

CHƯƠNG 5: TỔNG KẾT VÀ ĐÁNH GIÁ

5.1. KIẾN THỨC ĐẠT ĐƯỢC

Trong thời gian thực hiện khoá luận tốt nghiệp, chúng em đã được cung cấp thêm nhiều điều mới, kiến thức mới:

- Có những kiến thức tổng quan về dịch máy và các kỹ thuật trong mạng nơ-ron để dịch máy.
- Có được các kiến thức và kinh nghiệm trong việc thiết kế, xây dựng, triển khai và kiểm thử khi thực hiện phát triển một hệ thống cung cấp dịch vụ.
- Biết được quá trình phát triển của dịch máy.
- Học hỏi các quy trình phát triển dự án phần mềm như Kanban, Waterfall, Scrum và có cơ hội được áp dụng vào luận văn.
- Nâng cao khả năng làm việc nhóm và giao tiếp của các thành viên trở nên tốt hơn.
- Khả năng tìm kiếm, đọc tài liệu và sách báo nâng cao. Hình thành được các thói quen như trích dẫn tài liệu. Khả năng tổng hợp các kiến thức từ nhiều nguồn khác nhau được nâng cao.
- Nâng cao được khả năng lên kế hoạch và đảm bảo hoàn thành trong khoảng thời gian cho trước.
- Hình thành kỹ năng tự chủ, tự tập và tinh thần trách nhiệm công việc.
- Học hỏi được cách trình bày, viết tài liệu một cách hợp lý và đẹp mắt.
- Có thêm kinh nghiệm đọc hiểu, hiệu chỉnh từ mã nguồn đã được phát triển. Khả năng chỉnh sửa và khắc phục khi gặp lỗi được nâng cao.

5.2. SẢN PHẨM THU ĐƯỢC

5.2.1. Môi trường phát triển

- Hệ điều hành: Ubutu, MacOS và Windows
- Công cụ phát triển phần mềm: Visual Code, Jupyter notebook và Google Colab
- Các thư viện/nền tảng được sử dụng:

Tên thư viện / nền tảng	Tóm tắt chức năng
Python	Python là một ngôn ngữ lập trình thông dịch được sử dụng rất phổ biến trong lĩnh vực khoa học máy tính nhờ những ưu điểm như: Đa nền tảng, đơn giản, mã nguồn mở, nhiều thư viện hỗ trợ.
Flask	Thư viện Tensorflow được sử dụng trong việc tính toán các biểu đồ và các dữ liệu dưới dạng số hoá trong sản phẩm khoá luận. Là một thư viện mã nguồn mở hỗ trợ mạnh mẽ các phép toán học để tính toán trong máy học.

Matplotlib	Nó là một thư viện vẽ đồ thị rất mạnh mẽ, giúp ta có cái nhìn trực quan hơn về dữ liệu.
Numpy	NumPy là một thư viện cho ngôn ngữ lập trình Python, thêm hỗ trợ cho các mảng và ma trận lớn, đa chiều, cùng với một tập hợp lớn các hàm toán học cấp cao để hoạt động trên các mảng này.
Pandas	Pandas dùng để thao tác và phân tích dữ liệu. Cụ thể, nó cung cấp các cấu trúc dữ liệu và các thao tác để thao tác các bảng số và chuỗi thời gian.
Scipy	Scipy chứa các mô-đun để tối ưu hóa, đại số tuyến tính, tích hợp, nội suy, các chức năng đặc biệt, FFT, xử lý tín hiệu và hình ảnh, bộ giải ODE và các nhiệm vụ phổ biến khác trong khoa học và kỹ thuật.
Scikit-learn	Scikit-learn được thiết kế dựa trên nền Numpy và Scipy. Scikit-learn chứa hầu hết các thuật toán

	machine learning hiện đại nhất, đi kèm với tài liệu và luôn được cập nhật.
Tensorflow	Thư viện Tensorflow là thư viện mã nguồn mở dùng cho tính toán số học sử dụng đồ thị luồng dữ liệu.
Keras	Keras là một thư viện mạng nơ-ron mã nguồn mở được viết bằng Python. Được thiết kế để cho phép thử nghiệm nhanh với các mạng thần kinh sâu, nó tập trung vào việc thân thiện với người dùng, mô-đun và mở rộng.

5.2.2. Môi trường triển khai

- Nền tảng đám mây Amazon EC2
- Truy cập: các trình duyệt có kết nối mạng.

5.2.3. Chức năng đã cài đặt

- Cung cấp dịch vụ nhận vào văn bản tiếng Anh dạng text và trả về đối tượng text chứa văn bản tiếng Việt tương ứng.

5.2.4. Chức năng chưa cài đặt

5.2.5. So sánh các kết quả thu được với mục tiêu ban đầu

Mục tiêu ban đầu	Nhận xét mức độ hoàn thành
Trình bày lý do xây dựng mô hình dịch máy	Đã trình bày các lý do ở chương 1 của luận văn
Trình bày lý thuyết nền tảng và giải pháp để xử lý việc dịch một văn bản từ tiếng Anh sang tiếng Việt.	Đã trình bày ở chương 2
Xây dựng, thu thập dữ liệu và đào tạo mô hình để dịch một văn bản từ tiếng Anh sang tiếng Việt.	
Xây dựng một trang web demo việc sử dụng mô hình để dịch một văn bản từ tiếng Anh sang tiếng Việt.	
Viết 120 trang luận văn theo đúng chuẩn yêu cầu và trích dẫn các tài liệu tham khảo đầy đủ.	Luận văn được viết tương đối đầy đủ và chính xác.

5.3. KẾT QUẢ MÔ HÌNH HUẤN LUYỆN

(-----)

5.4. ĐỊNH HƯỚNG PHÁT TRIỂN VÀ NGHIÊN CỨU TRONG TƯƠNG LAI

- Cải thiện lại mã nguồn để dịch được chính xác và hợp lý hơn.
- Nghiên cứu kĩ hơn về lý thuyết nền tảng, từ đó có các bước cải thiện và thực hiện chức năng một cách đúng đắn.
- Chỉnh sửa các tài liệu nghiên cứu, hướng dẫn như luận văn, hướng dẫn sử dụng để giúp người dùng mau chóng nắm bắt được vấn đề.
- Hoàn thiện chức năng dịch văn bản, sửa một số lỗi còn tồn tại. Hoặc đổi phương pháp xây dựng mô hình để chuẩn xác hơn.
- Cải thiện tốc độ xử lý các tác vụ của ứng dụng, giúp ứng dụng chạy mượt mà và tạo trải nghiệm tốt hơn cho người dùng.
- Trình bày, chỉnh sửa mã nguồn theo khuôn mẫu để dễ dàng bảo trì và chỉnh sửa trong tương lai.
- Tích hợp tốt cho các dòng máy sử dụng hệ điều hành Android để cho điện thoại nào cũng có thể chạy ứng dụng một cách mượt mà nhất.

5.5. LỜI KẾT

Luận văn “Xây dựng mô hình dịch máy từ tiếng Anh sang tiếng Việt”, hệ thống cung cấp dịch vụ và ứng dụng được xây dựng là sản phẩm kết tinh của một quá trình học tập, làm việc và nghiên cứu nghiêm túc của nhóm sinh viên. Tuy hệ thống còn nhiều hạn chế về hệ thống lẫn khả năng xử lý của nó, song sản phẩm hệ thống cung cấp dịch vụ dịch máy từ tiếng Anh sang tiếng Việt đã đem lại cho nhóm sinh viên những kiến thức và kinh nghiệm quý báu cũng như cách để triển khai các dự án thực tế trong tương lai. Các hệ thống học sâu nói chung và dịch máy nói riêng hiện đang là những lĩnh vực nổi trội trên thế giới và nó đem lại lợi ích tuyệt vời trong cuộc sống. Những trải nghiệm trong luận văn là những kinh nghiệm quý báu cho chúng em để có những kiến thức và tiếp tục nghiên cứu và phát triển sự nghiệp của bản thân.