Smart Chatbot for Undergraduate Registry Services, Chiang Mai University

Author: Kungwansup Saelee
Advisor: Asst. Prof. Dr. Dussadee Praserttitipong

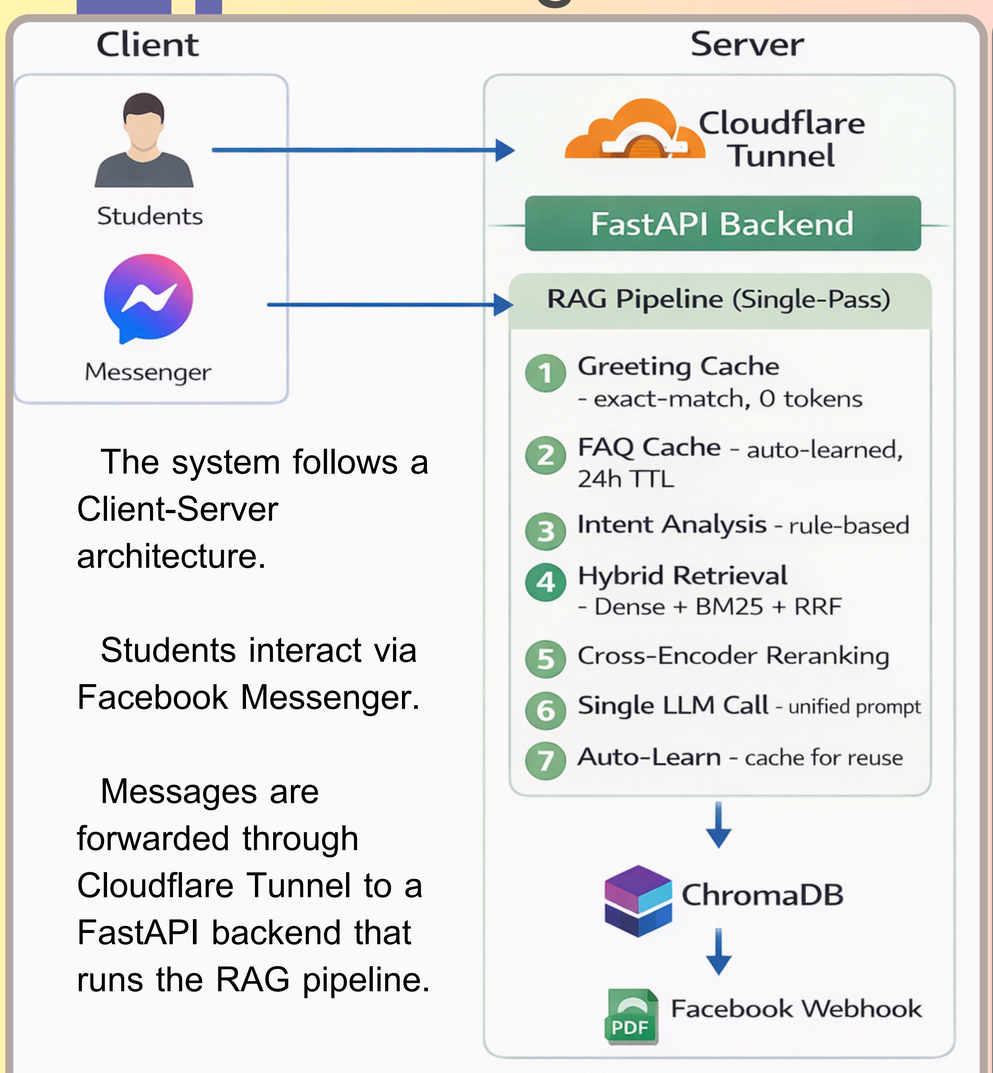
Abstract

This study proposes a Smart Chatbot for Undergraduate Registry Services at Chiang Mai University using a Single-Pass Retrieval-Augmented Generation (RAG) architecture requiring only one LLM call per query. The system integrates Hybrid Retrieval (Dense Retrieval + BM25 with Thai tokenization), Weighted Reciprocal Rank Fusion (0.7/0.3), Cross-Encoder reranking, and a 2-Tier Cache mechanism. Evaluation with 113 test questions across 16 categories achieved 93.8% overall accuracy. FAQ caching reduced latency by approximately 80% for repeated queries.

Introduction

Undergraduate registry services at Chiang Mai University rely heavily on staff to respond to repetitive questions regarding course registration, withdrawal deadlines, tuition payment, and academic documents, and this manual process creates bottlenecks during peak registration periods while limiting 24/7 accessibility. The project aims to develop an AI chatbot that answers registry queries via Facebook Messenger, ensures responses are grounded in official university documents, supports both Thai and English queries, and complies with PDPA. The proposed solution is a Single-Pass RAG-based chatbot integrating Hybrid Search, Cross-Encoder Reranking, and 2-Tier Caching, requiring only one LLM call per query.

Design

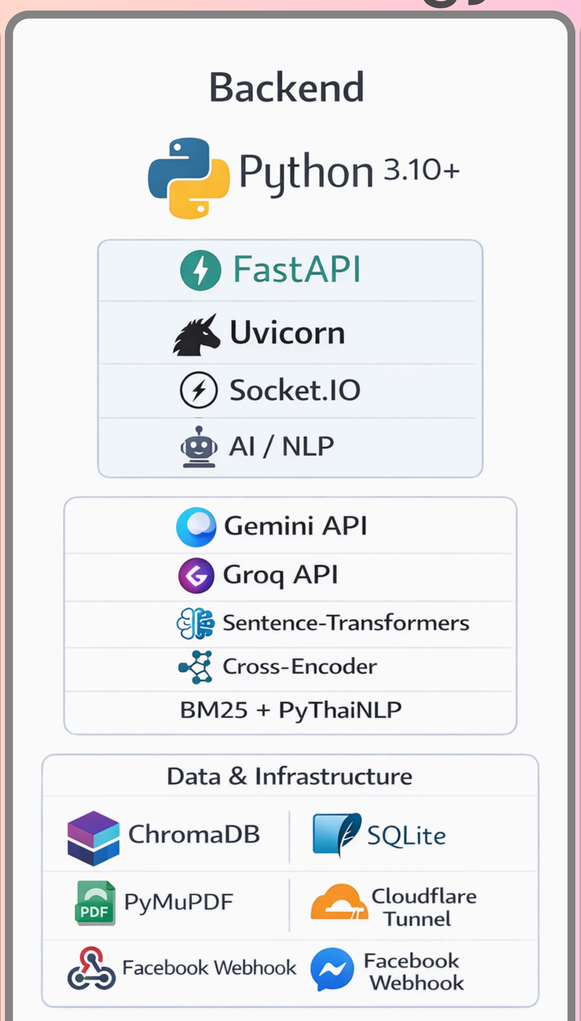


The system follows a Client-Server architecture.

Students interact via Facebook Messenger.

Messages are forwarded through Cloudflare Tunnel to a FastAPI backend that runs the RAG pipeline.

Technology

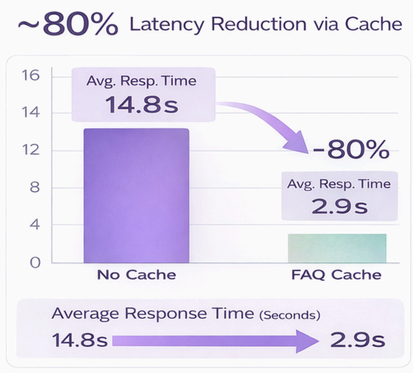


Results

- Round 1 covers calendar, payment, withdrawal, enrollment, greeting, and English queries.
- Round 2 includes spoken/informal Thai, typos, ambiguous inputs, off-topic, prompt injection, and multi-topic questions.
- Round 2 avg. response time was 7.47s (complex queries exceeding 5s target).

Test Round	Questions	Correct	Accuracy
Round 1 — Standard registry	59	58	**98.3%**
Round 2 — Challenging & adversarial	54	48	**88.9%**
Overall (16 categories)	**113**	**106**	**93.8%**





Conclusion

The proposed chatbot successfully answers undergraduate registry questions through Facebook Messenger using a Single-Pass RAG architecture that requires only one LLM call per question; testing with 113 questions across 16 categories achieved 93.8% overall accuracy (98.3% on standard questions and 88.9% on challenging ones), while the 2-Tier caching mechanism reduced latency by approximately 80% for repeated questions. The main weakness lies in very short or ambiguous inputs (1–2 words), where the retriever lacks sufficient context, resulting in 70.0% accuracy, and the Round 2 average response time of 7.47 seconds exceeded the 5-second target due to query complexity. Future work includes improving synonym mapping for informal or spoken Thai queries, adding additional communication channels such as LINE OA and a web interface, and containerizing the system with Docker for production deployment.

References

[1] Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274.
[2] Gupta, A., & Agarwal, P. (2021). AI chatbot platform for university educational guidance.
[3] LINE Chatbot for research service support. (2022). Rajamangala University of Technology Thanyaburi.

204499 Independent Study, Academic Year 2025
Department of Computer Science, Faculty of Science, Chiang Mai University