

1. PENDAHULUAN

1.1. Latar Belakang Masalah

Universitas Kristen Petra merupakan salah satu institusi perguruan tinggi swasta di Surabaya. Di dalam kegiatan studinya sendiri, UK Petra merancang kurikulum untuk memenuhi kebutuhan industri 4.0. Kurikulum tersebut diintegrasikan ke dalam silabus mata kuliah yang terbagi ke dalam mata kuliah wajib, konsentrasi, pilihan, lintas prodi, serta mata kuliah umum untuk memenuhi capaian pembelajaran sesuai dengan program studi yang diambil mahasiswa. Selain itu, terdapat juga program Merdeka Belajar Kampus Merdeka (MBKM) yang dapat diambil mahasiswa seperti Magang dan Studi Independen dan Bersertifikat (MSIB), Indonesia International Student Mobility Awards (IISMA), dan lainnya yang dapat dikonversikan ke beberapa mata kuliah terkait sebanyak maksimal 20 SKS dalam 1 semester (Kementerian Pendidikan dan Kebudayaan RI, 2022). Oleh karena pilihan yang luas dalam kurikulum itulah, mahasiswa diberikan kebebasan dalam kegiatan studinya untuk mengambil mata kuliah yang ditawarkan melalui PRS.

Pendaftaran Rencana Studi atau PRS merupakan salah satu ketentuan wajib yang perlu dilakukan mahasiswa tiap pergantian semester. Tentu mahasiswa setidaknya pernah menghadapi dilema dalam penyusunan strategi pengambilan PRS oleh karena beberapa faktor. Faktor-faktor yang mempengaruhi pengambilan mata kuliah dari mahasiswa sendiri antara lain motivasi karir, pemenuhan syarat program studi dan umum, serta minat pribadi mahasiswa (Lee et al., 2021). Selain itu, setiap mahasiswa memiliki kebebasan yang berbeda dalam pengambilan PRS karena ketentuan yang berbeda-beda antar program dan semester yang telah mereka tempuh (Lee et al., 2021) serta waktu pembukaan mata kuliah maupun jumlah kuota yang tersedia. Kemudian hal ini juga diperkuat dengan banyaknya mata kuliah konsentrasi dan pilihan yang wajib diambil mahasiswa sesuai program studinya diperkirakan sekitar 60% dalam kurikulum program studi UK Petra Surabaya (Universitas Kristen Petra, 2022). Performa mahasiswa pada semester sebelumnya pun menjadi faktor penting yang perlu dievaluasi dalam penyusunan rencana studi mereka kedepannya yang lebih baik lagi. Hal-hal tersebutlah yang membuat pengambilan PRS menjadi kompleks dan menjadi perhatian UK Petra Surabaya. Salah satu upaya UK Petra membantu mahasiswa dalam PRS adalah dengan menyediakan perwalian untuk berkonsultasi

terkait perencanaan PRS. Akan tetapi, perwalian sendiri memiliki beberapa keterbatasan seperti keterbatasan waktu dan jumlah mahasiswa yang banyak dengan kebutuhan yang berbeda-beda.

Untuk mengatasi permasalahan-permasalahan tersebut, *recommendation system* (RecSys) dapat menjadi solusi yang tepat. Pada penelitian yang dilakukan oleh (Al-Twijri et al., 2022) dan (Premalatha et al., 2022) dilakukan pembuatan *course recommendation system* menggunakan teknik-teknik *Data Mining* dan *Natural Language Processing* seperti (ES)²P (*Evolutionary Search of Emerging Sequential Patterns*) dan juga *Long Short Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU). (ES)²P pada penelitian (Al-Twijri et al., 2022) digunakan untuk mencari jalur perencanaan kurikulum berdasarkan jalur mahasiswa-mahasiswa yang telah ada sebelumnya maupun solusi jalur terbaik dari mahasiswa pendahulunya. Sementara itu, LSTM dan GRU pada penelitian (Premalatha et al., 2022) digunakan untuk mencari kesamaan dari tiap mata kuliah sesuai dengan *domain expertise* mahasiswa. Di dalam penelitian-penelitian tersebut, *recommendation system* juga sama-sama dibuat dengan mengevaluasi kegiatan studi mahasiswa pada semester sebelumnya. Hasil dari metode kedua penelitian tersebut mampu memberikan rekomendasi yang baik bagi mahasiswa dalam perencanaan studi.

Akan tetapi, seiring berkembangnya jaman muncullah LLM yang merupakan salah satu jenis *Generative Artificial Intelligence*. *Large Language Models* seringkali digunakan dalam aplikasi NLP, dimana output akan diberikan berdasarkan input *query user* (Sætra, 2023). Selain itu, dalam penelitian (Zhao et al., 2024) ditemukan bahwa LLM memiliki kemampuan *generalization* dan *reasoning* yang lebih baik dibandingkan dengan model *Deep Neural Network* biasa. Oleh karena kemampuan itulah penggunaan LLM memiliki peran yang krusial dalam mengembangkan *recommendation system* yang lebih cerdas dan *personalized*. Namun di balik kemampuannya tersebut LLM memiliki beberapa limitasi salah satunya seperti *data-cutoff*. Hal ini ditunjukkan model Chat GPT Open-AI yaitu GPT 4 Turbo sendiri di-*training* dengan data hingga bulan Desember 2023 (OpenAI, 2015-2024). Oleh karena itulah, LLM seringkali mengalami halusinasi ketika diberikan pertanyaan yang memerlukan informasi terkini di luar rentang waktu *training data*-nya. Selain itu, *specific domain task* juga menjadi keterbatasan dari penggunaan LLM (Lewis et al., 2021).

Demi implementasi *recommendation system* yang lebih interaktif dan meningkatkan partisipasi aktif dari mahasiswa, sistem tersebut dapat diaplikasikan ke dalam chatbot. Aplikasi chatbot merupakan program AI yang mensimulasikan percakapan dalam bahasa alami dengan *user* melalui aplikasi *messaging*, aplikasi *mobile*, maupun *handphone* (Bhoir et al., 2022).

Chatbot telah terbukti memberikan manfaat nyata, salah satunya dalam memberikan kepuasan dan meningkatkan loyalitas *customer* dalam penggunaannya di sektor bisnis (Jenneboer et al., 2022). Oleh karena itulah, chatbot ini sendiri akan didesain sebagai *private assistant* yang mampu membantu mahasiswa dalam mengambil keputusan perencanaan PRS serta memberikan evaluasi atau *feedback* terkait performa mahasiswa pada semester sebelumnya. Aplikasi chatbot ini dibuat bukan dengan tujuan untuk menggantikan dosen wali melainkan membantu dosen wali dalam perwalian yang lebih efektif. Aplikasi ini akan menggunakan *Large Language Models* sebagai basis model untuk peningkatan kemampuan *generalization* dan *reasoning* RecSys baik dalam perencanaan maupun evaluasi. Dalam mengatasi keterbatasan LLM, metode *Retrieval-Augmented Generation* dapat diterapkan.

RAG sendiri menggabungkan *memory* parametrik dan non-parametrik untuk memaksimalkan generasi output dari LLM (Lewis et al., 2021). Komponen utama dari metode tersebut adalah aspek *Retrieval* yang mengambil konteks dari sumber data eksternal untuk generasi output LLM. Pada umumnya RAG menggunakan *vector embeddings* yang disimpan dalam bentuk indeks sebagai representasi sumber data eksternalnya untuk implementasi *semantic similarity search*. Namun melalui hasil penelitian (Xu et al., 2024) ditemukan penggunaan RAG menggunakan *knowledge graph* untuk proses *retrieval* yang lebih akurat oleh karena pengolahan informasi relasi dan struktur data yang lebih baik. Selain itu, RAG juga memiliki perkembangan dalam arsitektur dari yang awalnya *single-agent* menjadi *multi-agent* untuk implementasi yang lebih fleksibel dan efisien dalam menyelesaikan tugas.

Pada skripsi ini, akan dibuat aplikasi chatbot *private assistant* sebagai *recommendation system* yang ditujukan untuk membantu pengambilan PRS mahasiswa. Aplikasi akan dibuat dengan menggunakan LLM dan metode *hybrid RAG* yang menggabungkan *vector database* dan *knowledge graph* sebagai indeks *knowledge base*. Arsitektur RAG yang akan diimplementasikan ialah *multi-agent* yang terdiri dari agen perencanaan, evaluasi, dan RAG. Agen perencanaan akan memberikan rekomendasi pengambilan PRS sesuai dengan preferensi *user*, sementara agen evaluasi akan memberikan kesimpulan komprehensif terkait performa *user* pada semester sebelumnya. Agen RAG akan digunakan untuk tanya jawab ataupun meningkatkan kualitas respons chatbot. Data yang digunakan pada *knowledge base* berupa data kurikulum program studi, kelas, kartu rencana studi mahasiswa, serta kartu rencana studi mahasiswa lainnya yang memiliki tingkat di atas mahasiswa *user* atau yang sudah lulus. Untuk mencapai performa terbaik, akan dilakukan perbandingan terhadap konfigurasi *hyper-parameter* RAG yaitu *chunk*

size, dan *top-k* dalam chatbot. Diharapkan melalui skripsi ini mahasiswa akan terbantu dan mengalami peningkatan yang signifikan dalam kegiatan studi mereka.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, rumusan masalah skripsi ini dirumuskan sebagai berikut:

1. Apa konfigurasi *hyper-parameter (chunk size, top-k) Hybrid Retrieval-Augmented Generation (RAG)* yang mampu memberikan performa terbaik dalam perencanaan pengambilan PRS mahasiswa?
2. Apakah aplikasi chatbot mampu memberikan rekomendasi yang sesuai dengan minat atau konsentrasi mahasiswa secara *long-term* dalam perencanaan pengambilan PRS mahasiswa?

1.3. Tujuan Penelitian

Skripsi ini dilakukan dengan tujuan utama melakukan perbandingan *hyper-parameter* untuk mendapatkan performa terbaik pada aplikasi chatbot dalam pemberian rekomendasi PRS. Kemudian melalui konfigurasi tersebut chatbot akan dievaluasi lebih lanjut terkait kemampuannya dalam memberikan rekomendasi yang sesuai dengan minat atau konsentrasi mahasiswa secara *long-term*.

1.4. Ruang Lingkup

1. Skripsi ini dibatasi pada pemberian rekomendasi PRS program sarjana (S1) Informatika UK Petra (Infor-SIB-DSA) yang terdiri dari Informatika, Sistem Informasi dan Bisnis (SIB), dan Data Science and Analytics (DSA)
2. *Database* yang digunakan adalah *MongoDB, Qdrant, dan neo4j*
3. Model *Large Language Models* menggunakan *LLama 3.1* dari *Ollama* dan *Embedding Model* menggunakan *intfloat/multilingual-e5-large* dari *FastEmbed*
4. Chatbot menggunakan metode *Hybrid Retrieval-Augmented Generation* yang menggabungkan *vector database* dengan *knowledge graph*
5. Arsitektur RAG chatbot menggunakan *multi-agent RAG* yang memiliki 3 agen dengan outputnya masing-masing yaitu:
 - a. Perencanaan: Rekomendasi pengambilan PRS yang sesuai dengan personalisasi mahasiswa *user*

- b. Evaluasi: Kesimpulan komprehensif terkait performa siswa terhadap semester sebelum-sebelumnya dan *planning* awal
 - c. RAG: Tanya jawab seputar kurikulum atau meningkatkan kualitas respons chatbot
- 6. Aplikasi Chatbot terbagi ke dalam beberapa hak akses dengan fiturnya seperti berikut:
 - a. *User*
 - i. Authentication
 - ii. Assessment
 - iii. Chat Session
 - b. *Admin*
 - i. Authentication
 - ii. Manajemen Knowledge Base
- 7. Input data yang digunakan sebagai *knowledge base* antara lain:
 - a. Data kurikulum mata kuliah yang digunakan adalah pedoman kurikulum dan transisi kurikulum atau data kurikulum terbaru yang diakses melalui API SIM Petra
 - b. Data jurusan terkait prospek karir serta konsentrasi yang didapatkan dari buku pedoman dan dokumen *prospectus* UK Petra
 - c. Data kontekstual seputar wawasan spesialis, LEAP, dan lainnya sebagai referensi pendukung dalam pemberian rekomendasi mata kuliah PRS
 - d. Data kelas mata kuliah yang terbuka pada waktu pengambilan PRS yang diakses melalui API SIM Petra
- 8. Input data mahasiswa yang digunakan merupakan basis data detail mahasiswa dan data KRS yang diakses melalui API SIM Petra secara anonim berupa:
 - a. Data mahasiswa *user* untuk personalisasi
 - b. Data mahasiswa lainnya yang digunakan menjadi acuan jalur studi untuk *long-term planning*
- 9. Wawancara sederhana dengan dosen wali akan dilakukan untuk mengetahui dokumen kontekstual tambahan apa saja yang diperlukan dalam *knowledge base* chatbot untuk mendukung performa *recommender system* PRS yang lebih baik.
- 10. *Pre-processing* data input menjadi *knowledge base* melalui *feature engineering*, *chunking*, *embeddings*, *indexing*, dan pembuatan *knowledge graph* akan dilakukan

11. Output aplikasi chatbot berupa jawaban dalam bentuk teks terkait rekomendasi PRS dan evaluasi performa mahasiswa
12. Pengujian aplikasi chatbot dilakukan dengan cara:
 - a. Percobaan menggunakan dataset *testing* untuk melakukan perbandingan dan menentukan *hyper-parameter* terbaik menggunakan *framework llama-index*
 - b. Percobaan mandiri menggunakan 33 data anonimus mahasiswa Infor-SIB-DSA untuk menguji kesesuaian rekomendasi chatbot dengan data mahasiswa tersebut untuk kategori IPS di bawah 3 & di atas 3 serta kategori angkatan 2021 & 2022
 - c. Pengujian aplikasi untuk mengukur performa aplikasi chatbot secara keseluruhan untuk tiap agen serta tingkat halusinasi
13. Dataset *testing* berupa pasangan pertanyaan dan jawaban dibuat menggunakan LLM dan dikurasi oleh penulis untuk pengujian performa chatbot
14. Skripsi ini dibatasi pada pembuatan aplikasi dan *API services pipeline* agen chatbot saja tidak termasuk implementasi ke aplikasi *Petra Mobile*

1.5. Manfaat Penelitian

Manfaat dari skripsi ini adalah peningkatan efisiensi pengambilan SKS mata kuliah mahasiswa yang terpersonalisasi serta performa mahasiswa secara maksimal. Selain itu juga proses perwalian akan menjadi lebih efisien dengan bantuan chatbot sebagai *companion* dosen wali.

1.6. Metodologi Penelitian

1. Studi Literatur
 - a. *Recommendation System (RecSys)*
 - b. *Large Language Model (LLM)*
 - c. *Vector Embeddings*
 - d. *Prompt Engineering*
 - e. *Knowledge Graph (KG)*
 - f. *Retrieval-Augmented Generation (RAG)*
 - g. *Multi-Agent RAG*
 - h. *Dataset*

5. Pengujian
 - a. Pengujian menggunakan dataset *testing* untuk melakukan perbandingan dan menentukan *hyper-parameter chunk size* dan *top-k* terbaik
 - b. Pengujian kesesuaian rekomendasi chatbot menggunakan data anonim mahasiswa
6. Analisis Data
 - a. Pengolahan hasil pengujian performa aplikasi chatbot dengan *test dataset*
 - b. Pengolahan hasil pengujian rekomendasi chatbot menggunakan data anonim mahasiswa
 - c. Pengolahan hasil survey tingkat kepuasan mahasiswa
7. Pengambilan Kesimpulan
 - a. Penarikan kesimpulan dari pengujian performa aplikasi chatbot dengan *test dataset*
 - b. Penarikan kesimpulan dari pengujian rekomendasi chatbot menggunakan data anonim mahasiswa
 - c. Penarikan kesimpulan dari survey tingkat kepuasan mahasiswa
8. Penulisan Laporan
 - a. Penulisan laporan hasil yang diperoleh

1.7. Sistematika Penulisan

Berikut adalah sistematika penulisan buku skripsi ini dibagi menjadi beberapa bab, yaitu:

1. BAB I: PENDAHULUAN

Bab I membahas terkait latar belakang, rumusan masalah, tujuan, ruang lingkup, manfaat, serta metodologi penelitian yang digunakan dalam skripsi.
2. BAB II: LANDASAN TEORI

Bab II membahas terkait teori-teori serta penelitian yang digunakan dalam skripsi.
3. Bab III: ANALISIS DAN DESAIN SISTEM

Bab III membahas terkait perencanaan analisis serta desain sistem yang digunakan dalam skripsi.
4. Bab IV: IMPLEMENTASI SISTEM

Bab IV membahas terkait implementasi dari sistem yang telah dibahas pada bab III sebelumnya.

5. Bab V: PENGUJIAN SISTEM

Bab V membahas terkait pengujian sistem aplikasi yang telah dibuat berdasarkan implementasi pada bab IV sebelumnya.

6. Bab VI: KESIMPULAN & SARAN

Bab VI membahas terkait kesimpulan dari pengujian yang telah dilakukan pada bab V dan saran untuk pengembangan mendatang.