Trưởng nhóm : Nguyễn Quốc Nam

Tên đề tài : Nghiên cứu về image captioning và ứng dụng

Tóm tắt nội dụng :  Tìm hiểu về các khái niệm cơ bản và công nghệ liên quan như Computer Vision và Natural Language Processing. Tiếp theo, tìm hiểu các phương pháp và mô hình chính như Encoder-Decoder, Attention Mechanism và Transformers để cải thiện hiệu suất mô hình. Việc sử dụng các bộ dữ liệu chuẩn như MS COCO, Flickr8k và Flickr30k là cần thiết, cùng với các bước tiền xử lý dữ liệu như cắt, xoay và chuẩn hóa ảnh. Trong quá trình huấn luyện, áp dụng các kỹ thuật học sâu và đánh giá mô hình bằng các chỉ số như BLEU, CIDEr, SPICE, METEOR và ROUGE-L.

Để triển khai mô hình, bạn cần chuẩn bị môi trường với Python và các thư viện như TensorFlow, PyTorch, và Flask. Tiếp theo, triển khai mô hình bằng cách tạo API với Flask hoặc sử dụng BentoML để đơn giản hóa quá trình đóng gói và triển khai trên các nền tảng serverless. Cuối cùng, kiểm thử và tối ưu hóa mô hình để đảm bảo hiệu suất, đồng thời giám sát và bảo trì để duy trì hiệu quả hoạt động.

Mục tiêu : Mục tiêu của đề tài Image Captioning là phát triển một mô hình trí tuệ nhân tạo có khả năng tự động sinh ra câu mô tả văn bản cho hình ảnh bằng tiếng việt , kết hợp giữa xử lý hình ảnh (CNN) và xử lý ngôn ngữ tự nhiên (LSTM hoặc Transformer). Đề tài hướng tới việc ứng dụng trong thực tiễn như trợ giúp người khiếm thị, tự động hóa phân tích nội dung hình ảnh, và tối ưu hóa tìm kiếm ảnh, đồng thời cung cấp nền tảng cho các ứng dụng AI sáng tạo hơn trong tương lai.

Kết quả dự kiến : Kết quả dự kiến của đề tài Image Captioning là xây dựng thành công mô hình AI có khả năng tự động sinh câu mô tả hình ảnh với độ chính xác cao, kèm theo mô tả chi tiết về nội dung hình ảnh. Một phần mềm hoàn chỉnh sẽ được phát triển, tích hợp nhiều mô hình khác nhau như CNN-LSTM, Transformer hoặc các mô hình tiên tiến khác để nâng cao độ chính xác và chất lượng chú thích. Phần mềm sẽ có giao diện thân thiện, được huấn luyện và kiểm thử trên các bộ dữ liệu thực tế, kèm theo báo cáo khoa học chi tiết cùng các đề xuất cải tiến và hướng phát triển trong tương lai.