

BAB 3 METODOLOGI PENELITIAN

3.1 Alur Penelitian

Pada bagian ini menjelaskan alur dalam mengerjakan penelitian. Alur yang diberikan merupakan jabaran dari awal hingga akhir dari pengerjaan penelitian. Dimulai dari pengumpulan data, perancangan, implementasi, penilaian, dan dokumentasi. Berikut adalah penjelasan dari alur penelitian.

1. Studi Literatur

Pada tahap ini dilakukan pengumpulan literatur yang berhubungan dengan penelitian yang digunakan. Literatur mengenai *large language model*, *vector database*, *cosine similarity*, teknik *retrieval augmented generation* dan penelitian tentang penggunaan *chatbot* sebelumnya.

2. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data yang berupa Undang - Undang dan Peraturan Pemerintah yang digunakan sebagai *knowledge base*. Data dikumpulkan dari *website* peraturan.bpk.go.id.

3. Perancangan

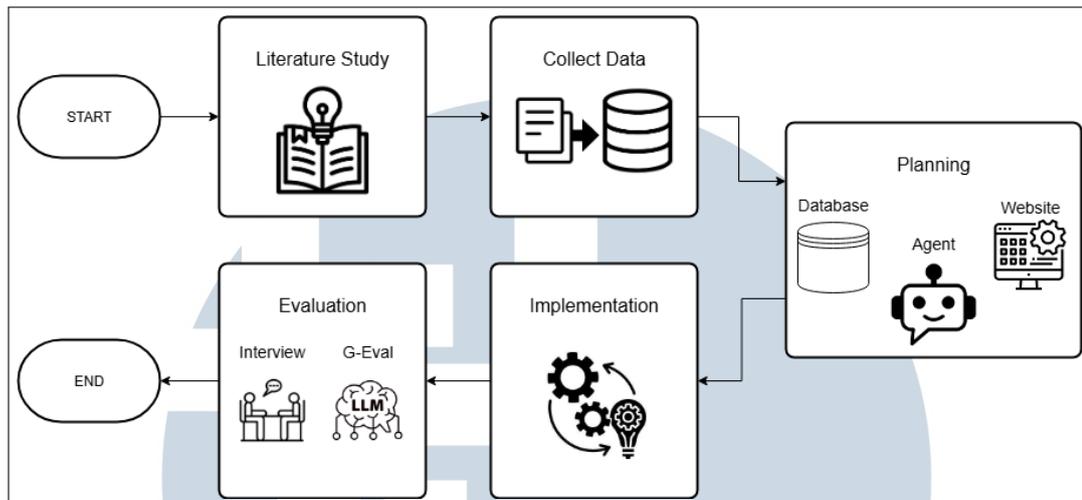
Pada tahap ini dilakukan proses perancangan dari sisi *chatbot* dan tampilan aplikasi.

4. Implementasi

Pada tahap ini adalah proses pembuatan aplikasi yang sudah dirancang pada tahap perancangan.

5. Pengujian

Pada tahap ini dilakukan pengujian *chatbot* dengan beberapa pertanyaan serta akurasi dari *chatbot*. Pengujian dilakukan dengan wawancara konsultan pajak dan mengukur akurasi dengan metrics G-Eval.



Gambar 3.1. Flowchart Alur Penelitian

3.2 Pengumpulan Data

Data yang digunakan adalah dokumen PDF yang berisikan undang-undang dan peraturan pemerintah tentang pajak. Data didapatkan melalui *website* BPK RI, yaitu peraturan.bpk.go.id. Data Dokumen yang dipilih memiliki informasi yang relevan dengan topik penelitian. Terdapat tiga Undang-Undang dan tiga Peraturan Pemerintah yang digunakan pada penelitian ini.

A Undang-Undang

Pertama adalah UU Nomor 7 Tahun 2021 yang mengatur ketentuan umum dan tata cara perpajakan, kuasa wajib pajak dan penentuan pidana pajak [16]. Kedua yaitu UU Nomor 11 Tahun 2020 yang mengatur cipta kerja untuk memudahkan berusaha, berinvestasi dan menciptakan lapangan pekerjaan [17]. Terakhir UU Nomor 36 Tahun 2008 yang mengatur berbagai aspek terkait subjek, objek, tarif, dan ketentuan perpajakan penghasilan di Indonesia [18]. Berikut adalah daftar Undang-Undang yang digunakan dalam bentuk daftar.

1. UU Nomor 7 Tahun 2021 [16]
2. UU Nomor 11 Tahun 2020 [17]
3. UU Nomor 36 Tahun 2008 [18]

B Peraturan Pemerintah

Pertama adalah PMK Nomor 243 Tahun 2014 yang mengatur tentang Surat Pemberitahuan (SPT) yang mewajibkan melaporkan perhitungan dan pembayaran pajak [19]. Kedua yaitu PP Nomor 50 Tahun 2022 yang mengatur tata cara pelaksanaan hak dan pemenuhan perpajakan bagi wajib pajak termasuk Nomor Pokok Wajib Pajak (NPWP), pelaporan Surat Pemberitahuan (SPT) dan tata cara membayar pajak [20]. Terakhir yaitu PP Nomor 58 Tahun 2023 yang mengatur tarif pemotongan Pajak Penghasilan (PPH) atas penghasilan pekerjaan, jasa atau kegiatan yang diterima oleh wajib pajak pribadi [21].

1. PMK Nomor 243 Tahun 2014 [19]
2. PP Nomor 50 Tahun 2022 [20]
3. PP Nomor 58 Tahun 2023 [21]

Dokumen yang digunakan akan dipecah menjadi per halaman untuk disimpan ke dalam *vector database*. Penyimpanan dalam bentuk yang lebih kecil untuk mempermudah dalam pencarian data. Selain itu, pembagian dokumen menjadi bagian yang lebih kecil mempermudah *Large Language Model (LLM)* menggunakan dokumen dalam menjawab pertanyaan.

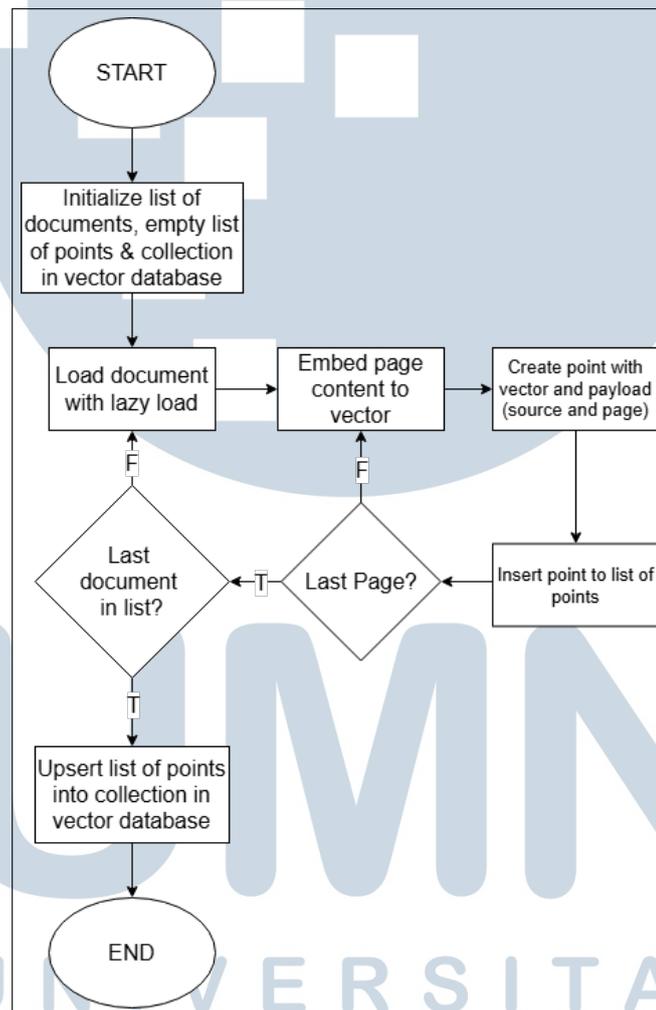
3.3 Perancangan

Pada bagian perancangan akan menjelaskan arsitektur sistem yang akan dikembangkan pada penelitian ini. Mulai dari *vector database*, model *Retrieval Augmented Generation (RAG)*, *chatbot* dan aplikasi. Desain sistem dirancang untuk memenuhi kebutuhan pada penelitian sebelumnya. Tujuan menggunakan *Vector database* dibandingkan dengan *database* biasa karena keperluan untuk melakukan *similarity* atau *semantic search*.

3.3.1 Rancangan Vector Database

Pada tahap ini, akan membahas proses persiapan *vector database* data yang digunakan berupa PDF dengan isi teks. PDF akan dipecah menjadi per halaman menggunakan *lazy load* dari langchain community. Teks dari PDF tidak melalui proses *filter* untuk mempermudah pengguna jika mau menambahkan PDF baru.

Teks per halaman akan disimpan pada *vector database* dalam bentuk *PointStruct* dengan isi random id, *payload* dan *vector* data. *Payload* berisikan teks asli dan metadata yang ada informasi *source* atau judul PDF serta halaman berapa dari teks tersebut. *Vector* akan diisi teks pdf yang sudah diubah bentuknya ke vektor melalui proses *embedding*. *Embedding* akan menggunakan model *intfloat/multilingual-e5-large* dengan dimensi 1024 dan jarak *Cosine Similarity*. Untuk *payload* akan diisi *source* dan *page* untuk mengetahui sumber dan halaman.

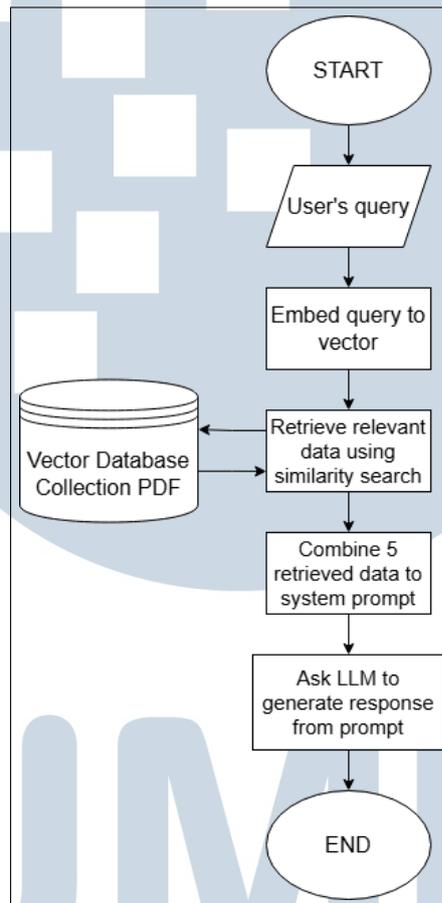


Gambar 3.2. Flowchart rancangan *vector database*

3.3.2 Rancangan RAG

Tahap awal pada perancangan teknik *Retrieval Augmented Generation (RAG)* pada chatbot dengan membuat arsitektur dahulu. Terdapat tiga proses utama pada teknik RAG, yaitu Pengambilan (*Retrieval*), Augmentasi (*Augmented*) dan

Generasi (*Generation*). Proses pengambilan adalah proses mengumpulkan data yang relevan dengan input di *vector database*. Pada proses augmentasi adalah pengabungan data yang sudah diambil dengan instruksi sistem. Terakhir pada generasi adalah proses chatbot menghasilkan jawaban dari input berdasarkan data yang didapat.

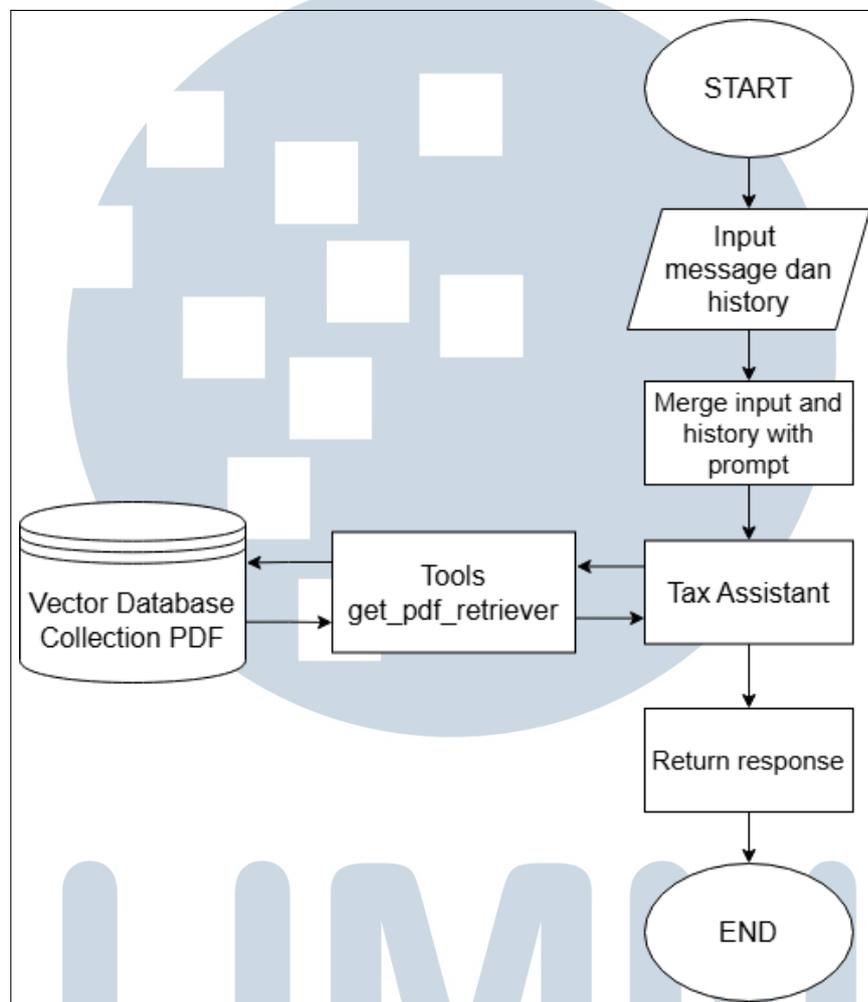


Gambar 3.3. Flowchart rancangan RAG

3.3.3 Rancangan Chatbot

Pada tahap ini dilakukan perancangan pada chatbot yang akan dibuat. Dari kebutuhan penelitian, chatbot hanya diperlukan untuk menjawab pertanyaan dengan bantuan mengambil data yang relevan. Jadi chatbot akan menggunakan 1 *agent* yang dilengkapi dengan *tools* untuk mengambil data. *Agent* akan dilengkapi dengan *LLM* model dari Gemini yaitu Gemini 2.0 Flash. *LLM* menggunakan API dari Gemini untuk menggunakan *LLM*. *Agent* adalah AI yang dapat dilengkapi *prompt* atau instruksi sistem sehingga dapat mengerjakan tugas yang kompleks serta

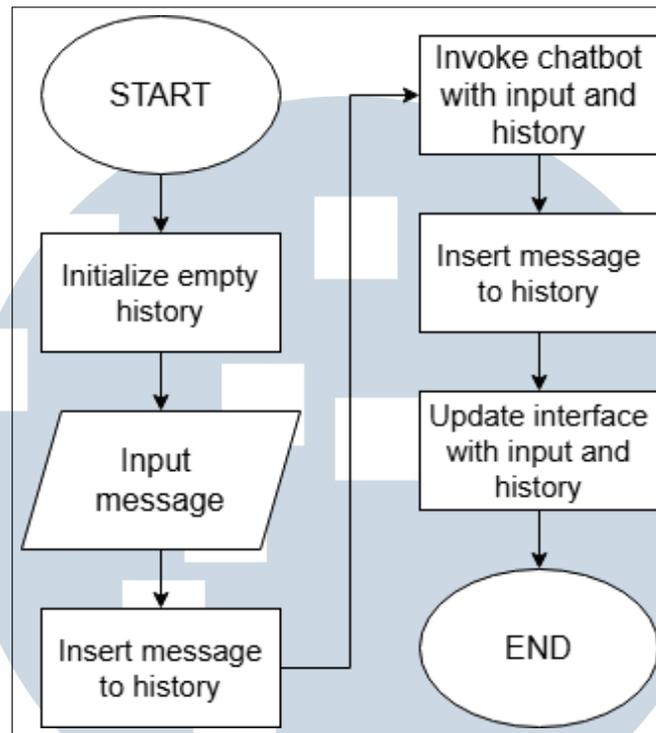
membuat keputusan. *Tools* adalah *function* yang dilengkapi dengan deskripsi yang dapat digunakan oleh *agent* sesuai kebutuhan agent tersebut.



Gambar 3.4. Flowchart rancangan chatbot

3.3.4 Rancangan Aplikasi

Rancangan aplikasi adalah rancangan tampilan yang digunakan untuk mengakses chatbot. Tampilan akan menggunakan *open-source* Streamlit untuk membantu *user* dalam mengakses chatbot. Tampilan dirancang agar dapat memberi input ke chatbot dan melihat *history* sebelumnya.



Gambar 3.5. Flowchart rancangan aplikasi

3.4 Evaluasi

Evaluasi diperlukan untuk menguji tingkat akurasi terhadap chatbot yang telah dibuat. Penilaian dilakukan dengan empat cara yaitu, wawancara, perbandingan, evaluasi G-Eval dan pengukuran kemiripan. Penilaian pertama yaitu dengan melakukan wawancara konsultan pajak terhadap jawaban chatbot dari daftar pertanyaan yang sudah disiapkan. Proses ini menguji apakah respon dari *Large Language Model (LLM)* dengan teknik *Retrieval Augmented Generation (RAG)* masih relevan menurut konsultan pajak tersebut. Wawancara dengan konsultan pajak diharapkan dapat menentukan akurasi dari jawaban chatbot.

Evaluasi kedua yaitu dengan menghitung *score* G-Eval menggunakan *framework DeepEval*. G-Eval adalah teknik evaluasi berbasis *LLM*. Evaluasi yang diuji adalah pertanyaan dan jawaban *chatbot* pada Tabel A1.

Evaluasi hanya menguji hasil dari *chatbot* dan tidak menguji hasil dari *LLM* karena bias. Sebelumnya telah dilakukan pengujian G-Eval pada *LLM* dengan mendapatkan *score* 100%. *LLM* cenderung memberikan skor yang lebih tinggi terhadap teks yang dihasilkan oleh *LLM* yang membuat ilusi skor yang lebih tinggi. Selain itu, *LLM* yang menghasilkan jawaban dan menguji jawaban menggunakan

model yang sama yaitu Gemini 2.0 Flash yang membuat hasilnya menjadi bias

Evaluasi ketiga yaitu dengan membandingkan hasil *chatbot* dengan LLM tanpa teknik *Retrieval-Augmented Generation*. LLM dan *chatbot* akan memberikan output dari input yang akan diberikan. Hasil dari LLM dan *chatbot* akan dibandingkan satu sama lain dengan dilihat dari akurasi dan kelengkapan jawaban.

Evaluasi keempat yaitu pengukuran pada pencarian data *vector database* dilakukan dengan analisa *cosine similarity* antara *query* dan hasil pencarian. *Score* yang didapat dari Qdrant juga diuji dengan rumus *cosine similarity*. Penilaian membandingkan nilai dan evaluasi sendiri konteks antara *query* dan hasil pencarian.

3.5 Dokumentasi

Tahap terakhir adalah dokumentasi yaitu penulisan laporan. Laporan berisikan rangkaian proses dalam mengerjakan penelitian dari latar belakang, landasan teori, perancangan, implementasi dan penilaian aplikasi. Semua informasi dirangkum sebaik mungkin agar mudah dipahami pembaca.

