perancangan chatbot kesehatan pencernaan. *JSAI (Journal Scientific and Applied Informatics)*, *8*(1), 181–188. https://doi.org/10.36085/jsai.v8i1.7678

Septian, I., Ariessanti, H. D., Herwanto, A., & Go, R. Y. (2026). Perancangan dan Implementasi Aplikasi Virtual Try-On Pakaian dengan Fitur Unggah Model 3D Berbasis Web Menggunakan React dan Three.js. *Journal Artificial: Informatika Dan Sistem Informasi*, *4*(1), 123–136. https://doi.org/10.54065/artificial.1142

Tohir, H., Merlina, N., & Haris, M. (2024). Utilizing retrieval-augmented generation in large language models to enhance Indonesian language NLP. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, *10*(2), 352–360. https://doi.org/10.33480/jitk.v10i2.5916

Wada, A., Tanaka, Y., Nishizawa, M., Yamamoto, A., Akashi, T., Hagiwara, A., Hayakawa, Y., Kikuta, J., Shimoji, K., Sano, K., Kamagata, K., Nakanishi, A., & Aoki, S. (2025). Retrieval-augmented generation elevates local LLM quality in radiology contrast media consultation. *npj Digital Medicine, 8*, 395. https://doi.org/10.1038/s41746-025-01802-z

Xu, L., Sanders, L., Li, K., & Chow, J. C. L. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer, 7*(4), e27850. https://doi.org/10.2196/27850

Yang, R., Tan, T. F., Lu, W., Thirunavukarasu, A. J., Ting, D. S. W., & Liu, N. (2023). Large language models in health care: Development, applications, and challenges. *Health Care Science, 2*(4), 255–263. https://doi.org/10.1002/hcs2.61