IMPLEMENTASI LARGE LANGUAGE MODELS MENGGUNAKAN FRAMEWORK LANGCHAIN PADA QUESTION ANSWERING SYSTEM

SKRIPSI

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:
MUHAMMAD FARIZ RAMADHAN
NIM D1041211059

FAKULTAS TEKNIK
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Muhammad Fariz Ramadhan

NIM : D1041211059

menyatakan bahwa dalam skripsi yang berjudul " IMPLEMENTASI LARGE LANGUAGE MODELS MENGGUNAKAN FRAMEWORK LANGCHAIN PADA QUESTION ANSWERING SYSTEM" tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi manapun. Sepanjang pengetahuan Saya, tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Demikian pernyataan ini dibuat dengan sebenar-benarnya. Saya sanggup menerima konsekuensi akademis dan hukum di kemudian hari apabila pernyataan yang dibuat ini tidak benar.

Pontianak, 31 Juli 2025



Muhammad Fariz Ramadhan
NIM D1041211059

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS TANJUNGPURA
FAKULTAS TEKNIK
Jalan Prof. Dr. H. Hadari Nawawi Pontianak 78124
Telepon (0561) 740186, WA: +6282152280907
Email : ft@untan.ac.id Website : http://teknik.untan.ac.id

HALAMAN PENGESAHAN

IMPLEMENTASI LARGE LANGUAGE MODELS MENGGUNAKAN FRAMEWORK LANGCHAIN PADA QUESTION ANSWERING SYSTEM

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:

Muhammad Fariz Ramadhan
NIM D1041211059

Telah dipertahankan di depan Penguji Skripsi pada tanggal 31 Juli 2025
dan diterima sebagai salah satu persyaratan untuk memperoleh gelar sarjana.

Susunan Penguji Skripsi:

Ketua,

Prof. Dr. Herry Sujaini, S.T., M.T.
NIP. 196806291997021001

Penguji Utama,

Dr. Arif Bijaksana Putra Negara, S.T., M.T.
NIP. 197208081998021002

Sekretaris,

Helen Sastypratiwi, S.T., M.Eng.
NIP. 198601172012122004

Penguji Pendamping,

Izuardo Zulkarnain, S.Pd., M.T.
NIP. 199405172024061001

Pontianak, 31 Juli 2025
Dekan,

Dr.-Ing. Ir. Slamet Widodo, M.T., IPM
NIP. 196712231992031002

HALAMAN PERSEMBAHAN

Skripsi ini Fariz persembahkan dan perjuangkan untuk kedua orang tua (Bapak H Ramdaniansyah, S.IP., M.AP., dan Ibu Hj Ellyani, S.Pd) dan saudara-saudaraku (Kakak Gita Amanda Putri (Rahimahullah) dan Abangku Muhammad Rifky Fikrullah, S.Sos). Telah banyak Cinta dan Kasih sayang yang diberikan kepada ku. Terima Kasih banyak

Dari titisan ilmu-Mu kupersembahkan goresan tanganku bagi ilmu pengetahuan Indonesia

KATA PENGANTAR

Puji dan Syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena berkat Rahmat dan Hidayah-Nya, saya dapat menyelesaikan skripsi dengan judul "IMPLEMENTASI LARGE LANGUAGE MODELS MENGGUNAKAN FRAMEWORK LANGCHAIN PADA QUESTION ANSWERING SYSTEM".

Penelitian ini mengimplementasikan tentang bagaimana Large Language Models dapat diimplementasikan sebagai model untuk Question Answering System yang dapat memberikan jawaban dan informasi seputar pedoman akademik Universitas Tanjungpura. Dengan memanfaatkan framework Langchain dan metode RAG, pengembangan aplikasi berbasis LLM dapat diimplementasikan dan digunakan.

Skripsi ini dibuat sebagai langkah terakhir dalam menyelesaikan kuliah pada program studi Sarjana Informatika dan sebagai salah satu syarat kelulusan. Penulisan skripsi ini tidak lepas dari berbagai pihak, oleh karena itu dalam kesempatan ini saya ingin menyampaikan terima kasih yang sebesar-besarnya kepada Prof. Dr. Herry Sujaini, S.T., M.T. dan Helen Sastypratiwi, S.T., M.Eng sebagai pembimbing utama dan pembimbing pendamping yang telah memberikan bimbingan, masukan, dan diskusi yang intensif mengenai penulisan serta alur penelitian. Terima kasih juga kepada Haried Novriando, S.Kom., M.Eng. selaku Dosen Pembimbing Akademik yang telah mengarahkan dan memberikan dukungan saya. Terima kasih banyak kepada kedua orang tua penulis yaitu H Ramdaniansyah S.IP., M.AP. dan Hj Ellyani serta saudara/saudari yang sangat kami sayangi yang telah memberikan dukungan penuh.

Penulis menyadari bahwa masih banyak terdapat kekurangan dalam tulisan ini, oleh karena itu dengan segenap kerendahan hati penulis mengharapkan kritik dan saran yang membangun demi kesempurnaan laporan ini dan penulis kedepannya. Penulis berharap semoga penelitian ini dapat bermanfaat bagi pengembangan ilmu pengetahuan dan bagi siapapun yang membacanya.

Pontianak, 31 Juli 2025
Penulis,

ABSTRAK

Informasi yang telah beredar luas dapat diperoleh dengan berbagai cara, salah satunya dengan bertanya kepada seseorang yang telah mengetahui informasi tersebut sebelumnya. Meskipun demikian, permasalahan terjadi saat seseorang yang mengetahui informasi tersebut tidak tersedia, sehingga memperoleh informasi dilakukan dengan cara membaca dokumen pedoman yang membutuhkan waktu lebih lama. Salah satu contohnya yaitu seseorang yang ingin mengetahui isi pedoman akademik Universitas Tanjungpura. Pedoman akademik merupakan dokumen yang berisi aturan, prosedur, dan informasi penting mengenai suatu institusi pendidikan. Teknologi saat ini telah berkembang untuk menyelesaikan masalah tersebut, salah satunya dengan membangun Question Answering System (QAS) yang dapat menjawab pertanyaan berdasarkan informasi aktual. Penelitian ini mengimplementasikan Large Language Models (LLM) dalam mengembangkan QAS menggunakan framework LangChain dengan pendekatan retrieval-augmented generation (RAG). Framework LangChain mempermudah proses pengembangan aplikasi yang memanfaatkan LLM. Proses yang diimplementasikan meliputi pembacaan dokumen PDF, pemisahan dokumen menjadi potongan (*chunks*), serta proses embedding dan penyimpanan ke dalam basis data vektor menggunakan Pinecone. LLM yang digunakan dalam pengujian yaitu llama3:8B (4.7GB), deepseek-r1:8B (5.2GB), dan phi-4:14B (9.1GB). Evaluasi sistem dilakukan menggunakan metrik ROUGEScore dan BERTScore, dengan hasil terbaik ditunjukkan oleh llama3:8B yang memperoleh skor ROUGE-1 sebesar 0,33 (33%), ROUGE-2 sebesar 0,243 (24%), dan BERTScore sebesar 0,73 (73%) dibandingkan dengan 2 model lainnya yaitu deepseek-r1:8b memperoleh skor ROUGE-1 sebesar 0,259 (25%), ROUGE-2 sebesar 0,154 (15%), dan BERTScore sebesar 0,689 (68%) dan phi-4:14b memperoleh skor ROUGE-1 sebesar 0,28 (28%), ROUGE-2 sebesar 0,208 (20%), dan BERTScore sebesar 0,722 (72%). Berdasarkan hasil evaluasi tersebut. llama3 dipilih sebagai LLM untuk QAS pedoman akademik karena mampu dalam menjawab pertanyaan berdasarkan dokumen dibandingkan 2 LLM lainnya. Kemudian dilakukan juga pengujian User Acceptance Testing (UAT) dengan 15 responden, yang menghasilkan rata-rata skor 3,96 atau setara dengan 79,2%. Hal tersebut menunjukkan kesimpulan bahwa sistem QAS diterima oleh pengguna karena dapat memberikan informasi dengan mudah, relevan, dan bermanfaat.

Kata kunci: Large Language Models (LLM), *Question Answering System (QAS), Langchain, retrieval-augmented generation, Pinecone,* pedoman akademik Universitas Tanjungpura, ROUGE-Score, BERTScore, *Flask*, UAT.

ABSTRACT

Information that is widely circulated can be obtained in various ways, one of which is by asking someone who already knows the information. However, problems arise when that person is not available, forcing individuals to obtain the information by reading official documents, which can be time-consuming. One such example is someone seeking to understand the contents of the academic guidelines of Universitas Tanjungpura. Academic guidelines are documents containing rules, procedures, and important information regarding an educational institution. Today's technology has advanced to address this issue, one solution being the development of a Question Answering System (QAS) that can answer questions based on factual information. This research implements Large Language Models (LLM) to develop a QAS using the LangChain framework with a retrieval-augmented generation (RAG) approach. LangChain simplifies the development process of applications that utilize LLMs. The implemented process includes reading PDF documents, splitting them into chunks, embedding the chunks, and storing them in a vector database using Pinecone. The LLMs used for testing include llama3:8B (4.7GB), deepseek-r1:8B (5.2GB), and phi-4:14B (9.1GB). System evaluation was conducted using ROUGE Score and BERTScore metrics. The best results were achieved by llama3:8B with a ROUGE-1 score of 0.33 (33%), ROUGE-2 score of 0.243 (24%), and BERTScore of 0.73 (73%), outperforming the other two models: deepseek-r1:8B with ROUGE-1 of 0.259 (25%), ROUGE-2 of 0.154 (15%), and BERTScore of 0.689 (68%); and phi-4:14B with ROUGE-1 of 0.28 (28%), ROUGE-2 of 0.208 (20%), and BERTScore of 0.722 (72%). Based on these evaluation results, llama3 was selected as the preferred LLM for the academic guideline QAS due to its superior performance in answering questions from the document. Additionally, a User Acceptance Testing (UAT) was conducted with 15 respondents, yielding an average score of 3.96 or 79.2%. This shows the conclusion that the QAS system is accepted by users because it can provide information easily, relevantly and usefully.

Keywords: Large Language Models (LLM), *Question Answering System (QAS), Langchain, retrieval-augmented generation, Pinecone,*Academic Guidelines Universitas Tanjungpura, ROUGE-Score, BERTScore, *Flask*, UAT

DAFTAR ISI

Halaman Pernyataan ii
Halaman Pengesahan iii
Halaman Persembahan iv
Kata Pengantar v
Abstrak vi
ABSTRACT vii
Daftar Isi viii
Daftar Gambar x
Daftar Tabel xi
Daftar Kode Program xii
Daftar Lampiran xiii
Daftar Notasi xiv

Bab I Pendahuluan 1

1.1 Latar Belakang 1
1.2 Perumusan Masalah 3
1.3 Tujuan Penelitian 4
1.4 Pembatasan Masalah 4
1.5 Sistematika Penulisan 5

Bab II Tinjauan Pustaka 6

2.1 Penelitian Terkait 6
2.2 Question Answering System 9
2.3 Large Language Models 10
2.4 Retrieval-Augmented Generation 11
2.5 Vector Database 12
2.6 LangChain 13
2.7 Pengembangan Pertanyaan 15
2.8 ROUGE – Score 16
2.9 BERTScore 16
2.10 Flask 17
2.11 *User Acceptance Testing* (UAT) 17

Bab III Metodologi Penelitian 18

3.1 Data Penelitian 18
3.2 Perangkat Penelitian 18
3.3 Arsitektur Sistem 19
3.4 Diagram Alir Penelitian 20
3.5 Persiapan Data 21
3.6 Embedding dan storage 21
3.7 Integrasi pada LLM dengan RAG dan Vector Database 22
3.8 Evaluasi Integrasi LLM dengan RAG 24

3.9 Implementasi Sistem Question Answering System 25
3.10 Pengujian Sistem dengan UAT 26

Bab IV Hasil dan PEMBAHASAN 27

4.1 Pemrosesan Data 27
4.2 Implementasi *Embbeddings* dan *Store Knowledge* 29
4.3 Integrasi Sistem LLMs dan RAG 32
4.4 Hasil Evaluasi ROUGE-Score dan BERT-Score 37
4.5 Hasil Implementasi LLM dengan QAS menggunakan Flask 39
4.6 Hasil Evaluasi Sistem dengan UAT 41

Bab V Kesimpulan dan Saran 44

5.1 Kesimpulan 44
5.2 Saran 45

Daftar Pustaka 46

Lampiran A Cover buku pedoman akademik untan 2023/2024 ..... A-1
Lampiran B List pertanyaan dan referensi jawaban B-1
Lampiran C Jawaban LLM llama3 ..... C-1
Lampiran D Jawaban LLM deepseek-r1 ..... D-1
Lampiran E Jawaban llm phi-4 ..... E-1
Lampiran F Sample record Chunks dan Embedding F-1
Lampiran G RESPONDEN UAT DAN JAWABAN F-1
Lampiran H DOKUMENTASI RESPONDEN UAT F-1

DAFTAR GAMBAR

Gambar 2.1 Arsitektur dari QAS sederhana 10
Gambar 2.2 Arsitektur dari RAG 12
Gambar 2.3 Proses Sederhana dari QAS.....15
Gambar 3.1 Arsitektur QAS dengan LangChain dan LLMs menggunakan metode RAG ……. 19
Gambar 3.2 Alur Penelitian Implementasi LLM dalam mengembangkan QAS menggunakan *framework LangChain*. 20
Gambar 3.3 Pemisahan dokumen ke bentuk sederhana dengan LangChain dengan kelas PyPDFLoader dan CharacterTextSplitter 21
Gambar 3. 4 *Embedding chunk* dengan model *embedding* dari *LangChain.Huggingface* 22
Gambar 3.5 *Store* data yang telah di *embedding*.....22
Gambar 3.6 Proses Integerasi LLM terhadap database dengan RAG.....23
Gambar 3.7 Platform Ollama 24
Gambar 4.1 *Chunks Document* ...……29
Gambar 4.2 Interface Pinecone 30
Gambar 4.3 API dari Pinecone.....31
Gambar 4.4 Proses pembuatan index untuk Pinecone 31
Gambar 4.5. Keluaran Jawaban dari LLM.....37
Gambar 4.6 Website Flask QAS pedoman Akademik Universitas Tanjungpura ..39
Gambar 4.7 Hasil Jawaban QAS Pertanyaan Sesuai Konteks 40
Gambar 4.8 Hasil Jawaban QAS Pertanyaan di Luar Konteks 40

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian Terkait……...……..7
Tabel 3.1 Tabel 10 Daftar Pertanyaan Yang Digunakan Sebagai Evaluasi……..24
Tabel 3.2 Daftar Kuesioner UAT QAS…….26
Tabel 4.1 Evaluasi ROUGE-Score dan BERTScore pada 3 LLM……38
Tabel 4.2 Hasil Kuesioner UAT……42

DAFTAR KODE PROGRAM

Kode Program 4.1. *Load* Data menggunakan fungsi *PyPDFLoader*.....27
Kode Program 4.2. Penggunaan *CharacterTextSplitter* untuk membuat *chunks*...27
Kode Program 4.3. Menerapkan Fungsi *text_splitter* Kepada Dokumen 28
Kode Program 4.4. Penggunaan *Embedding* dari *HuggingFace* 30
Kode Program 4.5. Import library dan penggunaan *PineconeVectorStore*.....32
Kode Program 4.6. Import library integrasi Sistem LLMs dan RAG.....33
Kode Program 4.7. Variabel *Embeddings* dan *Vectorstore* 33
Kode Program 4.8. *Template Prompt* untuk QAS 34
Kode Program 4.9. Membuat *PromptTempalte* sebagai input awal LLM.....35
Kode Program 4.10. Menggunakan Ollama sebagai pemilihan LLM lokal 35
Kode Program 4.11. Proses Mengambil Dokumen yang Relevan.....36
Kode Program 4.12. Menggabungkan Dokumen Relevan dengan Pertanyaan 36
Kode Program 4.13. Fungsi *invoke* untuk menjalankan LLM.....36

DAFTAR LAMPIRAN

LAMPIRAN A Cover buku pedoman akademik untan 2023/2024 ..... A-1
LAMPIRAN B List pertanyaan dan referensi jawaban B-1
LAMPIRAN C Jawaban LLM llama3 .....C-1
LAMPIRAN D Jawaban LLM deepseek-r1 ..... D-1
LAMPIRAN E Jawaban llm phi-4.....E-1
LAMPIRAN F Sample record Chunks dan Embedding F-1

DAFTAR NOTASI

BAB I
PENDAHULUAN

1.1 Latar Belakang

Teknologi yang sudah berkembang sangat pesat telah merubah tren penyebaran informasi yang mana dalam bentuk buku sekarang beralih ke dalam bentuk digital. Banyak sumber informasi yang telah menjadi inovasi teknologi dan diharapkan dapat memberikan informasi yang bermanfaat serta meningkatkan pengetahuan manusia melalui buku tersebut (Rashmi, 2015). Pedoman akademik merupakan dokumen yang berisi aturan, prosedur, dan informasi penting mengenai suatu institusi pendidikan. Dokumen ini bertujuan untuk mengatur seluruh kegiatan akademik agar berjalan dengan tertib dan sesuai dengan ketentuan yang berlaku. Di dalam buku tersebut terdapat informasi seperti visi dan misi institusi, sejarah, tata tertib, serta berbagai ketentuan akademik lainnya. Pedoman akademik Universitas Tanjungpura dapat diakses secara publik dan *online* pada website informatika untan di *https://informatika.untan.ac.id/dokumen_prodi/9*.

Informasi tersebut masih berupa teks sehingga apabila seseorang ingin memperoleh informasi tersebut, mereka harus membacanya dari awal dan membutuhkan waktu yang lebih lama dalam memperoleh informasi yang diinginkan. Salah satu sistem yang dapat membantu menyelesaikan dalam memperoleh informasi dari sekian banyaknya informasi adalah dengan membuat *Question Answering System* (QAS) yang dapat menjawab pertanyaan sesuai dengan informasi yang ada di dalam buku. Pedoman akademik dapat dimanfaatkan sebagai sumber utama dalam pengembangan QAS, yang memungkinkan pengguna memperoleh informasi akademik secara lebih mudah, cepat, dan interaktif. Dengan teknologi ini, seluruh pengguna dapat mengakses informasi secara efisien tanpa perlu membaca seluruh dokumen secara manual.

QAS merupakan sebuah teknologi yang diharapkan dapat memberikan informasi menjadi lebih interaktif dan sesuai dengan keinginan pengguna (Saha et al., 2024). QAS merupakan teknologi *Natural Language Processing* (NLP) telah mengalami perkembangan yang luar biasa dengan munculnya berbagai model

canggih seperti BERT dari Google dan GPT dari OpenAI dengan memanfaatkan arsitektur transformator dan telah dilatih dengan jumlah data yang sangat banyak sehingga disebut sebagai *Large Language Models* (LLM) (Zhao et al., 2024). LLM dapat menjadi solusi dalam pemilihan model untuk mengembangkan NLP.

LLM adalah model kecerdasan buatan berbasis *deep learning* yang dilatih dengan sejumlah besar data teks untuk memahami, menghasilkan, dan memproses bahasa manusia. LLM mampu melakukan berbagai tugas terkait bahasa, seperti menjawab pertanyaan, menerjemahkan teks, merangkum informasi, hingga membuat konten kreatif (James Thirunavukarasu et al., 2023). Banyak pengembang yang telah memberikan kemudahan untuk menggunakan LLM dalam mengembangkan aplikasi pribadi. Penggunaan LLM dapat digunakan dengan mengambil API dari LLM yang tersedia ataupun dengan menggunakan secara lokal tanpa API berbayar atau biasa disebut juga dengan *open source*. Pemilihan hal tersebut dapat disesuaikan dengan kebutuhan masing-masing.

Meskipun demikian, LLM masih memiliki keterbatasan contohnya adalah dengan menghasilkan output yang tidak sesuai fakta yang sedang terjadi atau output yang bersifat tidak masuk akal (disebut juga halusinasi) (Topsakal & Akinci, 2023). Mengembangkan LLM agar dapat mencapai solusi dan memberikan output yang diinginkan dapat dilakukan dengan memanfaatkan kerangka kerja yang bernama Langchain.

LangChain yang merupakan kerangka kerja *open source* hadir sebagai solusi yang bertujuan untuk menyederhanakan proses kompleks dalam mengembangkan aplikasi yang didukung oleh LLM(Pandya & Mahavidyalaya, 2023). LangChain membantu dalam mengembangkan aplikasi untuk memanfaatkan potensi besar dari LLM dalam menyelesaikan tugas-tugas seperti menganalisis dokumen, mengembangkan chatbot, menganalisis kode dan lain-lain (Singh et al., 2024). Penggunaan teknologi LangChain ditujukan agar model LLM dapat dikembangkan menjadi aplikasi sistem tanya jawab dan diharapkan mampu memberikan jawaban atas pertanyaan berdasarkan data-data yang dilatih sebelumnya seperti dari file *Portable Document Format* (PDF), CSV, JSON, dll. LangChain juga menyediakan beberapa kebutuhan yang diperlukan dalam

mengembangkan aplikasi berbasis LLM seperti *document parsing, agent tools, chatLLM*, dan lain-lain (Pandya & Mahavidyalaya, 2023).

Beberapa penelitian telah memanfaatkan teknologi LangChain dan pendekatan *Retrieval Augmented Generation* (RAG) untuk meningkatkan akurasi LLM dalam pengembangan QAS. Penelitian "A RAG-based Question Answering System Proposal for Understanding Islam: MufassirQAS LLM" oleh Ahmet Yusuf Alan (2024), mengembangkan MufassirQAS dengan memanfaatkan LLM dan RAG berbasis vector database untuk menjawab pertanyaan kompleks tentang Islam dengan referensi buku-buku terbuka di konteks Turki. Penelitian oleh O. Cárdenas, dkk.(2024), berjudul "Development of a ChatBot model for health telecare" mengintegrasikan LangChain, embeddings, Pinecone, dan GPT-3.5 untuk membangun chatbot di bidang kesehatan, dengan Pinecone sebagai vector database dan LangChain sebagai alat pemroses data. Sementara itu, penelitian oleh Suci Rahayu, dkk. (2024), dalam "Penerapan Teknologi LangChain pada Question Answering System Fikih Empat Madzhab" merancang sistem QAS menggunakan LangChain, FAISS sebagai vector database, dan diuji menggunakan BERT serta ROUGE, yang menunjukkan tingkat akurasi yang cukup tinggi.

Berdasarkan uraian yang telah dijelaskan diatas, maka penulis tertarik untuk melakukan penelitian yang berjudul " Implementasi Large Language Models Menggunakan Framework Langchain Pada Question Answering System." dengan fokus dalam memanfaatkan teknologi kecerdasan buatan terbaru yaitu LLM dan dalam penelitian ini dapat menghasilkan model yang bisa memberikan informasi dengan tanya jawab.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka rumusan masalah dalam penelitian ini dapat disimpulkan sebagai berikut.

1. Apakah implementasi Large Language Models (LLM) dengan bantuan framework LangChain dan metode RAG dapat mempermudah

pengembangan Question Answering System (QAS) dan meningkatkan kemudahan serta akurasi dalam memperoleh informasi?

2. Apakah kinerja QAS yang dikembangkan dapat dievaluasi berdasarkan akurasi dan relevansi jawabannya terhadap pertanyaan pengguna?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditentukan maka penelitian ini bertujuan untuk mengembangkan sistem Question Answering System (QAS) dengan mengimplementasikan Large Language Model (LLM) menggunakan framework LangChain untuk mempermudah pencarian informasi dari dokumen Pedoman Akademik Universitas Tanjungpura. Dalam prosesnya, LangChain dimanfaatkan sebagai kerangka kerja untuk memproses data, mengintegrasikan arsitektur Retrieval-Augmented Generation (RAG), serta mendukung penggunaan LLM agar pengembangan QAS menjadi lebih efisien dan akurat. Kemudian dilakukan evaluasi terhadap kinerja QAS yang dikembangkan dalam memberikan jawaban yang akurat dan relevan terhadap pertanyaan pengguna.

1.4 Pembatasan Masalah

Beberapa hal yang menjadi batasan dalam penelitian ini adalah sebagai berikut.

1. Framework yang digunakan adalah *framework LangChain*
2. LLM yang akan digunakan berupa open source LLM yang diunduh di Ollama yaitu llama3:8b (4,7 GB) , deepseek-r1:8b (5,2 GB), dan phi4:14b (9,1 GB) karena keterbatasan *hardware*.
3. Metode yang digunakan dalam pengembangan LLM adalah *Retrieval Augmented Generation* (RAG)
4. *File pdf* yang digunakan adalah Buku Pedoman Akademik Universitas Tanjungpura.

5. Metrik evaluasi yang digunakan adalah metrik evaluasi ROUGE-Score dan BERTScore

6. Implementasi QAS menggunakan aplikasi *Flask.*

7. Pengujian implementasi QAS menggunakan *User Acceptance Testing* (UAT)

1.5 Sistematika Penulisan

Sistematika dari penulisan skripsi terdiri dari lima bab yaitu, Bab 1 Pendahuluan, Bab II Tinjauan Pustaka, Bab III Metodologi Penelitian, Bab IV Hasil dan Pembahasan, dan Bab V Penutup.