

3. ANALISIS DAN DESAIN SISTEM

3.1 Analisa Permasalahan

Ada beberapa masalah yang dihadapi oleh universitas X dalam menjawab pertanyaan dari calon mahasiswa baru. Saat ini universitas X masih mengandalkan petugas administrasi untuk menjawab pertanyaan yang diajukan oleh calon mahasiswa secara manual. Hal ini mengakibatkan pemborosan waktu dan sumber daya manusia yang seharusnya dapat digunakan untuk kegiatan lain yang lebih produktif. Keterbatasan petugas administrasi dapat menghambat responsivitas dalam memberikan informasi kepada calon mahasiswa yang membutuhkan jawaban secara cepat dan akurat. Akibatnya, ada beberapa pertanyaan dari calon mahasiswa yang belum terjawab oleh petugas administrasi dengan baik.

Petugas administrasi Universitas X juga terbatas oleh waktu buka kampus, sehingga hanya dapat memberikan jawaban terhadap pertanyaan calon mahasiswa selama jam tersebut. Namun, pertanyaan dari calon mahasiswa dapat datang kapan saja, tanpa memperdulikan waktu buka kampus. Akibatnya, calon mahasiswa harus menunggu waktu buka kampus untuk mendapatkan jawaban dari petugas administrasi. Namun, karena harus menunggu, pertanyaan-pertanyaan tersebut masuk dalam antrian yang dapat memperpanjang waktu tunggu untuk mendapatkan jawaban. Kondisi ini berpotensi menimbulkan kehilangan minat dari calon mahasiswa, karena mereka mungkin telah menemukan informasi serupa di universitas lain yang menyediakan layanan informasi yang lebih cepat dan fleksibel. Dalam hal ini, kecepatan dan ketersediaan informasi menjadi faktor kunci dalam mempengaruhi keputusan calon mahasiswa untuk mendaftar di Universitas X.

3.2 Analisa Kebutuhan

Dari masalah yang telah dijelaskan pada bagian analisis permasalahan, maka Universitas X memerlukan implementasi chatbot yang dapat memberikan respon yang cepat dan efisien terhadap pertanyaan calon mahasiswa. *Chatbot* ini akan memastikan bahwa semua pertanyaan yang masuk dapat dijawab dengan baik tanpa harus menunggu waktu atau ketersediaan sumber daya manusia.

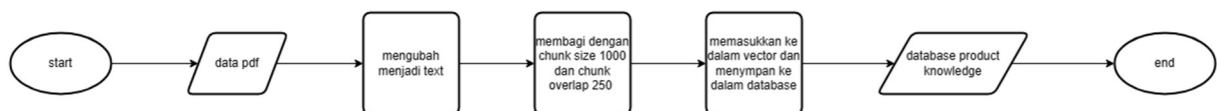
Dengan demikian, universitas X dapat meningkatkan layanan informasi kepada calon mahasiswa dan memastikan bahwa setiap pertanyaan mendapat tanggapan yang memuaskan.

Selain itu, kehadiran *chatbot* juga memungkinkan calon mahasiswa untuk mendapatkan jawaban atas pertanyaan mereka tanpa terikat oleh waktu buka kampus. Dengan adanya layanan informasi yang dapat diakses kapan pun, calon mahasiswa memiliki fleksibilitas dalam mencari informasi yang mereka butuhkan, baik itu pada jam operasional maupun di luar jam operasional. Ini akan meningkatkan kepuasan dan pengalaman calon mahasiswa dalam proses pendaftaran di universitas, karena dapat memperoleh informasi yang mereka perlukan dengan lebih mudah dan nyaman.

Dalam pembuatan *chatbot* dibutuhkan data yang berisi *product knowledge* mengenai universitas X berupa prodi yang tersedia, keunggulan, biaya, persyaratan, langkah mendaftar. Data *product knowledge* akan digunakan sebagai pengetahuan untuk menjawab pertanyaan calon mahasiswa atau orang tua calon mahasiswa dalam *Prompt engineering*. Data *conversation history* juga dibutuhkan untuk melakukan *training* pada *fine tune model* yang akan dibuat. Data *conversation history* berisi percakapan sebelumnya yang dilakukan oleh calon mahasiswa atau orang tua calon mahasiswa dengan petugas administrasi dari universitas X.

3.3 Analisis Data

Data yang akan digunakan untuk skripsi ini berupa pdf yang berisi informasi mengenai program studi yang tersedia di universitas X dan data dari *conversation history*. Setiap pdf berisi informasi singkat, *career prospect*, kurikulum, dan *tuition fee* untuk tiap program studi yang tersedia. Data pdf ini digunakan sebagai *product knowledge* yang akan menjadi *context* untuk *Prompt engineering*. Sebelum digunakan sebagai *context* data akan di *preprocessing* terlebih dahulu.



Gambar 3.1 Flowchart Data Preprocessing

Flowchart pada Gambar 3.1 menggambarkan proses *Data preprocessing* yang dilakukan untuk data pdf. Pertama data akan diubah terlebih dahulu menjadi text secara manual. Setelah itu

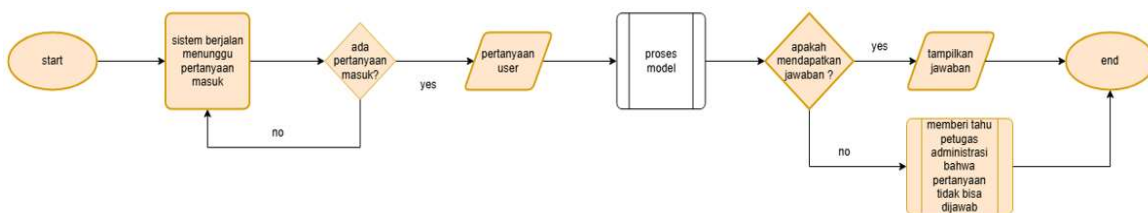
data yang sudah berupa *text* akan di *load* lalu di *split* dengan *chunk size* 1000 dan *chunk overlap* 250. Data yang telah di *split* akan dimasukkan ke dalam *vector* dan disimpan ke dalam *database chroma database*).

Untuk data dari *conversation history* berisi pertanyaan dari calon mahasiswa dan jawaban dari petugas administrasi universitas X. Data diambil dari Mekari Qontak, CRM yang digunakan oleh Universitas X. Mekari Qontak memiliki beberapa *channel* sosial media yang bisa digunakan oleh *customer*. Platform yang digunakan oleh universitas X yaitu *whatsapp*, *instagram*, dan *email*. Data yang didapat berupa excel terdiri dari kumpulan percakapan yang diambil dalam rentang waktu januari 2023 hingga desember 2023. Data ini akan digunakan sebagai *data training* untuk *fine tuning*. *Data preprocessing* yang dilakukan dilakukan secara manual karena harus membaca satu persatu pertanyaan dan jawaban untuk dimasukkan ke dalam *format training* untuk *fine tuning* dari *open ai*. *Data preprocessing* yang dilakukan hanya membenarkan kata kata yang memiliki salah penulisan. Setelah *data training* sudah sesuai *format* yang ada maka akan di *upload* dan menunggu proses model di *fine tune* sesuai dengan *data training* yang ada

3.4 Desain Sistem

Desain sistem membahas bagaimana pertanyaan akan diproses hingga menghasilkan jawaban yang baik akan dijelaskan menggunakan *flowchart*. Sistem mencakup memahami pertanyaan yang diajukan, mencari data yang dibutuhkan, dan memberikan jawaban yang sesuai dengan pertanyaan yang diberikan.

3.4.1 Prompt Engineering

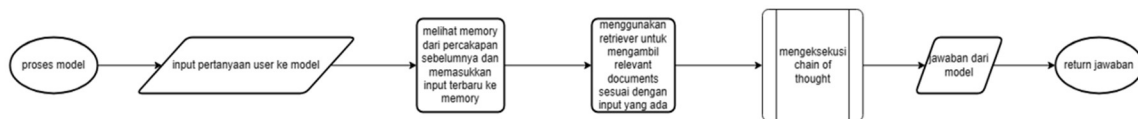


Gambar 3.2 Flowchart User untuk Prompt Engineering

Pada gambar 3.2, sistem akan berjalan di *background* menunggu pertanyaan yang masuk dari calon mahasiswa berupa teks. Saat ada pertanyaan yang masuk, maka pertanyaan itu akan menjadi input sistem. Pertanyaan akan di proses dengan model yang sudah dibuat dan model akan

menghasilkan antara 2 jawaban. Saat model bisa mendapatkan jawaban dari pertanyaan yang diajukan, maka output akan berupa jawaban dari pertanyaan tersebut. Jika model tidak bisa menemukan jawaban dari pertanyaan, maka akan masuk proses untuk memberitahu petugas administrasi untuk menjawab pertanyaan yang ada. Langkah yang berwarna orange pada gambar 3.2 akan dikerjakan oleh Michael Christian Wijaya dengan nrp C14200087 dalam tugas akhir berjudul *Leads Categorization* dengan *Support Vector Machine* dan *Gaussian Naive Bayes* untuk Student Admission pada Universitas X.

Ada beberapa proses model yang akan dicoba pada pengerjaan skripsi ini. Proses ini akan menggunakan *chain of thought prompting* untuk mencari jawaban dari pertanyaan calon mahasiswa. Model yang pertama akan menggunakan prompt yang memiliki *instruction* untuk mencari jawaban yang ada. Model yang kedua menggunakan prompt yang berisi contoh pertanyaan, contoh proses berpikir, dan contoh jawaban yang diharapkan. Model yang ketiga menggunakan prompt yang berisi contoh pertanyaan, dan contoh jawaban yang diharapkan tanpa contoh proses berpikir seperti di model kedua. Dari ketiga model yang ada akan diberikan ke pihak universitas X untuk *testing*. Universitas X akan memilih 1 model terbaik dari 3 model yang ada dan sehingga model yang dipilih akan digunakan sebagai model *chatbot* untuk membantu universitas X.



Gambar 3.3 Proses Model A

Pada gambar 3.3 akan dijelaskan alur dari proses model A. Pertanyaan dari calon mahasiswa akan dijadikan sebagai input untuk model yang ada. Model juga akan melihat *memory* dari percakapan sebelumnya dan menjadikannya input. *Memory* dapat diatur kapasitasnya sesuai dengan kebutuhan. Model akan menganalisa apakah pertanyaan baru berhubungan dengan pertanyaan yang diajukan sebelumnya. Pertanyaan baru juga akan disimpan untuk ke dalam *memory* untuk pertanyaan selanjutnya. Model akan memahami pertanyaan dan mencari dokumen yang relevan dengan pertanyaan yang ada. Setelah mendapatkan dokumen yang dibutuhkan, model akan berusaha mencari jawaban dari pertanyaan menggunakan *chain of thought prompting*. Alur dari *chain of thought prompting* dapat dilihat pada Gambar 3.4.



Gambar 3.4 Flowchart Chain of Thought Prompting A

Model akan memahami pertanyaan yang ada dengan baik terlebih dahulu. Model lalu akan memikirkan beberapa jawaban yang mungkin untuk menjawab pertanyaan yang ada. Model akan mengevaluasi dan memberikan alasan dari jawaban yang sudah dipikirkan sebelumnya. Terakhir model akan memilih jawaban terbaik dari hasil evaluasi yang sudah dilakukan dan menjadikan jawaban tersebut menjadi *output*. Berikut contoh pertanyaan dan jawaban dari *Chain of Thought Prompting A*.

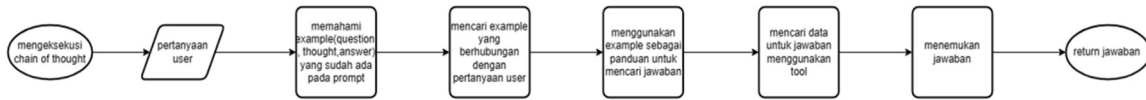
- Pertanyaan: saya tertarik dengan informatika
- Jawaban: Universitas X menawarkan program studi Informatika yang mencakup berbagai topik penting seperti Data Science, Cyber Law, Internet of Things, dan IT Governance. Dengan fokus pada kecerdasan buatan dan ilmu data, program ini dirancang untuk mempersiapkan mahasiswa dalam menghadapi tuntutan industri yang semakin kompleks dan berkembang pesat. Dengan didukung oleh dosen-dosen kompeten dan kurikulum yang relevan, program Informatika di unversitas X dapat menjadi pilihan yang tepat bagi Anda yang tertarik dalam bidang teknologi informasi.



Gambar 3.5 Proses Model B dan Proses Model C

Pada gambar 3.5 akan dijelaskan alur dari proses model B dan C. Proses model B dan C sama persis, perbedaan terletak pada *chain of thought* saja. Pertanyaan dari calon mahasiswa akan dijadikan sebagai input untuk model yang ada. Model juga akan melihat *memory* dari percakapan sebelumnya dan menjadikannya input. Model akan menganalisa apakah pertanyaan baru berhubungan dengan pertanyaan yang diajukan sebelumnya. Pertanyaan baru juga akan disimpan untuk ke dalam *memory* untuk pertanyaan selanjutnya. Model akan berusaha mencari jawaban dari pertanyaan menggunakan *chain of thought prompting*. Alur dari *chain of thought prompting* model

B dapat dilihat pada Gambar 3.6 dan Gambar 3.7 untuk Alur dari *chain of thought prompting* model C.

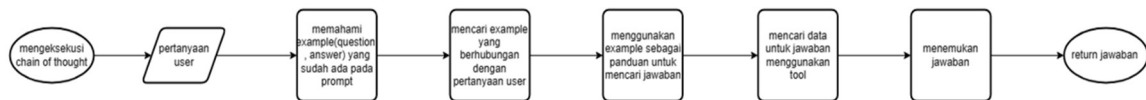


Gambar 3.6 Flowchart Chain of Thought Prompting Model B

Model akan memahami *example(question, thought, answer)* yang ada pada *prompt* terlebih dahulu. Model lalu akan mencari *example* yang berhubungan dengan pertanyaan yang ada. Model akan menggunakan *example* tersebut sebagai panduan untuk mencari jawaban. Model akan mempelajari *thought* yang ada pada *example*. Model akan menggunakan tool untuk mencari data yang dibutuhkan untuk menjawab pertanyaan yang ada. Model akan menemukan jawaban dan menjadikan jawaban tersebut sebagai *output*. Berikut contoh pertanyaan dan jawaban dari *Chain of Thought Prompting B*.

- Pertanyaan: saya tertarik dengan informatika
- Jawaban: Jurusan yang cocok untuk Anda di Universitas X adalah Teknik Informatika. Program studi Teknik Informatika di Universitas X menawarkan berbagai konsentrasi seperti Cyber Security, Technopreneurship, dan Artificial Intelligence and Data Science. Anda akan mempelajari tentang pengembangan perangkat lunak, keamanan informasi, dan analisis data. Selain itu, kurikulum kami dirancang untuk memenuhi standar internasional dan memberikan pengalaman belajar internasional selama satu tahun di universitas mitra kami di luar negeri.

Lulusan Teknik Informatika di Universitas X memiliki prospek karir yang sangat baik di berbagai industri, seperti perusahaan teknologi, perbankan, konsultan IT, dan banyak lagi. Jika Anda tertarik untuk mengembangkan keterampilan di bidang teknologi informasi, Teknik Informatika di Universitas X adalah pilihan yang tepat untuk Anda.



Gambar 3.7 Flowchart Chain of Thought Prompting Model C

Model akan memahami *example (question, answer)* yang ada pada *prompt* terlebih dahulu. Model lalu akan mencari *example* yang berhubungan dengan pertanyaan yang ada. Model akan menggunakan *example* tersebut sebagai panduan untuk mencari jawaban. Model akan

menggunakan tool untuk mencari data yang dibutuhkan untuk menjawab pertanyaan yang ada. Model akan menemukan jawaban dan menjadikan jawaban tersebut sebagai *output*. Berikut contoh pertanyaan dan jawaban dari *Chain of Thought Prompting C*.

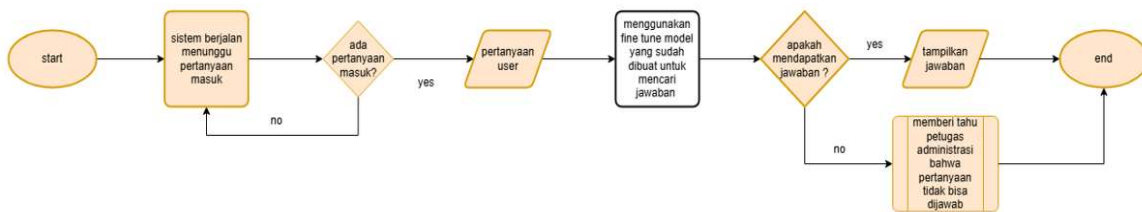
- Pertanyaan: saya tertarik dengan informatika
- Jawaban: Jurusan yang cocok untuk Anda di Universitas X adalah Teknik Informatika. Program studi Teknik Informatika di Universitas X dirancang untuk mempersiapkan mahasiswa dengan pengetahuan dan keterampilan dalam pengembangan perangkat lunak, jaringan komputer, keamanan informasi, dan teknologi informasi terkini. Anda akan belajar tentang pemrograman, desain basis data, keamanan jaringan, dan teknologi cloud computing.

Mengapa memilih di Universitas X?

- Kurikulum yang terus diperbarui sesuai dengan perkembangan teknologi informasi terkini.
- Pengajar yang berpengalaman dan ahli di bidangnya.
- Kesempatan untuk mendapatkan pengalaman internasional melalui pertukaran pelajar atau magang di perusahaan teknologi terkemuka.
- Prospek karir yang baik di berbagai industri, termasuk teknologi informasi, perbankan, e-commerce, dan lainnya.

Jika Anda tertarik dengan bidang informatika, maka Teknik Informatika di Universitas X adalah pilihan yang tepat untuk Anda.

3.4.2 Fine Tuning



Gambar 3.8 Flowchart User untuk Fine Tuning

Sistem akan berjalan di *background* menunggu pertanyaan yang masuk dari calon mahasiswa berupa teks seperti pada Gambar 3.8. Saat ada pertanyaan yang masuk, maka pertanyaan itu akan menjadi input sistem. Pertanyaan akan di proses menggunakan model yang sudah di *fine tune* sebelumnya menggunakan data dari *conversation history*. Saat model bisa mendapatkan jawaban

dari pertanyaan yang diajukan, maka output akan berupa jawaban dari pertanyaan tersebut. Jika model tidak bisa menemukan jawaban dari pertanyaan, maka akan masuk proses untuk memberitahu petugas administrasi untuk menjawab pertanyaan yang ada. Langkah yang berwarna orange pada gambar 3.8 akan dikerjakan oleh Michael Christian Wijaya dengan nrp C14200087 dalam tugas akhir berjudul *Leads Categorization* dengan *Support Vector Machine* dan *Gaussian Naive Bayes* untuk Student Admission pada Universitas X.



Gambar 3.9 Flowchart Fine Tuning

Alur dari pembuatan *fine tune* model dapat dilihat pada gambar 3.9. Model membutuhkan *data training* yang didapatkan dari *conversation history*. *Data conversation history* meliputi pertanyaan seputar biaya kuliah, prosedur pendaftaran, jurusan yang tersedia, dan tes masuk universitas X dari calon mahasiswa dan jawaban dari staf admisi universitas X. *Data conversation history* harus diubah sesuai format *data training* dari OpenAI. *Data training* akan di *upload* ke server OpenAI agar bisa digunakan untuk membuat *fine tune* model. OpenAI akan mengirimkan *email* saat *fine tune* model sudah selesai dibuat.