

KẾ HOẠCH THỰC HIỆN MÔN DỰ ÁN CNTT- HK1/2024-2025

THÔNG TIN NHÓM SINH VIÊN

Mã nhóm

Tên sinh viên 1 (nhóm trưởng): Hoàng Đình Quý Vũ

MSSV: 521H0517

Số ĐT: 0868245465

Email: 521h0517@student.tdtu.edu.vn

Tên sinh viên 2: Trần Quốc An

MSSV: 521H0385

Số ĐT: 0383474552

Email: 521h0385@student.tdtu.edu.vn

THÔNG TIN ĐỀ TÀI

GV hướng dẫn: PGS.TS Lê Anh Cường

Số ĐT: 0902134662

Email: leanhcuong@tdtu.edu.vn

Mã đề tài: ĐAA005

Tên đề tài: Building a virtual assistant to support admissions information using RAG techniques

Ngày bắt đầu: 2/11/2024

Tuần	Từ	Đến	Công việc cần làm	Công việc đã làm	% hoàn thành
1	2/11/2024	16/11/2024	Tìm kiếm dữ liệu tuyển sinh của các trường đại học tại TP HCM, kết hợp việc tìm hiểu các kỹ thuật cần thiết trong việc xây dựng dự án	Tìm kiếm dữ liệu tuyển sinh của 15 trường đại học tại TP HCM. Tìm hiểu kỹ thuật RAG và prompt engineering	100%
2	17/11/2024	20/11/2024	Thực hiện áp dụng kỹ thuật RAG với bộ dữ liệu tuyển sinh	Thực hiện áp dụng kỹ thuật RAG với bộ dữ liệu tuyển sinh	90%
3	21/11/2024	21/11/2024	BÁO CÁO TIẾN ĐỘ		
4	23/11/2024	26/11/2024	Thực hiện kỹ thuật RAG với kỹ thuật prompt engineering nhằm tăng cường kết quả	Thực hiện kỹ thuật RAG với kỹ thuật prompt engineering nhằm tăng cường kết quả	90%
5	27/11/2024	30/11/2024	Thực hiện các kỹ thuật nhằm tăng cường kết quả truy vấn: query router, query transformation, hybrid search, semantic chunking. Sử dụng mô hình gemini-1.5-flash-8b	Thực hiện các kỹ thuật nhằm tăng cường kết quả truy vấn: query router, query transformation, hybrid search, semantic chunking	100%
6	1/12/2024	1/12/2024	BÁO CÁO TIẾN ĐỘ		

7	2/12/2024	7/12/2024	Ứng dụng kỹ thuật Advanced RAG để tăng cường hệ thống qua việc tăng độ hiệu quả của truy vấn so với naïve RAG, ứng dụng kỹ thuật few-shot prompt engineering. Sử dụng mô hình gemini-1.5-pro thay gemini-1.5-flash cho việc điều hướng và xử lý truy vấn từ người dùng, mô hình gemma-2-9b-it cho việc sinh câu trả lời dựa trên truy vấn và tài liệu	Ứng dụng kỹ thuật Advanced RAG để tăng cường hệ thống qua việc tăng độ hiệu quả của truy vấn so với naïve RAG, ứng dụng kỹ thuật few-shot prompt engineering. Sử dụng mô hình gemini-1.5-pro thay gemini-1.5-flash cho việc điều hướng và xử lý truy vấn từ người dùng, mô hình gemma-2-9b-it cho việc sinh câu trả lời dựa trên truy vấn và tài liệu	80%
8	8/12/2024	13/12/2024	Tiếp tục cải thiện hệ thống, cải thiện prompt. Xây dựng bộ dữ liệu điểm chuẩn các ngành lưu vào database, hoàn thiện luồng chạy cho hệ thống (RAG retrieval, SQL retrieval, Websearch retrieval).tiến hành xây dựng giao diện dựa vào streamlit	Tiếp tục cải thiện hệ thống, cải thiện prompt. Xây dựng bộ dữ liệu điểm chuẩn các ngành lưu vào database, hoàn thiện luồng chạy cho hệ thống (RAG retrieval, SQL retrieval, Websearch retrieval).tiến hành xây dựng giao diện dựa vào streamlit	100%
9	13/12/2024	16/20/1900	Tiếp tục thử nghiệm hệ thống, tiến hành xử lý các truy vấn dạng so sánh (giải quyết multi query). Xây dựng bộ dữ liệu groundtruth nhằm đo lường hiệu năng của hệ thống, đánh giá kết quả truy vấn	Tiếp tục thử nghiệm hệ thống, tiến hành xử lý các truy vấn dạng so sánh (giải quyết multi query). Xây dựng bộ dữ liệu groundtruth nhằm đo lường hiệu năng của hệ thống, đánh giá kết quả truy vấn	90%
	16/12/2024	20/12/2024	Tiến hành đánh giá hệ thống, đánh giá hệ thống với human evaluation	Tiến hành đánh giá hệ thống, đánh giá hệ thống với human evaluation	100%
	21/12/2024	23/12/2024	Hoàn thành bộ đánh giá kết quả, và thử nghiệm hệ thống	Hoàn thành bộ đánh giá kết quả, và thử nghiệm hệ thống	100%
	24/12/2024	4/2/2025	Nghỉ tết		
	5/2/2025	10/2/2025	Hoàn thành báo cáo và nộp kết quả	Hoàn thành báo cáo và nộp kết quả	100%

XÁC NHẬN CỦA GIẢNG VIÊN HƯỚNG DẪN

A handwritten signature in black ink, appearing to read 'Cuong', is written over a horizontal line. The signature is stylized with a large 'C' and a long horizontal stroke extending to the right.