

BÁO CÁO TIẾN ĐỘ DỰ ÁN HỎI ĐÁP TỰ ĐỘNG LỊCH SỬ - DU LỊCH

Thành viên thực hiện: Nguyễn Vĩnh Hưng*, Nguyễn Hòa An[†]

Giảng viên hướng dẫn: Trần Thanh Phước[‡]

Tuần 7 – Giai đoạn 2 – Ngày 26/8/2025

Tóm tắt nội dung

Trong tuần này, nhóm tiếp tục triển khai mô hình RAG + KG, lần này sử dụng Phi-3-mini-4k-instruct để thay thế Mistral-7B do giới hạn tài nguyên. Đồng thời, nhóm đã chỉnh sửa và bổ sung các phần còn thiếu trong báo cáo, dẫn nguồn tham khảo, hình ảnh, bảng biểu đúng chuẩn.

1 Tổng quan

Trong tuần, nhóm tập trung thực hiện hai nhiệm vụ chính:

1. Xây dựng và triển khai mô hình pretrained RAG + KG dựa trên Phi-3-mini-4k-instruct.
2. Chỉnh sửa báo cáo: bổ sung mô tả chi tiết pipeline RAG, dẫn hình ảnh minh họa vào phần văn bản.

2 Tiến độ công việc

2.1 Xây dựng mô hình pretrained RAG kết hợp KG

ử dụng Phi-3-mini-4k-instruct (3.8B tham số, context 4096 token) thay thế Mistral-7B để phù hợp tài nguyên Kaggle T4 (16GB).

Mô tả pipeline chi tiết:

*Email: 52200097@student.tdtu.edu.vn

[†]Email: 52200182@student.tdtu.edu.vn

[‡]Email: tranthanhphuoc@tdtu.edu.vn.

- Chunking tài liệu lịch sử (500–800 token).
- Embedding: dangvantuan/vietnamese-embedding.
- Retriever: ChromaDB (vector) + Neo4j (graph query bằng Cypher)

2.2 Chỉnh sửa nội dung báo cáo

- Bổ sung mô tả chi tiết cho phần trích xuất JSON bằng ChatGPT, bao gồm kiểm thử tự động, xử lý lỗi cú pháp và bước xác nhận của tester.
- Chỉnh sửa lại đảm bảo khi nhắc đến hình minh họa đều gắn tham chiếu

2.3 Chuẩn hóa tài liệu tham khảo

- Đảm bảo hình và bảng đều dẫn nguồn rõ ràng trong caption.

3 Kế hoạch cho tuần tiếp theo

Trong tuần tới, nhóm sẽ:

- Sửa lỗi model embedding bị lỗi không thể mã hóa các đoạn chunk.