

บทสรุปผู้บริหาร

โครงการนี้เป็นการออกแบบและพัฒนาระบบผู้ช่วยตอบคำถาม Chatbot สำหรับสาขาวิชาวิศวกรรมไฟฟ้า เพื่อแก้ไขปัญหาการเข้าถึงข้อมูลของภาควิชาที่มีอยู่หลายแหล่ง เช่น ข้อมูลหลักสูตร รายวิชา ข้อมูลบุคลากร และข้อมูลประชาสัมพันธ์ ซึ่งอาจทำให้ผู้ใช้งานใช้เวลาในการค้นหาข้อมูลและได้รับข้อมูลที่ไม่สะดวกต่อการใช้งาน ระบบนี้จึงถูกพัฒนาขึ้นเพื่อช่วยให้ผู้ใช้งานสามารถค้นหาและเข้าถึงข้อมูลได้อย่างรวดเร็ว ถูกต้อง และมีประสิทธิภาพมากขึ้น

ระบบที่พัฒนาขึ้นประยุกต์ใช้เทคนิค Retrieval-Augmented Generation (RAG) ร่วมกับ Large Language Model (LLM) เพื่อให้ระบบสามารถค้นคืนข้อมูลจากฐานความรู้และนำข้อมูลที่เกี่ยวข้องมาใช้ในการสร้างคำตอบ โดยใช้ n8n Workflow Automation เป็นเครื่องมือควบคุมกระบวนการทำงาน และใช้ Pinecone Vector Database สำหรับจัดเก็บข้อมูลในรูปแบบเวกเตอร์ ระบบแบ่งการทำงานออกเป็น 2 ส่วนหลัก ได้แก่ กระบวนการจัดเตรียมข้อมูล (Ingestion Pipeline) สำหรับรวบรวมและจัดการข้อมูลจากเอกสารต้นทาง และกระบวนการตอบคำถาม (Query Pipeline) สำหรับรับคำถาม ค้นคืนข้อมูล และสร้างคำตอบพร้อมแหล่งอ้างอิง

ระบบถูกออกแบบให้มีความสามารถในการตรวจสอบความปลอดภัยของข้อความ การคัดกรองประเภทคำถาม และการประมวลผลข้อมูลเอกสาร PDF เพื่อเพิ่มความถูกต้องและความปลอดภัยของระบบ โดยมีการประเมินประสิทธิภาพด้วยกรอบการประเมิน RAGAS ซึ่งกำหนดเป้าหมายคะแนนความถูกต้องไม่น้อยกว่าร้อยละ 50 รวมถึงพิจารณาด้านระยะเวลาการตอบสนอง การจัดการข้อมูล และต้นทุนการพัฒนา เพื่อให้ได้ระบบ Chatbot ที่สามารถช่วยลดภาระงานของบุคลากร และเพิ่มประสิทธิภาพการให้บริการข้อมูลของสาขาวิชาอย่างสูงสุด

ปิยพนธ์ บุญโสภิต
วชิรพล พลารักษ์
วิชญ์ หมั่นนิยม
วีรวุฒิ อาจใจ

EXECUTIVE SUMMARY

This project involves the design and development of a Chatbot question-answering system for the Program in Electrical Engineering to resolve issues regarding access to program information stored across multiple sources, such as curricula, course information, personnel details, and public relations announcements, which can cause users to spend time searching for information and receive inconvenient service. This system was therefore developed to help users search for and access information more quickly, accurately, and efficiently.

The developed system applies Retrieval-Augmented Generation (RAG) techniques in combination with Large Language Models (LLMs) to retrieve information from a knowledge base and use relevant data to generate answers. It utilizes n8n Workflow Automation as a workflow control tool and Pinecone Vector Database for storing data in vector format. The system architecture is divided into two main parts: the Ingestion Pipeline for gathering and managing data from source documents, and the Query Pipeline for receiving queries, retrieving information, and generating answers with reference sources.

The system is designed with capabilities for message safety verification, query classification, and PDF document processing to enhance accuracy and security. Performance is evaluated using the RAGAS evaluation framework with a target accuracy score of no less than 50%, alongside considerations of response time, data management, and development cost, to deliver a Chatbot system that reduces personnel workload and maximizes the efficiency of program information delivery.

Piyaphon Boonsopa

Wachiraphon Phalarak

Witsanu Manniyom

Wirawut Atchai