



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Dhani Dwi Septian Bangun

NIM : 211402120

Judul diajukan oleh* : ☐ Dosen

☐ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) :

1. Data Science and Intelligent System
2. Machine Learning

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I:
Mohammad Fadly Syah Putra, M.Sc
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II:
Dr. Muhammad Anggia Muchtar, S.T., MM.IT.

Paraf Calon Dosen Pembimbing I

Medan, November 2024
Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

** Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT)
NIP.198908172019032023



RINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	Implementasi Graph-Based Retrieval Augmented Generation (GraphRAG) pada Large Language Model (LLM) pada Pengembangan Chatbot Virtual Assistant Universitas Sumatera Utara
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Teknologi informasi berkembang pesat dan telah membawa dampak yang signifikan di berbagai bidang, termasuk bidang pendidikan. Teknologi membantu pekerjaan manusia dalam berbagai bidang menjadi lebih cepat terselesaikan. Di era digital ini, universitas dituntut untuk menyediakan layanan informasi yang cepat, akurat, dan efisien guna memenuhi kebutuhan mahasiswa dan masyarakat luas (Khan, 2018). Universitas Sumatera Utara (USU) memiliki berbagai layanan berbasis website seperti portal akademik, situs resmi universitas, situs fakultas, dan layanan lainnya. Namun, banyaknya website yang tersedia sering kali membuat mahasiswa, dosen, maupun staf kesulitan mengakses informasi yang mereka butuhkan secara cepat dan efisien. Pengguna harus mengunjungi beberapa situs atau aplikasi berbeda untuk menemukan informasi spesifik, yang dapat mengakibatkan kebingungan dan pemborosan waktu. Untuk mengatasi kendala ini, Universitas Sumatera Utara perlu mengembangkan solusi yang lebih terintegrasi dan ramah pengguna.</p> <p>Salah satu alternatif yang dapat diterapkan adalah pengembangan virtual assistant berbasis kecerdasan buatan menggunakan chatbot. Chatbot memiliki kemampuan untuk memahami permintaan pengguna dan memberikan respons yang relevan dalam waktu singkat, sehingga menghemat waktu dibandingkan pencarian manual melalui situs web atau mesin pencari seperti Google. Dengan kemampuan chatbot untuk merespons pertanyaan dengan cepat dan tepat, diharapkan dapat mengurangi beban kerja staf administrasi serta meningkatkan pengalaman pengguna dalam mengakses informasi akademik (Hussain et al., 2021). Salah satu penerapan Chatbot yaitu menggunakan Large Language Models (LLM). LLM adalah salah satu jenis model AI yang dapat memproses dan menghasilkan teks bahasa. Model pada LLM umumnya dilatih dengan menggunakan sejumlah besar data teks dan menggunakan teknik pembelajaran mendalam atau deep learning untuk mempelajari pola dan struktur bahasa (M.U. Hadi et al, 2021). Kehadiran LLM sendiri menjadi sebuah fenomena luar biasa pada beberapa tahun terakhir ini, tepatnya setelah kehadiran ChatGPT yang dirilis oleh perusahaan OpenAI pada bulan November 2022 yang lalu. Hanya dalam kurun waktu dua bulan setelah perilisannya, pada Januari 2023 jumlah pengguna ChatGPT telah mencapai 100 juta pengguna di seluruh dunia (Hu & Krystal, 2023).</p> <p>Hal ini mengindikasikan bahwa chatbot dapat menjadi solusi efektif untuk meningkatkan aksesibilitas informasi, khususnya di lingkungan universitas. Chatbot semakin populer digunakan dalam bidang Pendidikan karena dapat meningkatkan interaksi antara institusi dan pengguna. Dalam studi yang dilakukan oleh Fatima et al. (2020), ditemukan bahwa chatbot dapat membantu mengurangi waktu tunggu untuk mendapatkan informasi dan menyediakan dukungan 24/7 kepada pengguna. Hal ini sangat penting dalam konteks akademik, di mana informasi sering kali dibutuhkan di luar jam kerja.</p> <p>Dalam penelitian ini, metode Graph-Based Retrieval Augmented Generation (GraphRAG) diimplementasikan untuk mendukung performa chatbot berbasis LLM. Selain itu, pendekatan <i>GraphRAG</i> tidak hanya meningkatkan akurasi jawaban, tetapi juga memastikan bahwa</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI SI TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

informasi yang disampaikan selalu relevan dan terkini. Hal ini menjadi penting dalam lingkungan universitas, di mana kebutuhan informasi dapat sangat dinamis, mencakup jadwal akademik, pengumuman, prosedur administratif, hingga data penelitian. Dengan menerapkan GraphRAG, chatbot mampu mengakses dan menghubungkan berbagai sumber data internal Universitas Sumatera Utara secara efektif, sehingga memberikan pengalaman yang lebih personal dan interaktif bagi penggunanya.

Berdasarkan permasalahan yang sudah diuraikan sebelumnya, maka sangat diperlukan solusi untuk mengatasinya. Salah satu cara untuk meminimalisir waktu dalam penyampaian informasi akademik di Universitas Sumatera Utara dengan penerapan chatbot, diharapkan dapat mengatasi masalah keterlambatan dalam mengakses informasi akademik.

Terdapat penelitian terkait yang dilakukan oleh Lilit Galstyan dan Hovhannes Martirosyan (2024) yang berjudul “SmartAdvisor University Chatbot Spring 2024”. Penelitian ini memanfaatkan kemampuan model pembuatan teks untuk menghadirkan chatbot penasihat akademik yang inovatif dan personal. Tujuan utama penelitian ini untuk mengembangkan sumber daya menyeluruh yang menggabungkan pengetahuan tentang universitas program studi untuk mahasiswa. Penelitian ini memanfaatkan basis data vektor bersama dengan kerangka kerja Retrieval-Augmented Generation (RAG).

Penelitian lainnya dilakukan oleh Rachmat & Kesuma (2024) yang berjudul Implementasi *Large Language Models Gemini* Pada Pengembangan Aplikasi Chatbot Berbasis Android. Penelitian ini dilakukan untuk memahami bagaimana implementasi API key Gemini. API key Gemini dapat dilakukan pada aplikasi AI, khususnya pada aplikasi chatbot, dan apakah aplikasi chatbot yang dibangun menggunakan Gemini dapat berfungsi sesuai dengan fungsionalitas Gemini LLM itu sendiri. Setelah aplikasi chatbot yang dimaksud selesai dibuat dan diuji coba, didapatkan bahwa aplikasi chatbot yang dihasilkan aplikasi chatbot yang dihasilkan berfungsi sesuai dengan yang diharapkan dan mampu memanfaatkan fungsionalitas LLM Gemini.

Penelitian lainnya dilakukan oleh Julius Odede & Ingo Frommholz (2024) yang berjudul JayBot – Aiding University Students and Admission with an LLM-based Chatbot. Penelitian ini menyajikan JayBot yaitu sistem chatbot berbasis LLM yang bertujuan untuk meningkatkan pengalaman pengguna dosen, staf, calon mahasiswa dan mahasiswa di sebuah universitas di Inggris.

Penelitian lainnya dilakukan oleh Radhakrishnan & Dias (2023) yang berjudul Use of a ChatBot-Based Advising System for the Higher-Education System. Penelitian ini memperkenalkan mekanisme yang layak untuk sistem pemberian saran berbasis chatbot. penelitian ini menggunakan Large Language Model (LLM) yang bersifat opensource yang dikombinasikan dengan basis pengetahuan khusus untuk mengatasi aspek-aspek penting yang diperlukan untuk sistem pemberian saran berbasis chatbot, seperti personalisasi, memori percakapan, dan kemudahan pemeliharaan. Sistem yang diusulkan menunjukkan tingkat akurasi respons sebesar 89% yang membuktikan bahwa pendekatan baru dari arsitektur berbasis komponen ini unggul dalam hal kinerja jika dibandingkan dengan pendekatan serupa.

Berdasarkan dari latar belakang dan beberapa penelitian terdahulu yang telah diuraikan di atas, dapat ditarik kesimpulan bahwa sistem *Natural Language Processing* merupakan suatu solusi yang tepat. Maka penulis mengajukan untuk melakukan penelitian mengenai pembuatan system **“Implementasi Graph-Based Retrieval Augmented Generation (GraphRAG) pada Large Language Model (LLM) pada Pengembangan Chatbot Virtual Assistant Universitas Sumatera Utara”**.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Penelitian Terdahulu

No.	Penulis	Judul	Tahun
1.	Lilit Galstyan dan Hovhannes Martirosyan	SmartAdvisor University Chatbot Spring 2024	2024
2.	Rachmat & Kesuma	Implementasi <i>Large Language Models Gemini</i> Pada Pengembangan Aplikasi Chatbot Berbasis Android	2024
3.	Leonardo pasquarelli, Charles Koutcheme, Arto Hellas	Comparing the Utility, Preference, and Performance of Course Material Search Functionality and Retrieval-Augmented Generation Large Language Model (RAG-LLM) AI Chatbots in Information-Seeking Tasks	2024
4.	Odede Ingo Frommholz JAOdede	JayBot – Aiding University Students and Admission with an LLM-based Chatbot	2024
5.	Joni Salminen, Soon-gyo Jung, Johanne Medina, Kholoud Aldous, Jinan Azem, Waleed Akhtar, Bernard J. Jansen	Using Cipherbot: An Exploratory Analysis of Student Interaction with an LLM-Based Educational Chatbot	2024
6.	Radhakrishnan & Dias	Use of a ChatBot-Based Advising System for the Higher-Education System	2023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	7.	Pandey & Sharma	A comparative study of retrieval-based and generative-based chatbots using Deep Learning and Machine Learning	2023
	8.	Sruti Srinivasa Ragavan, Mohammad Amin Alipour	Revisiting Human Information Foraging: Adaptations for LLM-based Chatbots	2024
Rumusan Masalah	<p>Universitas Sumatera Utara (USU) menyediakan berbagai layanan berbasis website seperti portal akademik, situs resmi universitas, dan situs fakultas. Namun, dengan banyaknya situs yang tersedia, sering kali mahasiswa, dosen, dan staf mengalami kesulitan dalam mengakses informasi dengan cepat dan efisien. Proses pencarian informasi yang tersebar di berbagai platform dapat membingungkan pengguna dan membuang waktu yang seharusnya bisa digunakan untuk tujuan lainnya. Oleh karena itu dibutuhkan sebuah sistem chatbot berbasis kecerdasan buatan (AI) yang dapat menyederhanakan proses pencarian dengan memberikan jawaban yang cepat, relevan, dan mudah diakses oleh pengguna. Penerapan teknologi ini dalam chatbot di lingkungan universitas dapat meningkatkan pengalaman pengguna dalam mengakses informasi akademik dan administratif secara lebih cepat dan akurat.</p>			



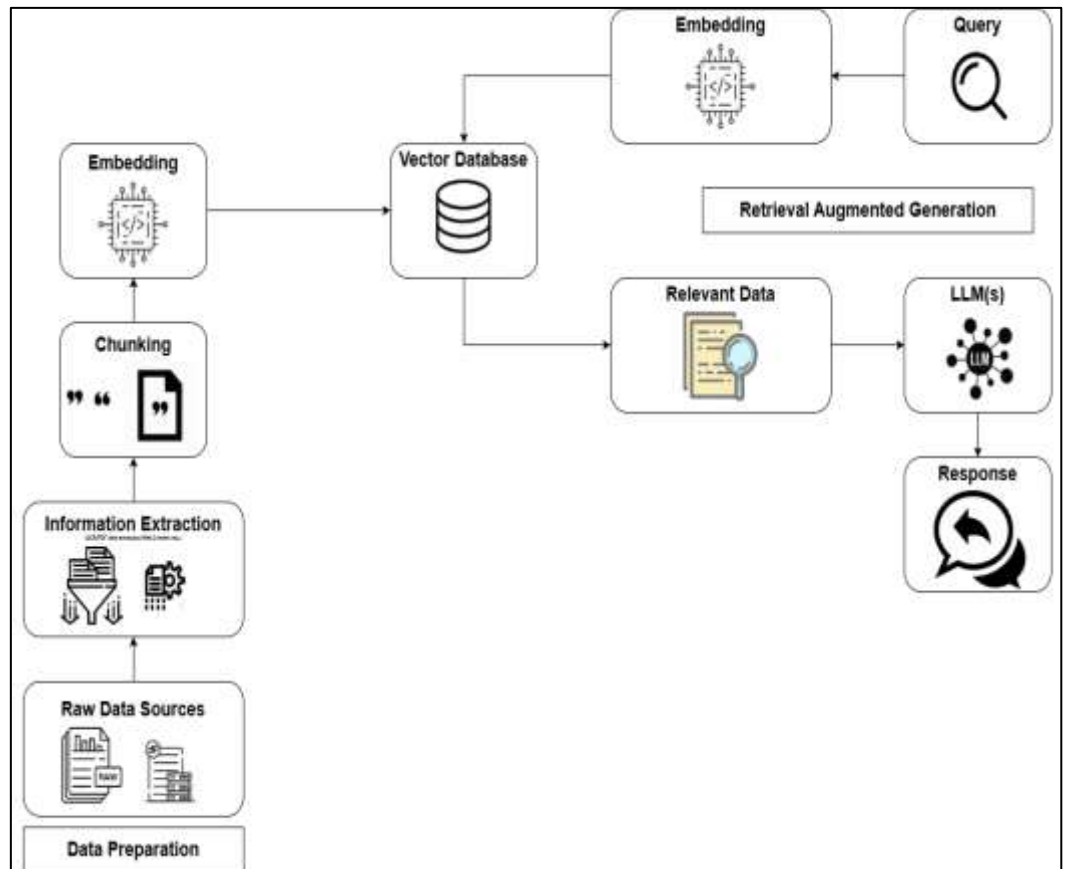
KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Metodologi



1. Raw Data Sources: Data mentah dikumpulkan dari berbagai sumber, seperti dokumen teks, PDF, database, atau file lainnya. Data ini biasanya tidak terstruktur atau semi-terstruktur.
2. Data Preparation: Data mentah dibersihkan dan diubah ke format yang lebih terstruktur, seperti menghapus duplikasi, mengisi data yang hilang, atau mengatur ulang formatnya agar siap untuk diproses lebih lanjut.
3. Information Extraction: Informasi penting diambil dari data yang telah dipersiapkan menggunakan teknik seperti OCR, Natural Language Processing (NLP), atau ekstraksi pola tertentu. Contoh hasilnya adalah entitas seperti nama, tanggal, atau angka.
4. Chunking: Data yang diekstraksi dipecah menjadi unit kecil (chunks) untuk memudahkan pengolahan. Potongan-potongan ini dapat berupa kalimat, paragraf, atau segmen teks.
5. Embedding: Setiap potongan data diubah menjadi representasi numerik atau vektor menggunakan algoritma embedding. Representasi ini memungkinkan komputer memahami hubungan antar data dalam bentuk matematika.
6. Vector Database: Vektor hasil embedding disimpan dalam basis data khusus untuk mendukung pencarian cepat berdasarkan kesamaan atau jarak antar vektor.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI SI TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	<ol style="list-style-type: none">7. Query: Pengguna memberikan input berupa pertanyaan atau kata kunci yang ingin dicari informasinya dalam sistem.8. Embedding (Query): Pertanyaan atau query pengguna diubah menjadi vektor embedding menggunakan metode yang sama seperti pada data, sehingga dapat dibandingkan dengan data yang ada.9. Retrieval: Sistem membandingkan vektor query dengan vektor data dalam database untuk menemukan data yang paling relevan. Proses ini dilakukan dengan mengukur kesamaan antar vektor.10. Relevant Data: Data yang dianggap paling relevan dengan query pengguna diambil dari database untuk diproses lebih lanjut.11. LLM(s): Data relevan yang telah ditemukan diproses oleh model bahasa besar (Large Language Model), seperti GPT, untuk menghasilkan jawaban atau respons yang sesuai dengan konteks query.12. Response: Jawaban atau informasi yang telah diolah oleh model bahasa besar diberikan kembali kepada pengguna dalam format yang mudah dipahami.
Referensi	<p>Khan, M.S. (2018). Impact of Information Technology on Education Sector. <i>International Journal of Advanced in Computer Science and Management Studies</i>, 6(5), 1-5</p> <p>Hussain, W., Shad, A. B., & Jabeen, F. (2021). Chatbots in Higher Education: A review. <i>International Journal of Education Technology in Higher Education</i>, 18(1), 1-5.</p> <p>Fatima, S., Aslam, N., & Hussain, M. (2020). Evaluating the Performance of Chatbots in Education: A systematic Review. <i>Journal of Education Computing Research</i>, 58(5), 1007-1027.</p> <p>M. U. Hadi, Q. Al-Tashi, R. Qureshi dan e.a., "Large Language Models: A Comprehensive Applications, Challenges, Limitations, and Future Prospects, "<i>TechRxiv</i>, vol.4, September 2023, https://doi.org</p>

Medan, 22 Januari 2025
Mahasiswa yang mengajukan,

(Dhani Dwi Septian Bangun)

NIM.211402120