

Implementasi *Semantic Retrieval* Berbasis *Embedding* dan *LLM* pada Aplikasi Web Keuangan Pribadi

Jimmy Putra Alam dan Akhmad Budi*)

Program Studi Sistem Informasi, Institut Bisnis dan Informatika Kwik Kian Gie, Jl. Yos Sudarso Kav 87, Sunter Jakarta 14350, Indonesia.

*) Surel korespondensi : akhmad.budi@kwikkiangie.ac.id

Abstract. *This study developed a web-based personal finance application integrated with semantic retrieval using embedding and Large Language Models to address limitations in conventional keyword-based search. The system was developed using the Extreme Programming approach to ensure iterative development, rapid feedback, and adaptability to user requirements. The application was designed using a client-server architecture and implemented to support transaction recording, financial management, and contextual data retrieval. Financial transaction data were preprocessed and converted into vector embeddings to capture semantic meaning, enabling more flexible and accurate search using natural language queries. The system utilized a vector database with pgvector extension to store embeddings and perform similarity search. Evaluation was conducted using the RAGAs framework with 100 financial transaction data and 20 query scenarios. The results showed that the system achieved a context recall of 0.81, context precision of 0.70, faithfulness of 0.83, and answer relevancy of 0.84. These results indicate that the system was able to retrieve relevant contextual information and generate consistent and accurate responses. Furthermore, the implementation demonstrated improved effectiveness in organizing, managing, and retrieving financial data compared to traditional methods. Overall, the proposed system successfully provided a more efficient and context-aware solution for personal financial management and information retrieval.*

Keywords: *personal finance, semantic retrieval, embeddings, web application, vector database*



This work is licensed under Attribution-NonCommercial-ShareAlike 4.0 International. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Diterbitkan oleh LPPM Institut Bisnis dan Informatika Kwik Kian Gie, Jl. Yos Sudarso Kav 87, Sunter Jakarta 14350, Indonesia.
DOI : <https://doi.org/10.46806/jib.v15i1.2027>

1. Pendahuluan

Dalam kehidupan sehari-hari, setiap orang berhadapan dengan aktivitas keuangan seperti menerima pendapatan, melakukan pengeluaran, serta berusaha menyeimbangkan keduanya agar kebutuhan dapat terpenuhi. Namun, banyak individu masih menghadapi kesulitan dalam mengelola keuangan pribadinya secara teratur. Sebagian besar hanya mengandalkan ingatan untuk mengetahui penggunaan uang mereka, tanpa adanya pencatatan yang sistematis. Akibatnya, mereka sering kehilangan kendali terhadap arus kas, tidak mengetahui secara pasti tujuan pengeluaran, dan sulit menilai kondisi keuangan pribadi secara objektif.

Ketidakteraturan dalam mencatat transaksi menyebabkan banyak orang tidak menyadari pola pengeluaran mereka sendiri. Pendapatan yang diterima setiap bulan sering kali habis tanpa diketahui secara jelas penggunaannya. Hal ini membuat

seseorang kesulitan menetapkan prioritas pengeluaran, mengatur kebutuhan jangka panjang, dan menjaga keseimbangan antara pemasukan dan pengeluaran.

Bagi sebagian individu yang telah berusaha mencatat transaksi, permasalahan lain muncul dari metode pencatatan yang digunakan. Pencatatan secara manual, baik di buku tulis maupun lembar kerja digital seperti spreadsheet, masih menimbulkan kendala dalam hal kepraktisan dan efisiensi. Data transaksi yang terus bertambah dari waktu ke waktu membuat catatan menjadi tidak terstruktur dan sulit dibaca. Saat pengguna ingin menelusuri pengeluaran tertentu, mereka harus membuka banyak halaman atau file secara satu per satu. Proses ini memakan waktu dan menyulitkan pengguna dalam menemukan kembali informasi yang dibutuhkan.

Selain kesulitan dalam pencarian, data keuangan yang tersimpan juga sering kali tidak disusun dengan konteks yang jelas. Banyak pengguna hanya memiliki daftar transaksi tanpa hubungan antar kategori atau tanpa penjelasan makna di balik setiap pengeluaran. Kondisi ini membuat informasi keuangan menjadi sekadar kumpulan angka yang tidak mudah dipahami. Akibatnya, pengguna sulit melihat pola kebiasaan pengeluaran, menganalisis pengeluaran, maupun memahami makna dari data yang telah dikumpulkan.

Masalah lain yang muncul adalah rendahnya kemampuan sistem pencatatan dalam menelusuri data keuangan berdasarkan konteks atau makna tertentu. Proses pencarian transaksi masih bergantung pada kesesuaian kata secara literal, bukan pada makna yang terkandung di dalam data. Hal ini menyebabkan pengguna kesulitan dalam menemukan informasi yang relevan apabila terdapat perbedaan istilah, format, atau penulisan pada saat pencatatan dilakukan. Kondisi tersebut semakin memperbesar hambatan dalam proses pelacakan dan analisis keuangan pribadi.

2. Tinjauan Pustaka

2.1. Sistem Informasi

Menurut Kenneth C. Laudon dan Jane P. Laudon (2020:16), "Sebuah sistem informasi dapat didefinisikan secara teknis sebagai seperangkat komponen yang saling berhubungan yang mengumpulkan (atau mengambil kembali), memproses, menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan pengendalian dalam sebuah organisasi".

2.2. Aplikasi

Menurut Afriezal Zein, et al. (2023:4), "Aplikasi atau program aplikasi adalah perangkat lunak yang dirancang untuk melakukan tugas tertentu pada sebuah komputer atau perangkat elektronik lainnya. Aplikasi ini berfungsi sebagai perangkat lunak pendukung yang memungkinkan pengguna untuk melakukan aktivitas tertentu, seperti menulis dokumen, mengedit gambar, bermain game, mengakses internet, dan lain sebagainya".

2.3. Website

Menurut Elgamar (2020:3), "Website merupakan sebuah media yang memiliki banyak halaman yang saling terhubung (*hyperlink*), dimana website memiliki fungsi dalam memberikan informasi berupa teks, gambar, video, suara dan animasi atau penggabungan dari semuanya".

2.4. Aplikasi Web

Menurut Anshul Verma dan Pradeepika Verma (2025:175), "Aplikasi web adalah jenis perangkat lunak yang sering diakses menggunakan peramban web. Aplikasi web biasanya menggunakan kombinasi skrip yang berjalan di server dan di komputer pengguna".

2.6. Semantic Retrieval

Menurut Huanran Zheng, et al. (2023:62), "*Semantic retrieval* adalah teknologi temu kembali berbasis vektor yang mengembalikan hasil paling relevan berdasarkan maksud dan konteks dari kueri pencarian, bukan hanya pencocokan kata kunci. Keunggulan semantic retrieval adalah kemampuannya untuk meningkatkan akurasi dan relevansi hasil pencarian, serta menangani beberapa kueri bahasa alami yang kompleks, seperti parafrase, koreksi kesalahan, transformasi, dan generalisasi".

2.7. Embeddings

Menurut Vicki Boykis (2024:5), "*Embeddings* secara umum merupakan data yang telah diubah menjadi matriks berdimensi-n yang digunakan dalam perhitungan *deep learning* atau pembelajaran mendalam".

2.8. Vector Database

Menurut Abhinav Kimothi (2025:99), "Basis data vektor (*vector database*) dibangun untuk menangani vektor berdimensi tinggi. Basis data ini memiliki spesialisasi dalam pengindeksan dan penyimpanan *vector embeddings* untuk memungkinkan proses pencarian dan pengambilan informasi semantik yang cepat".

2.9. Large Language Models (LLMs)

Menurut Oswlad Campesato (2024:122), "Model bahasa besar (*Large Language Models/LLMs*) adalah subset dari model fondasi yang secara khusus berfokus pada pemrosesan dan pembuatan bahasa manusia. Model-model ini biasanya merupakan jaringan saraf yang sangat besar yang dilatih pada korpus teks yang luas untuk melakukan berbagai tugas yang berhubungan dengan bahasa".

3.0. Retrieval-Augmented Generation (RAG)

Menurut Obie Fernandez (2024:87), "*RAG* adalah sebuah teknik yang menggabungkan kekuatan *LLM* dengan kemampuan pengambilan pengetahuan dari sumber eksternal. Alih-alih hanya bergantung pada pengetahuan pra-pelatihan dari *LLM*, *RAG* secara dinamis mengambil informasi yang relevan dari sumber eksternal dan mengintegrasikannya ke dalam proses generasi".

3. Metode

Penelitian ini akan berfokus pada pembangunan aplikasi web keuangan pribadi dengan fitur *Semantic Retrieval* berbasis *embedding* dan *LLM*. Proses pengembangan dilakukan secara iteratif untuk memastikan efektivitas fitur pencarian semantik dan integrasi kecerdasan buatan dalam sistem.

3.1. Metode Pengembangan Sistem

Metodologi pengembangan sistem yang digunakan dalam penelitian ini adalah *Extreme Programming (XP)*, yang merupakan salah satu pendekatan dalam kerangka kerja *Agile*. *Extreme Programming* dirancang untuk meningkatkan kualitas perangkat lunak serta kemampuan sistem dalam merespons perubahan kebutuhan pengguna secara cepat. Metodologi ini menekankan pengembangan perangkat lunak secara iteratif, umpan balik yang berkelanjutan, serta penyederhanaan desain yang berkembang seiring waktu.

3.2. Teknik Pengumpulan Data

Proses pengumpulan data dalam penelitian ini dilakukan melalui dua pendekatan utama. Peneliti melakukan simulasi data keuangan pribadi untuk menghasilkan kumpulan transaksi yang aman digunakan dalam konteks penelitian. Mengingat bahwa data finansial merupakan informasi sensitif, seluruh catatan transaksi disusun secara mandiri dengan variasi konteks yang beragam, sehingga sistem dapat diuji secara optimal dalam memahami pertanyaan pengguna melalui pencarian semantik.

3.3. Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan melalui beberapa tahapan yang dirancang untuk mendukung proses pencarian semantik berbasis *embedding*. Data transaksi yang diperoleh dari wawancara pengguna terlebih dahulu diproses agar memiliki struktur yang konsisten dan dapat diolah secara komputasional. Selanjutnya, data dikonversi menjadi representasi numerik dalam bentuk vektor *embedding* untuk menangkap makna semantik. Dengan langkah sebagai berikut:

1. Pembersihan Data, dilakukan dengan menghapus karakter tidak relevan, menyeragamkan format teks, serta mengonversi data numerik menjadi bentuk integer agar konsisten dan siap diproses.
2. Konversi ke *Embedding*, teks hasil pembersihan dikonversi menjadi vektor *embedding* menggunakan model *text-embedding-3-small* untuk menangkap hubungan semantik.
3. Penyimpanan *Embedding*, vektor *embedding* disimpan dalam *PostgreSQL* dengan ekstensi *pgvector* untuk mendukung pencarian berbasis *cosine similarity*.
4. Evaluasi Pencarian Semantik, evaluasi dilakukan menggunakan *framework RAGAs* untuk mengukur relevansi hasil pencarian terhadap kueri pengguna.

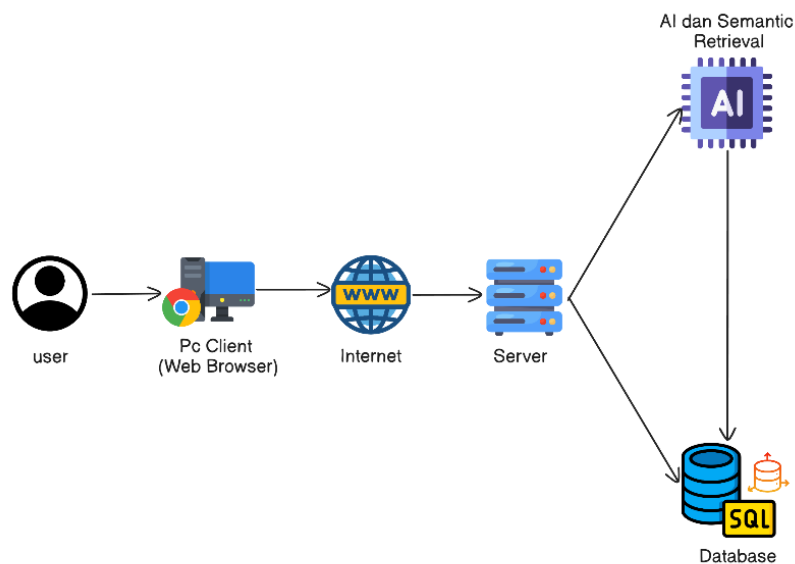
3.4. Teknik Pengukuran Data

Penelitian ini menggunakan *RAGAs (Retrieval-Augmented Generation Assessment)* sebagai kerangka evaluasi utama untuk menilai kinerja sistem *semantic retrieval* dan *Retrieval-Augmented Generation (RAG)*. *RAGAs* merupakan *framework* evaluasi yang dirancang untuk mengukur kualitas pengambilan konteks, kesesuaian jawaban, serta keterkaitan antara kueri pengguna, hasil *retrieval*, dan keluaran model bahasa. Evaluasi dilakukan secara komprehensif menggunakan beberapa metrik utama, yaitu *context precision* untuk menilai relevansi konteks yang diambil, *context recall* untuk mengukur kelengkapan informasi yang berhasil ditemukan, *faithfulness* untuk memastikan jawaban sesuai dengan konteks tanpa halusinasi, serta *answer relevance* untuk menilai kesesuaian jawaban terhadap maksud kueri pengguna.

4. Hasil

4.1. Arsitektur Sistem

Secara umum, sistem dibangun menggunakan arsitektur *client-server*, di mana pengguna berinteraksi melalui peramban web sebagai antarmuka utama untuk mengakses fitur aplikasi, seperti pencatatan transaksi dan pencarian data keuangan. Setiap permintaan pengguna dikirim melalui jaringan internet ke server yang berperan sebagai pusat pengolahan, mencakup logika bisnis, pemrosesan data, serta pengelolaan komunikasi dengan basis data dan modul kecerdasan buatan.

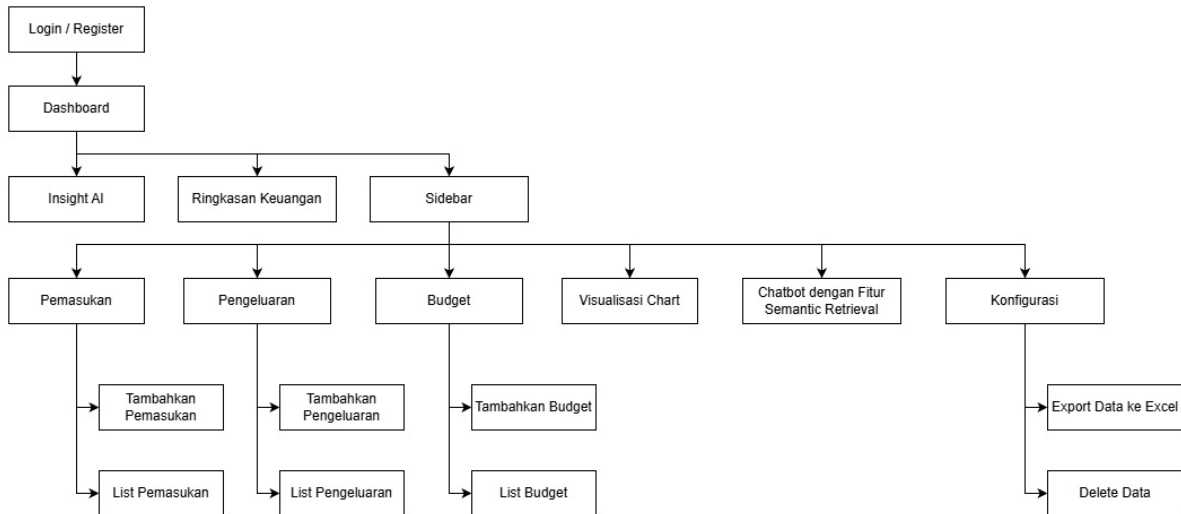


Gambar 1. Arsitektur Sistem

4.2. Struktur Navigasi

Rancangan struktur navigasi aplikasi web keuangan pribadi ini disusun secara hierarkis untuk memastikan kemudahan akses, kejelasan alur, dan konsistensi interaksi pengguna. Alur dimulai dari halaman login dan registrasi sebagai mekanisme autentikasi, kemudian pengguna diarahkan ke dashboard yang berfungsi

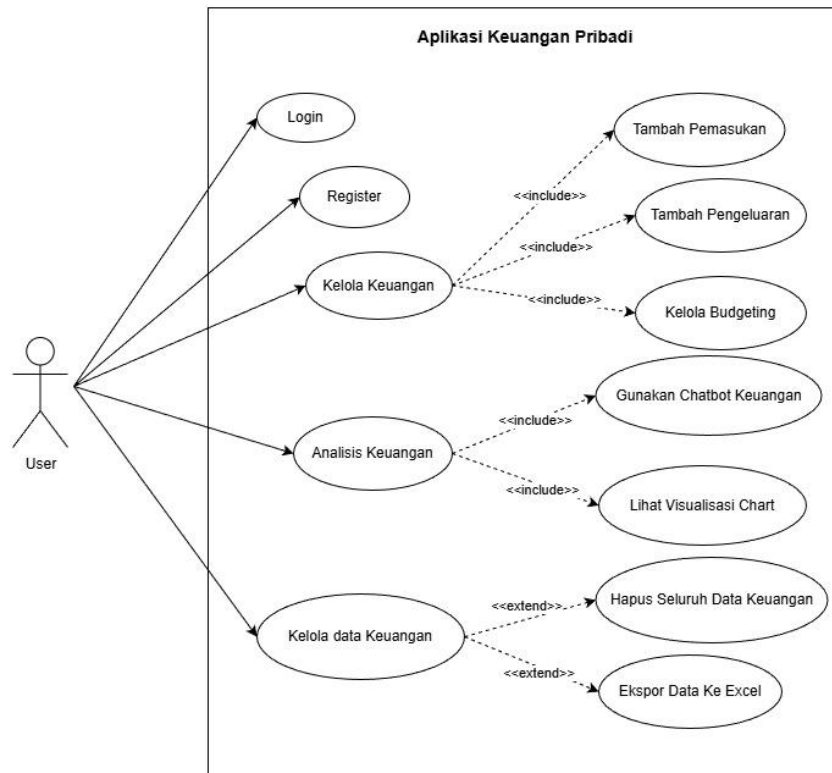
sebagai pusat informasi dan navigasi. Dashboard menyediakan ringkasan keuangan, insight berbasis kecerdasan buatan, serta akses ke berbagai modul melalui sidebar, meliputi pemasukan, pengeluaran, budget, visualisasi chart, chatbot dengan fitur *semantic retrieval*, dan konfigurasi. Setiap modul dirancang untuk mendukung pengelolaan data keuangan secara terstruktur, mulai dari pencatatan transaksi, perencanaan anggaran, hingga penyajian data secara visual dan pencarian berbasis bahasa alami, sehingga pengguna dapat memperoleh informasi secara lebih efektif dan kontekstual.



Gambar 2. Struktur Navigasi

4.3. Use Case Diagram

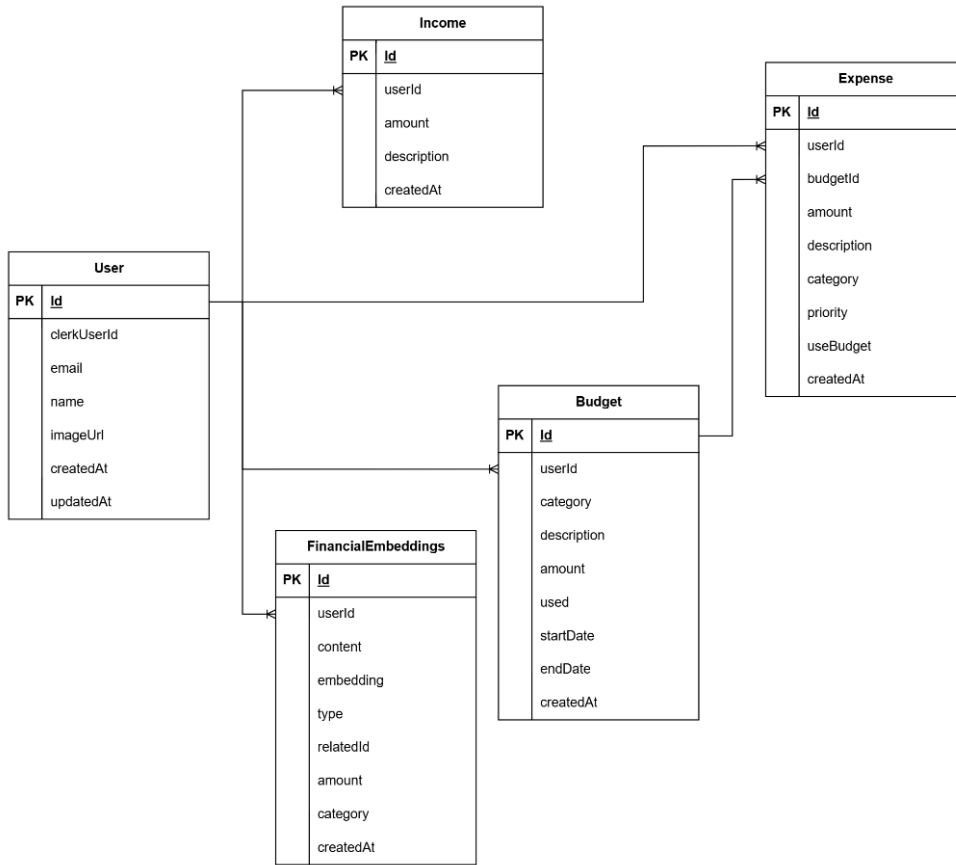
Use case diagram pada aplikasi web keuangan pribadi ini menggambarkan interaksi fungsional antara pengguna sebagai aktor utama dengan sistem. Proses diawali dengan autentikasi melalui login dan registrasi sebagai prasyarat akses. Setelah itu, pengguna dapat melakukan pengelolaan keuangan yang mencakup pencatatan pemasukan, pengeluaran, dan pengelolaan anggaran yang saling terintegrasi (*include*). Selain itu, sistem menyediakan fitur analisis keuangan yang memanfaatkan chatbot berbasis *semantic retrieval* dan visualisasi chart untuk mendukung pemahaman data. Sistem juga menyediakan pengelolaan data keuangan secara menyeluruh dengan fitur opsional (*extend*), seperti penghapusan data dan ekspor ke format Excel, sehingga mendukung kebutuhan pengelolaan dan pelaporan data secara fleksibel.



Gambar 3. Use Case Diagram

4.3. Rancangan Database

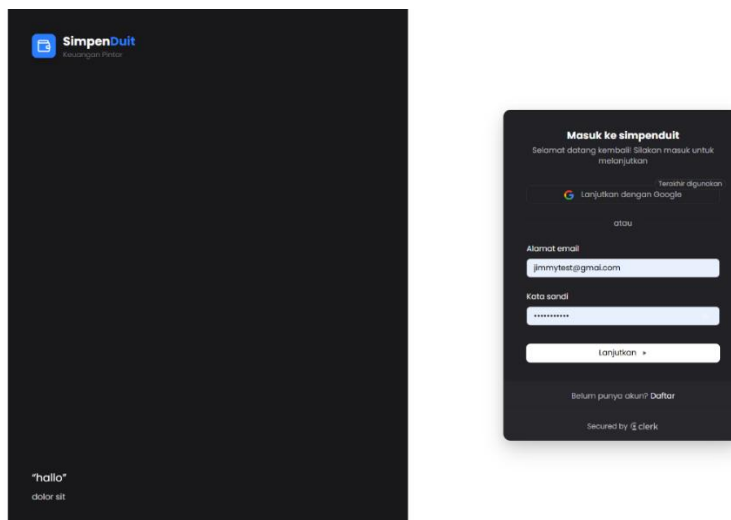
Rancangan basis data pada aplikasi web keuangan pribadi ini menggunakan model relasional yang terintegrasi untuk mengelola data pengguna, transaksi keuangan, anggaran, serta *embedding* berbasis *semantic retrieval*. Tabel *User* berperan sebagai entitas utama yang memiliki relasi satu ke banyak dengan tabel *Income*, *Expense*, *Budget*, dan *FinancialEmbedding*. Tabel *Income* dan *Expense* masing-masing digunakan untuk menyimpan data pemasukan dan pengeluaran, di mana *Expense* dilengkapi atribut kategori dan prioritas serta memiliki relasi opsional dengan *Budget*. Tabel *Budget* berfungsi untuk menyimpan perencanaan anggaran berdasarkan kategori dan periode tertentu, yang dapat digunakan oleh beberapa data pengeluaran. Selain itu, tabel *FinancialEmbedding* menyimpan representasi vektor dari data keuangan untuk mendukung proses pencarian berbasis makna secara efisien.



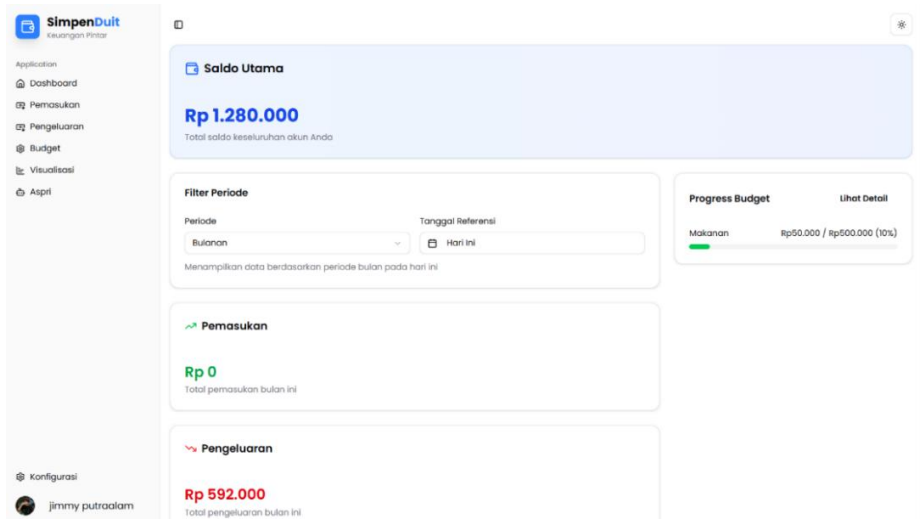
Gambar 4. Rancangan Database

4.4. Implementasi Sistem

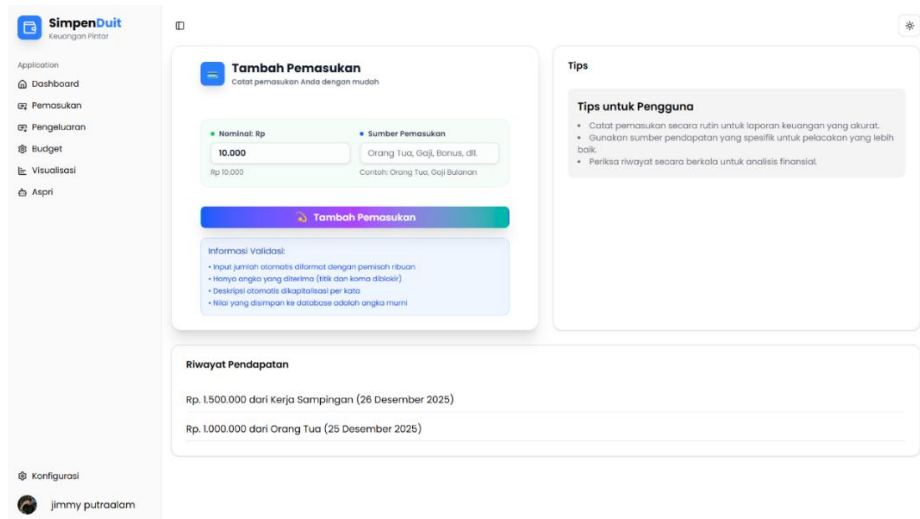
Berikut adalah hasil implementasi sistem berupa aplikasi web yang disajikan melalui gambar.



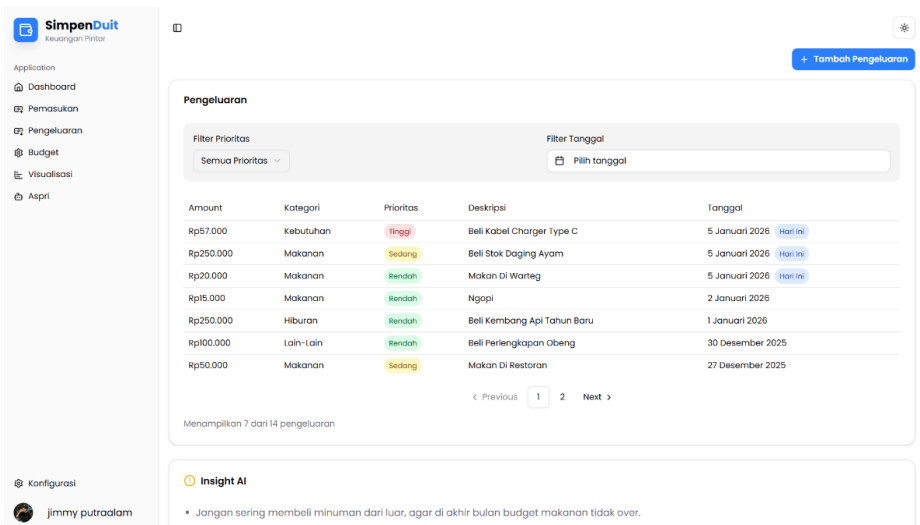
Gambar 5. Tampilan Halaman Login



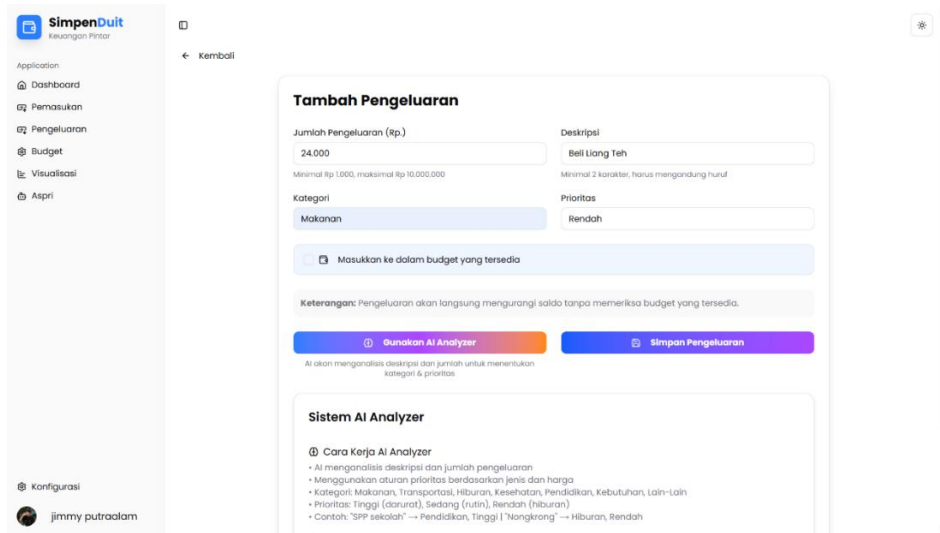
Gambar 6. Tampilan Halaman Dashboard



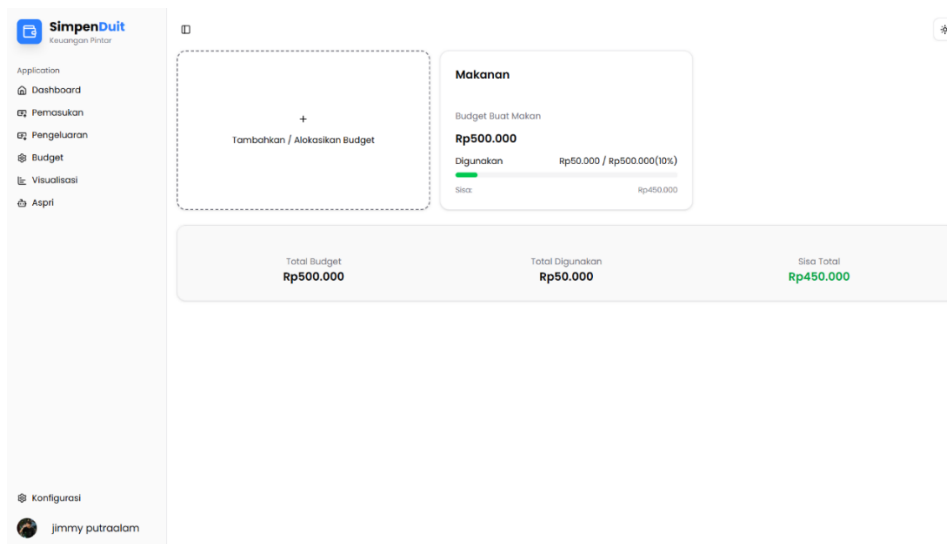
Gambar 7. Tampilan Halaman Pemasukan



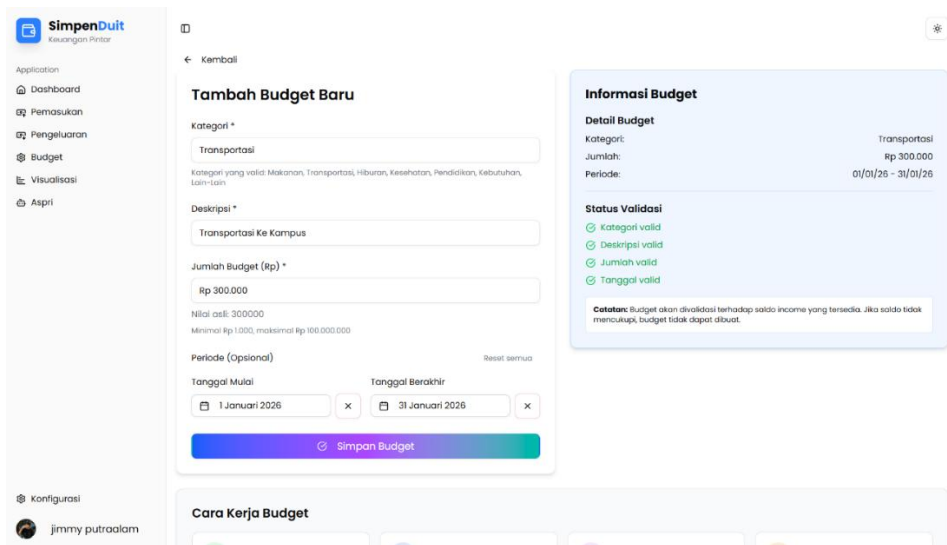
Gambar 8. Tampilan Halaman Pengeluaran



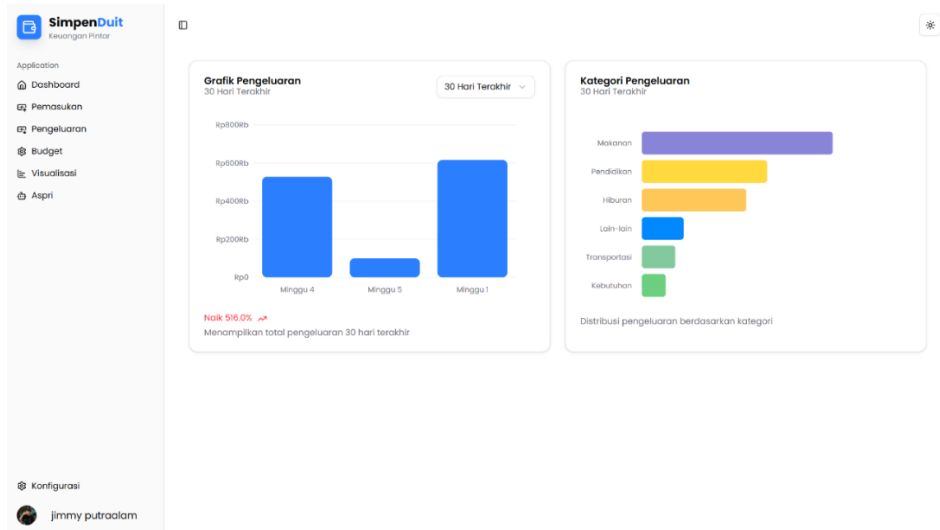
Gambar 9. Tampilan Halaman Tambah Pengeluaran



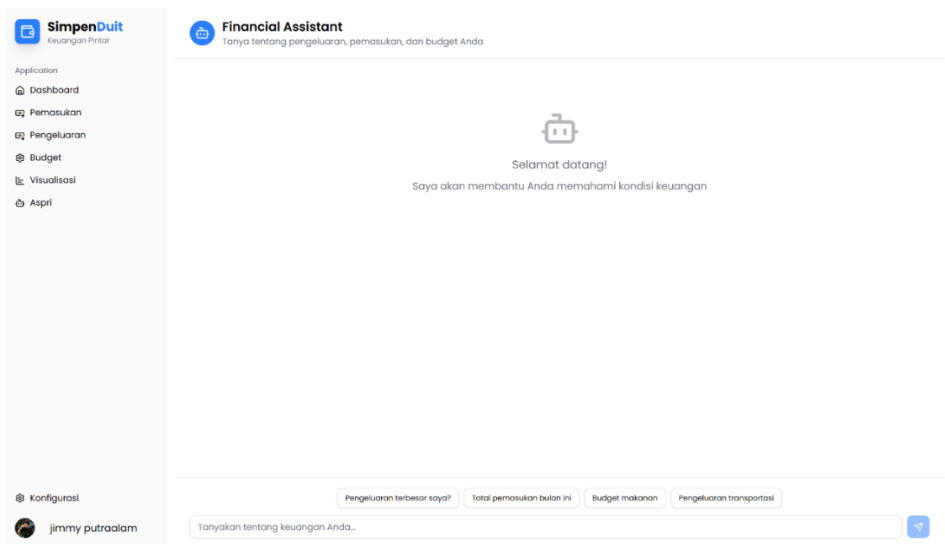
Gambar 10. Tampilan Halaman Budget



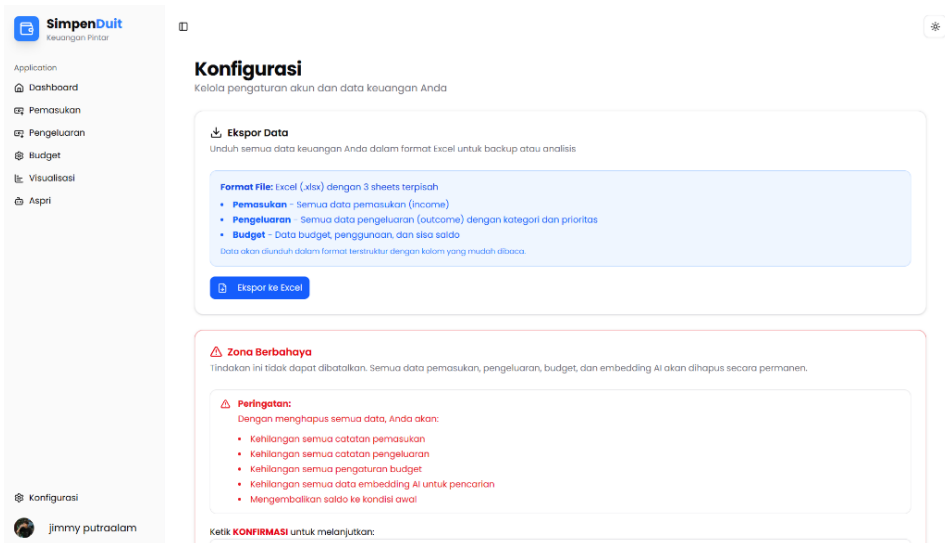
Gambar 11. Tampilan Halaman Tambah Budget



Gambar 12. Tampilan Halaman Visualisasi

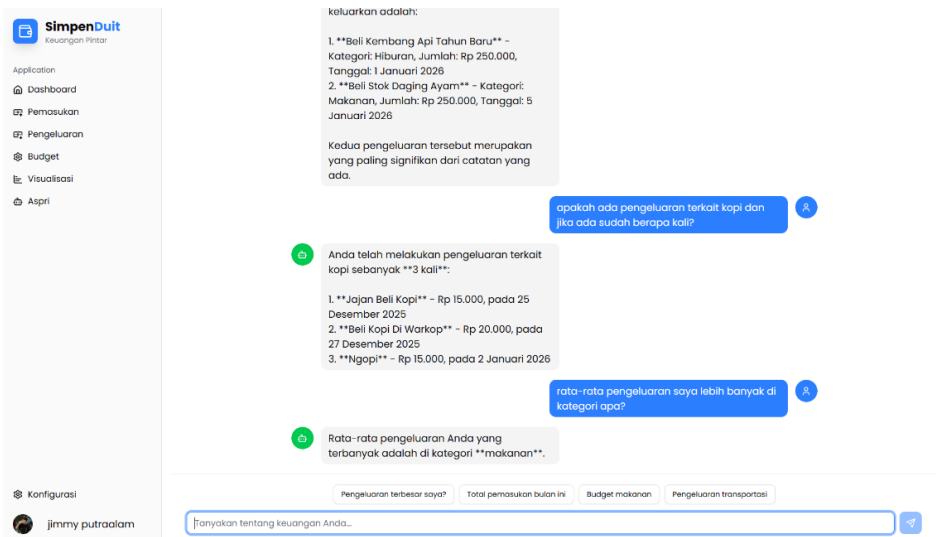


Gambar 13. Tampilan Halaman Chatbot



Gambar 14. Tampilan Halaman Konfigurasi

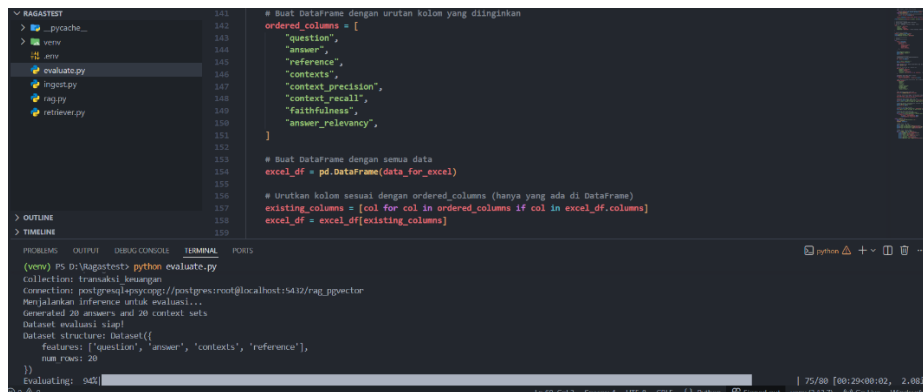
Implementasi modul pencarian berbasis *semantic retrieval* dilakukan untuk mengatasi keterbatasan metode pencarian berbasis kata kunci dengan memungkinkan pengguna menggunakan bahasa alami yang lebih fleksibel. Berikut gambar hasil implementasi:



Gambar 15. Tampilan Implementasi Semantic Retrieval

4.6. Pengujian Sistem Menggunakan RAGAs

Pengujian sistem dilakukan menggunakan 100 data transaksi keuangan yang telah disimpan dalam basis data *PostgreSQL* dengan ekstensi *pgvector* dan direpresentasikan dalam bentuk *embedding* vektor berdimensi 1536. Evaluasi dilakukan menggunakan 20 pertanyaan yang disusun berdasarkan skenario penggunaan, seperti pencarian berdasarkan jenis, kategori, tanggal, dan deskripsi transaksi. Metrik evaluasi yang digunakan dalam pengujian ini meliputi *context precision* untuk mengukur ketepatan konteks yang diambil, *context recall* untuk menilai kelengkapan konteks relevan, *faithfulness* untuk memastikan kesesuaian jawaban terhadap konteks, serta *answer relevancy* untuk mengukur relevansi jawaban terhadap pertanyaan. Nilai evaluasi diperoleh dari rata-rata skor seluruh pertanyaan yang diuji, sehingga memberikan gambaran menyeluruh mengenai kinerja sistem.



4.6. Hasil Pengujian RAGAs

Berikut adalah hasil pengujian sistem menggunakan RAGAs yang akan disajikan dalam bentuk tabel.

Tabel 1. Hasil Pengujian RAGAs

Metrik Evaluasi	Nilai Rata-rata
Context Precision	0,70
Context Recall	0,81
Faithfulness	0,83
Answer Relevancy	0,84

5. Pembahasan

Hasil menunjukkan bahwa aplikasi web keuangan pribadi yang dikembangkan telah berhasil diimplementasikan dengan baik, khususnya dalam mendukung fitur pencarian transaksi berbasis *semantic retrieval*. Sistem mampu mengelola data keuangan secara terstruktur sekaligus menyediakan mekanisme pencarian yang lebih fleksibel dan kontekstual. Keberhasilan ini terlihat dari kemampuan sistem dalam memahami maksud pengguna melalui bahasa alami serta menyajikan informasi yang relevan sesuai kebutuhan.

Secara kuantitatif, nilai *context recall* sebesar 0,81 menunjukkan bahwa sistem mampu mengambil sebagian besar konteks transaksi yang relevan, sehingga meminimalkan kehilangan informasi penting dalam proses pencarian. Sementara itu, nilai *context precision* sebesar 0,70 mengindikasikan bahwa meskipun masih terdapat sebagian konteks yang kurang relevan, sistem tetap mampu menjaga keseimbangan antara kelengkapan dan ketepatan informasi. Nilai *faithfulness* sebesar 0,83 menunjukkan bahwa jawaban yang dihasilkan sistem konsisten dengan konteks yang digunakan dan tidak menyimpang dari data yang tersedia. Selain itu, nilai *answer relevancy* sebesar 0,84 menegaskan bahwa jawaban yang diberikan memiliki tingkat kesesuaian yang tinggi terhadap pertanyaan pengguna.

Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa modul pencarian berbasis *semantic retrieval* yang diimplementasikan mampu meningkatkan kualitas pencarian transaksi keuangan secara signifikan. Sistem tidak hanya mampu menemukan data yang relevan berdasarkan makna, tetapi juga menghasilkan jawaban yang akurat dan kontekstual, sehingga dapat menjawab permasalahan penelitian serta meningkatkan efektivitas dalam pengelolaan dan pencarian informasi keuangan pribadi.

6. Kesimpulan

Berdasarkan hasil perancangan, implementasi, dan pengujian, dapat disimpulkan bahwa sistem web keuangan pribadi yang dikembangkan telah berhasil memenuhi tujuan penelitian. Sistem ini mampu menyediakan pencatatan transaksi yang terstruktur dan terpusat, sehingga memudahkan pengguna dalam memantau arus

kas serta memahami kondisi keuangan secara lebih objektif. Implementasi berbasis web juga memungkinkan pengelolaan data yang lebih rapi, efisien, dan minim kesalahan dibandingkan metode konvensional.

Selain itu, penerapan modul *semantic retrieval* terbukti meningkatkan kualitas pencarian data keuangan. Hasil evaluasi menunjukkan nilai *context recall* 0,81, *context precision* 0,70, *faithfulness* 0,83, dan *answer relevancy* 0,84, yang mengindikasikan bahwa sistem mampu mengambil konteks relevan dan menghasilkan jawaban yang sesuai dengan kebutuhan pengguna. Penggunaan *embedding* vektor juga memungkinkan data transaksi dipahami berdasarkan kesamaan makna, sehingga memberikan analisis yang lebih kontekstual. Secara keseluruhan, sistem ini mampu memberikan solusi yang lebih efektif dalam pencatatan dan pencarian data keuangan dibandingkan pendekatan konvensional.

Di sisi lain, penggunaan representasi *embedding* dalam bentuk vektor memungkinkan sistem untuk mengelompokkan dan menghubungkan data transaksi berdasarkan kesamaan makna. Pendekatan ini memberikan nilai tambah dalam analisis data keuangan, karena pengguna dapat memahami hubungan antar transaksi serta pola pemasukan dan pengeluaran secara lebih menyeluruh.

Secara keseluruhan, sistem web keuangan pribadi yang dikembangkan, khususnya dengan dukungan modul pencarian berbasis *semantic retrieval*, telah terbukti mampu memberikan solusi yang lebih efektif dalam pencatatan, pengelolaan, dan pencarian data keuangan. Sistem ini tidak hanya meningkatkan efisiensi dan akurasi pengelolaan data, tetapi juga memberikan pengalaman pengguna yang lebih baik melalui pencarian yang fleksibel dan berbasis makna, sehingga mampu melampaui keterbatasan metode konvensional.

Daftar Pustaka

- Abhinav Kimothi. (2025). *A Simple Guide to Retrieval Augmented Generation*. Manning Publications Co.
- Afriezal Zein. et al. (2023). *Konsep Dasar Rekayasa Perangkat Lunak*. Yayasan Cendikia Mulia Mandiri.
- Anshul Verma dan Pradeepika Verma. (2025). *Research Advances in Network Technologies: Volume 2, Edisi Ke-1*. CRC Press.
- Elgamar. (2020). *Konsep Dasar Pemrograman Web Dengan PHP*. Malang: CV. Multimedia Edukasi.
- Huanran Zheng. et al. (2023). "ECNU-LLM@CHIP-PromptCBLUE: Prompt Optimization and In-Context Learning for Chinese Medical Tasks", Springer. https://doi.org/10.1007/978-981-97-1717-0_5
- Kenneth C. Laudon. dan Jane P. Laudon. (2020). *Management information systems: managing the digital firm*. Edisi Ke-16, New York: Pearson.
- Obie Fernandez. (2024), *Patterns of Application Development Using AI*. Leanpub.
- Oswlad Campesato. (2024). *Large Language Models an Introduction*. Boston: Mercury Learning and Information.

Vicki Boykis. (2024). What are embeddings. Creative Commons.

Umar, R., Hadi, A., Widiandana, P., & Anwar, F. (2019). Perancangan Database Point of Sales Apotek Dengan Menerapkan Model Data Relasional. *Query: Journal of Information Systems*, 33–41.

Zahra, R. A. (2016). *Penerapan Algoritma Random Forest Dengan Hyperparameter Tuning Untuk Memprediksi Harga Sewa Kost Di Kota Bandar Lampung*. Skripsi.