

**TRƯỜNG ĐẠI HỌC NAM CẦN THƠ KHOA CÔNG NGHỆ THÔNG TIN**

BÁO CÁO THỰC TẬP

Họ và tên sinh viên: Nguyễn Tuấn Đạt Mã số SV: 219486

Lớp: DH21KMT01 Ngành: Khoa học máy tính

Họ và tên giảng viên hướng dẫn: Đặng Mạnh Huy

Họ và tên cán bộ hướng dẫn tại nơi thực tập: Võ Khương Duy

**Cần Thơ, 19 tháng 05 năm 2025**

**TRƯỜNG ĐẠI HỌC NAM CẦN THƠ**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO THỰC TẬP**

Họ và tên sinh viên: Nguyễn Tuấn Đạt Mã số SV: 219486

Lớp: DH21KMT01 Ngành: Khoa học máy tính

Họ và tên giảng viên hướng dẫn : Đặng Mạnh Huy

Họ và tên cán bộ hướng dẫn tại nơi thực tập : Võ Khương Duy

**LỜI CẢM ƠN**

Trong quá trình thực hiện đề tài “Hệ thống điền biểu mẫu tự động dựa trên lịch sử”, em đã nhận đƣợc rất nhiều sự quan tâm, hướng dẫn và giúp đỡ quý báu từ các thầy cô, bạn bè, gia đình và đơn vị thực tập. Em xin trân trọng gửi lời cảm ơn sâu sắc đến:

Trước hết, em xin gửi lời cảm ơn chân thành và sâu sắc nhất đến thầy Đặng Mạnh Huy – giảng viên hướng dẫn đã tận tình giúp đỡ. Nhờ sự hướng dẫn tận tâm, những góp ý quý báu và sự đồng hành của thầy, em đã có thể từng bước hoàn thiện đề tài thực tập một cách hiệu quả. Bên cạnh kiến thức chuyên môn, em còn học hỏi được nhiều kinh nghiệm thực tiễn và kỹ năng làm việc quan trọng cho con đường phát triển sau này. Em xin trân trọng cảm ơn thầy vì sự hỗ trợ và tạo điều kiện thuận lợi trong suốt thời gian thực hiện báo cáo này.

Em cũng xin gửi lời cảm ơn đến quý Thầy Cô trong khoa Công nghệ Thông tin – Trường Đại học Nam Cần Thơ, những người đã truyền đạt cho em những kiến thức nền tảng và chuyên môn quý báu trong suốt quá trình học tập, giúp em có đủ năng lực để áp dụng và phát triển đề tài một cách hiệu quả.

Bên cạnh đó, em xin chân thành cảm ơn Công Ty TNHH Một Thành Viên Công Nghệ Kỹ Thuật Tiên Phong – nơi em có cơ hội thực tập và tiếp cận với môi trƣờng làm việc thực tế. Em đặc biệt cảm ơn anh Võ Khương Duy – cán bộ hướng dẫn thực tập tại công ty, đã tạo điều kiện thuận lợi và hỗ trợ em trong suốt quá trình hoàn thành đề tài.

Cuối cùng, em xin kính chúc thầy luôn mạnh khỏe, hạnh phúc và thành công trong công tác giảng dạy và nghiên cứu.

Em chân thành cảm ơn!

Sinh viên thực hiện

**PHIẾU ĐÁNH GIÁ BÁO CÁO KẾT QUẢ THỰC TẬP**

**Học kỳ III năm học 2024 - 2025**

**(Dùng cho giảng viên phản biện – tối đa 6 điểm)**

Họ và tên giảng viên Phản biện: Huỳnh Võ Hữu Trí

Họ tên sinh viên thực tập: Nguyễn Tuấn Đạt Mã số SV: 219486

Lớp: DH21KMT01 Ngành: Khoa học máy tính

|  |  |  |
| --- | --- | --- |
| **Nội dung đánh giá** | **Điểm tối đa** | **Điểm chấm** |
| **I. Hình thức trình bày** | **0.5** |  |
| I.1 Đúng format của Khoa: Trang bìa, trang lời cảm ơn, 02 trang đánh giá thực tập của GVHD và GVPB (mẫu M-TT-04), trang tổng hợp điểm (M-TT-05), trang mục lục và các nội dung báo cáo. Sử dụng đúng mã và font tiếng Việt (Unicode Times New  Roman, Size 13) | 0.25 |  |
| I.2 Trình bày mạch lạc, súc tích, không có lỗi chính tả | 0.25 |  |
| **II. Phiếu theo dõi** | **0.5** |  |
| II.1 Có lịch làm việc đầy đủ cho 8 tuần | 0.25 |  |
| II.2 Số buổi thực tập tại cơ quan trong 1 tuần >=6; ít hơn 6 buổi  0.0 điểm | 0.25 |  |
| **III. Nội dung thực tập (quyển báo cáo)** | **3.0** |  |
| - Có được sự hiểu biết tốt về cơ quan nơi thực tập. | 0.5 |  |
| - Phương pháp thực hiện phù hợp với nội dung công việc được | 0.5 |
| giao. |  |
| - Kết quả củng cố lý thuyết. | 0.5 |
| - Kết quả rèn luyện kỹ năng thực hành. | 0.5 |
| - Kinh nghiệm thực tiễn thu nhận được. | 0.5 |
| - Kết quả công việc có đóng góp cho cơ quan nơi thực tập | 0.5 |
| **IV. Kết quả chất vấn sinh viên** | **2.0** |  |
| **TỔNG CỘNG** | **6.0** |  |
| Điểm trừ |  |  |
| **Điểm còn lại** |  |  |

Cần Thơ, ngày….tháng….năm 20….

GIẢNG VIÊN PHẢN BIỆN

*(ký tên)*

**PHIẾU ĐÁNH GIÁ BÁO CÁO KẾT QUẢ THỰC TẬP**

**Học kỳ III năm học 2024 - 2025**

**(Dùng cho giảng viên Hướng dẫn – tối đa 6 điểm)**

Họ và tên giảng viên Hướng dẫn: Đặng Mạnh Huy

Họ tên sinh viên thực tập: Nguyễn Tuấn Đạt Mã số SV: 219486

Lớp: DH21KMT01 Ngành: Khoa học máy tính

|  |  |  |
| --- | --- | --- |
| **Nội dung đánh giá** | **Điểm tối đa** | **Điểm chấm** |
| **I. Hình thức trình bày** | **0.5** |  |
| I.1 Đúng format của Khoa: Trang bìa, trang lời cảm ơn, 02 trang đánh giá thực tập của GVHD và GVPB (mẫu M-TT-04), trang tổng hợp điểm (M-TT-05), trang mục lục và các nội dung báo cáo. Sử dụng đúng mã và font tiếng Việt (Unicode Times New  Roman, Size 13) | 0.25 |  |
| I.2 Trình bày mạch lạc, súc tích, không có lỗi chính tả | 0.25 |  |
| **II. Phiếu theo dõi** | **0.5** |  |
| II.1 Có lịch làm việc đầy đủ cho 8 tuần | 0.25 |  |
| II.2 Số buổi thực tập tại cơ quan trong 1 tuần >=6; ít hơn 6 buổi  0.0 điểm | 0.25 |  |
| **III. Nội dung thực tập (quyển báo cáo)** | **3.0** |  |
| - Có được sự hiểu biết tốt về cơ quan nơi thực tập. | 0.5 |  |
| - Phương pháp thực hiện phù hợp với nội dung công việc được | 0.5 |
| giao. |  |
| - Kết quả củng cố lý thuyết. | 0.5 |
| - Kết quả rèn luyện kỹ năng thực hành. | 0.5 |
| - Kinh nghiệm thực tiễn thu nhận được. | 0.5 |
| - Kết quả công việc có đóng góp cho cơ quan nơi thực tập | 0.5 |
| **IV. Kết quả chất vấn sinh viên** | **2.0** |  |
| **TỔNG CỘNG** | **6.0** |  |
| Điểm trừ |  |  |
| **Điểm còn lại** |  |  |

Cần Thơ, ngày….tháng….năm 20….

GIẢNG VIÊN HƯỚNG DẪN

*(ký tên)*

**PHIẾU ĐIỂM TỔNG HỢP KẾT QUẢ THỰC TẬP**

**Học kỳ III năm học 2024 - 2025**

**(Dùng cho giảng viên Hướng dẫn tổng hợp – tối đa 10 điểm)**

Họ và tên giảng viên Hướng dẫn: Đặng Mạnh Huy

Họ tên sinh viên thực tập: Nguyễn Tuấn Đạt Mã số SV: 219486

Lớp: DH21KMT01 Ngành: Khoa học máy tính

|  |  |
| --- | --- |
| **Nguồn điểm** | **Điểm** |
| 1. Điểm của cán bộ hướng dẫn tại nơi thực tập (A) |  |
| 2. Điểm của giảng viên hướng dẫn (B) |  |
| 3. Điểm của giảng viên phản biện (C) |  |
| **Điểm tổng hợp= A+(B+C)/2 *(làm tròn tới 1 chữ số thập phân)*** |  |

Cần Thơ, ngày….tháng….năm 20….

GIẢNG VIÊN HƯỚNG DẪN

*(ký tên)*

**PHIẾU GIAO VIỆC CHO SINH VIÊN THỰC TẬP**

**Học kỳ …. năm học…………..**

Họ và tên sinh viên: .......................................................................... MSSV:......................

Lớp……………………………….Ngành………………………………………………… Cơ quan thực tập:.................................................................................................................

Họ và tên cán bộ hướng dẫn tại nơi thực tập: ......................................................................

Thời gian thực tập: từ ngày ………………. đến ……………...

|  |  |  |
| --- | --- | --- |
| Tuần | Nội dung công việc được giao (phù hợp với đề cương) | Số buổi hoặc giờ sinh viên làm việc tại cơ quan trong 1 tuần (phải >=6 buổi tương đương 24 giờ) |
| 1  Từ ngày  ..............  đến ngày  ............. |  |  |
| 2  Từ ngày  ..............  đến ngày  ............. |  |  |
| 3  Từ ngày  ..............  đến ngày  ............. |  |  |
| 4  Từ ngày  ..............  đến ngày  ............. |  |  |
| 5  Từ ngày  ..............  đến ngày  ............. |  |  |
| 6  Từ ngày  ..............  đến ngày  ............. |  |  |
| 7  Từ ngày  ..............  đến ngày  ............. |  |  |
| 8  Từ ngày  ..............  đến ngày  ............. |  |  |

Xác nhận của cơ quan *............., ngày......tháng.* *năm 20…*

Thủ trưởng Sinh viên Cán bộ hướng dẫn

*(Ký tên, đóng dấu) (Ký tên và ghi họ tên) (Ký tên và ghi họ tên)*

**MỤC LỤC**

[CHƯƠNG 1 7](#_Toc198532300)

[GIỚI THIỆU CƠ QUAN THỰC TẬP 7](#_Toc198532301)

[1.1 Tổng quan 7](#_Toc198532302)

[1.2 Chức năng, nhiệm vụ 7](#_Toc198532303)

[1.4 Thông tin liên hệ 7](#_Toc198532304)

[CHƯƠNG 2 9](#_Toc198532305)

[NỘI DUNG NGHIÊN CỨU 9](#_Toc198532306)

[2.1 Giới thiệu chung 9](#_Toc198532307)

[2.2 Lý do chọn đề tài 10](#_Toc198532308)

[2.3 Đối tượng và phạm vi nghiên cứu 10](#_Toc198532309)

[2.4 Mô tả chi tiết đề tài 10](#_Toc198532310)

[CHƯƠNG 3 PHƯƠNG PHÁP THỰC HIỆN 12](#_Toc198532311)

[3.1 Công cụ sử dụng 12](#_Toc198532312)

[3.2 Cơ sở lý luận 13](#_Toc198532313)

[CHƯƠNG 4 16](#_Toc198532314)

[KẾT QUẢ WEBSITE 16](#_Toc198532315)

[4.1 Chức năng hệ thống 16](#_Toc198532316)

[CHƯƠNG 5 24](#_Toc198532317)

[KẾT LUẬN NGHIÊN CỨU 24](#_Toc198532318)

[CHƯƠNG 6 25](#_Toc198532319)

[KẾT QUẢ ĐẠT ĐƯỢC QUA THỰC TẬP 25](#_Toc198532320)

# CHƯƠNG 1

# GIỚI THIỆU CƠ QUAN THỰC TẬP

## Tổng quan

Công ty TNHH Một Thành Viên Công Nghệ Kỹ Thuật Tiên Phong (tên quốc tế: Advanced High Technology One Member Limited Liability Company, viết tắt: ADHIGHTECH) đƣợc thành lập vào ngày 05 tháng 04 năm 2017. Trụ sở công ty đặt tại Số 10, đƣờng số 08, khu dân cƣ Long Thịnh, Phƣờng Phú Thứ, Quận Cái Răng, Thành phố Cần Thơ, Việt Nam.

Công ty hoạt động theo loại hình công ty trách nhiệm hữu hạn một thành viên ngoài nhà nước. Doanh nghiệp chuyên hoạt động trong lĩnh vực dịch vụ hỗ trợ kinh doanh khác chƣa đƣợc phân vào đâu, cụ thể là xuất nhập khẩu các mặt hàng mà công ty đang kinh doanh. Công ty hiện do ông Ngô Hồ Anh Khôi làm ngƣời đại diện pháp luật và chịu sự quản lý của Đội Thuế liên huyện Cái Răng - Phong Điền.

## Chức năng, nhiệm vụ

#### Chức năng

Công ty TNHH Một Thành Viên Công Nghệ Kỹ Thuật Tiên Phong có chức năng chính là cung cấp các dịch vụ hỗ trợ kinh doanh và tổ chức hoạt động xuất nhập khẩu. Ngoài ra, công ty còn đóng vai trò là đơn vị kết nối thị trƣờng trong và ngoài nƣớc thông qua các giải pháp công nghệ tiên tiến và dịch vụ kỹ thuật.

#### Nhiệm vụ

Triển khai các hoạt động xuất nhập khẩu một cách chuyên nghiệp, đảm bảo tuân thủ các quy định pháp lý hiện hành.

Hỗ trợ các doanh nghiệp trong và ngoài nước thực hiện hoạt động thương mại và hậu cần.

Xây dựng mối quan hệ hợp tác dài hạn và bền vững với đối tác, khách hàng.

Thực hiện đầy đủ nghĩa vụ thuế và tuân thủ quy định của các cơ quan quản lý nhà nước.

* 1. **Định hướng phát triển**

Công ty TNHH Một Thành Viên Công Nghệ Kỹ Thuật Tiên Phong hướng đến: Mở rộng quy mô hoạt động xuất nhập khẩu và đa dạng hóa mặt hàng kinh doanh.Tăng cường ứng dụng công nghệ trong công tác quản lý và điều hành doanh nghiệp. Phát triển mạng lưới đối tác trong và ngoài nƣớc, nâng cao năng lực cạnh tranh trên thị trƣờng quốc tế.Đào tạo nguồn nhân lực có kiến thức và kỹ năng chuyên sâu về công nghệ và thương mại.Trở thành doanh nghiệp tiên phong trong lĩnh vực dịch vụ hỗ trợ kinh doanh và kết nối giao thƣơng tại khu vực Đồng bằng sông Cửu Long.

## Thông tin liên hệ

Tên công ty: Công ty TNHH Một Thành Viên Công Nghệ Kỹ Thuật Tiên Phong (ADHIGHTECH)

Mã số thuế: 1801526082

Địa chỉ trụ sở chính: Số 10, đƣờng số 08, khu dân cƣ Long Thịnh, Phƣờng Phú Thứ, Quận Cái Răng, Thành phố Cần Thơ, Việt Nam

Điện thoại: 0916.416.409

# CHƯƠNG 2

# NỘI DUNG NGHIÊN CỨU

## Giới thiệu chung

Trí tuệ nhân tạo là lĩnh vực nghiên cứu phát triển các hệ thống máy tính có khả năng “học”, suy luận, và tự ra quyết định tương tự con người. Với khả năng học từ dữ liệu và dự đoán hành vi người dùng, AI đang dần trở thành công nghệ cốt lõi, góp phần thay đổi cách con người tương tác với hệ thống và xử lý công việc. Một trong những ứng dụng thực tiễn và tiềm năng của AI chính là trong hệ thống điền biểu mẫu tự động. Hiện nay, người dùng thường xuyên phải tương tác với nhiều loại biểu mẫu khác nhau trong công việc, học tập và cuộc sống hàng ngày. Việc điền thông tin vào các biểu mẫu này một cách thủ công không chỉ tốn thời gian mà còn dễ gây ra sai sót, ảnh hưởng đến hiệu suất và trải nghiệm người dùng.

Tại Trường Đại học Nam Cần Thơ (DNC), sinh viên và cán bộ thường xuyên phải sử dụng các biểu mẫu trực tuyến cho nhiều mục đích khác nhau như đăng ký môn học, đăng ký xét tốt nghiệp, khảo sát, thu thập thông tin phản hồi, v.v. Việc điền thông tin lặp đi lặp lại vào các trường tương tự trên các biểu mẫu khác nhau có thể gây ra sự nhàm chán và lãng phí thời gian.

Hệ thống điền biểu mẫu tự động ứng dụng AI ra đời như một giải pháp hiệu quả nhằm: tự động điền thông tin vào biểu mẫu dựa trên dữ liệu đã từng được nhập trước đó, học hỏi từ lịch sử người dùng để đưa ra các gợi ý phù hợp cho từng trường dữ liệu, tiết kiệm thời gian và công sức, đặc biệt trong các công việc hành chính có tính lặp lại cao, giảm thiểu sai sót nhập liệu, từ đó nâng cao độ chính xác và hiệu quả xử lý thông tin, cải thiện trải nghiệm người dùng, nhất là trong các hệ thống trực tuyến và ứng dụng di động. Không chỉ dừng lại ở việc hỗ trợ điền văn bản, hệ thống còn có khả năng xuất kết quả ở nhiều định dạng khác nhau như .docx, .pdf, hoặc ảnh, tùy theo nhu cầu sử dụng. Đây là một bước tiến quan trọng trong việc ứng dụng AI vào thực tiễn, hướng đến số hóa và tự động hóa quy trình làm việc, giúp tổ chức hoạt động linh hoạt và chuyên nghiệp hơn.

## Lý do chọn đề tài

Là một sinh viên ngành Công nghệ Thông tin, em nhận thấy rằng việc tối ưu hóa quá trình nhập liệu và xử lý dữ liệu là một trong những vấn đề mang tính thực tiễn cao. Trong quá trình thực tập và làm việc thực tế, em gặp rất nhiều tình huống phải điền lại các biểu mẫu có nội dung gần như giống nhau, điều này gây mất thời gian và dễ gây nhầm lẫn.

Với sự phát triển mạnh mẽ của trí tuệ nhân tạo và học máy, em nhận thấy có thể tận dụng các thuật toán học từ dữ liệu lịch sử để hỗ trợ người dùng một cách thông minh trong việc hoàn thiện biểu mẫu. Chính vì vậy, em quyết định chọn đề tài này với mong muốn xây dựng một hệ thống hỗ trợ tự động, thân thiện, dễ áp dụng trong thực tế, đồng thời nâng cao kỹ năng lập trình, xử lý dữ liệu, và triển khai mô hình AI.

## Đối tượng và phạm vi nghiên cứu

❖ Đối tượng nghiên cứu:

Quy trình điền biểu mẫu trực tuyến của sinh viên và cán bộ tại Trường Đại học Nam Cần Thơ.

Các phương pháp và công nghệ liên quan đến việc tự động điền biểu mẫu, quản lý lịch sử người dùng và gợi ý thông tin.

Các hệ thống điền biểu mẫu tự động hiện có và đánh giá hiệu quả của chúng.

❖ Phạm vi nghiên cứu:

Biểu mẫu đầu vào là các biểu mẫu dạng văn bản (text-based) phổ biến như hợp đồng, đăng ký thông tin, phiếu yêu cầu,...

Hệ thống được xây dựng với giao diện đơn giản, dễ sử dụng.

Hệ thống có khả năng lưu trữ dữ liệu lịch sử người dùng để phục vụ cho quá trình học và gợi ý.

Hệ thống có thể xuất kết quả ra các định dạng phổ biến như .docx, .pdf, hoặc ảnh.

Không đi sâu vào các phương pháp xử lý ngôn ngữ nâng cao hoặc hệ thống bảo mật dữ liệu phức tạp, nhưng đảm bảo tuân thủ các nguyên tắc bảo mật cơ bản về thông tin người dùng.

## Mô tả chi tiết đề tài

#### Chức năng chính cho ứng dụng

Quản lý người dùng: Đăng ký, đăng nhập, đăng xuất, xác thực OAuth, phân quyền, quản lý hồ sơ

Quản lý biểu mẫu: Tải lên docx, trích xuất trường, nhập liệu với gợi ý AI, lưu trữ lịch sử

Xử lý tài liệu: Tạo docx từ biểu mẫu và dữ liệu, tải xuống, quản lý lượt tải xuống

Tích hợp AI gợi ý điền giá trị: Gợi ý thông minh dựa trên ngữ cảnh, phân tích nội dung, cá nhân hóa gợi ý

Triển khai: Đóng gói Docker, cấu hình môi trường, tối ưu hóa hiệu suất

#### Yêu cầu đề tài

Xây dựng Website cho hệ thống điền biểu mẫu tự động dựa trên lịch sử.

Cấu hình Website, chỉnh sửa thông tin User, phân quyền web, API docs, APP login google, Đóng gói docker.

# CHƯƠNG 3 PHƯƠNG PHÁP THỰC HIỆN

## Công cụ sử dụng

Các công cụ sử dụng trong quá trình hoàn thành luận văn là một điều vô cùng cần thiết, các công cụ chiếm một vai trò quan trọng nhất bên cạnh hình thành ý tưởng của luận văn và góp phần xây dựng nên một chương trình hoàn thiện, đáp ứng nhu cầu người sử dụng. Việc lựa chọn những công cụ tốt, thông dụng và dễ dàng sử dụng nhất góp phần không nhỏ vào việc hoàn thiện luận văn tốt nghiệp. Dự án “Hệ thống điền biểu mẫu tự động dựa trên lịch sử người dùng” sử dụng nhiều công nghệ hiện đại, phản ánh xu hướng phát triển phần mềm thông minh và đáp ứng nhu cầu thực tiễn.

*Bảng 3. 1: Công cụ sử dụng*

|  |  |  |
| --- | --- | --- |
| Loại công nghệ | Công cụ / Công nghệ | Mô tả |
| Ngôn ngữ lập trình | Python | Ngôn ngữ chính để xây dựng backend và xử lý AI, dữ liệu. |
| Backend | Flask | Framework web nhẹ, phù hợp xây dựng API và ứng dụng web. |
| Frontend | HTML, CSS, JavaScript | Các công nghệ nền tảng xây dựng giao diện người dùng. |
| Tailwind CSS | CSS framework dạng tiện ích giúp phát triển UI hiện đại, linh hoạt. |  |
| Cơ sở dữ liệu | SQLite | Dùng cho môi trường phát triển, đơn giản, dựa trên tệp. |
| PostgreSQL / MySQL | Cơ sở dữ liệu dùng trong môi trường sản xuất, mạnh mẽ và mở rộng tốt. |  |
| Xử lý tài liệu | python-docx | Thư viện dùng để tạo, chỉnh sửa và xuất file Word (DOCX). |
| AI / Machine Learning | OpenAI API (GPT-4) | Sử dụng AI để đưa ra gợi ý điền mẫu dựa trên lịch sử nhập liệu. |
| Xác thực người dùng | Flask-Login | Quản lý đăng nhập và phiên làm việc của người dùng. |
| OAuth (Google) | Cho phép đăng nhập bằng tài khoản Google, tiện lợi và bảo mật. |  |
| Triển khai | Docker, Docker Compose | Container hóa ứng dụng giúp triển khai đồng nhất trên các môi trường. |
| Quản lý mã nguồn | Git | Theo dõi phiên bản mã nguồn, hỗ trợ cộng tác và kiểm soát thay đổi |

*.*

## Cơ sở lý luận

#### Đặt vấn đề nghiên cứu

Trong thực tế, việc điền biểu mẫu lặp đi lặp lại là một công việc tiêu tốn thời gian và dễ mắc lỗi. Trong các tổ chức hoặc doanh nghiệp, nhân viên thường xuyên phải hoàn thành các biểu mẫu với thông tin trùng lặp hoặc chỉ thay đổi một phần nhỏ, làm giảm năng suất và dễ dẫn đến nhầm lẫn. Điều này đặt ra nhu cầu cấp thiết về một hệ thống thông minh có khả năng tự động hóa việc nhập liệu, đồng thời đưa ra các gợi ý dựa trên lịch sử sử dụng trước đó.

Việc nghiên cứu và xây dựng hệ thống điền biểu mẫu tự động ứng dụng AI không chỉ giải quyết vấn đề tiết kiệm thời gian và công sức mà còn góp phần tăng cường độ chính xác, hiệu quả công việc và trải nghiệm người dùng trong môi trường số. Đặc biệt, việc tích hợp các công nghệ trí tuệ nhân tạo tiên tiến như OpenAI API (GPT-4) hứa hẹn mang lại khả năng gợi ý và điền thông tin một cách ngữ cảnh và chính xác, nâng cao đáng kể trải nghiệm người dùng.

Tuy nhiên, việc khai thác hiệu quả lịch sử người dùng để dự đoán và điền thông tin phù hợp, cũng như việc tích hợp liền mạch các công nghệ AI vào quy trình điền biểu mẫu vẫn còn nhiều vấn đề cần được nghiên cứu. Cụ thể, làm thế nào để hệ thống có thể học hỏi từ các lần điền biểu mẫu trước đó của người dùng một cách hiệu quả? Làm thế nào để tận dụng sức mạnh của GPT-4 trong việc hiểu ngữ cảnh của biểu mẫu và đưa ra các gợi ý phù hợp? Những thách thức về bảo mật dữ liệu lịch sử người dùng và quản lý phiên làm việc cũng cần được xem xét kỹ lưỡng.

Xuất phát từ những vấn đề trên, đề tài "Hệ thống điền biểu mẫu tự động dựa trên lịch sử người dùng" được thực hiện nhằm nghiên cứu và đề xuất một giải pháp tự động hóa quá trình điền biểu mẫu, tận dụng các công nghệ phát triển web hiện đại và sức mạnh của trí tuệ nhân tạo, hướng đến việc nâng cao hiệu quả và trải nghiệm người dùng tại Trường Đại học Nam Cần Thơ.

#### Mục tiêu nghiên cứu

❖ Nghiên cứu và phân tích

Tìm hiểu về quy trình điền biểu mẫu hiện tại của sinh viên và cán bộ tại Trường Đại học Nam Cần Thơ.

Nghiên cứu các công nghệ phát triển web (Python, Flask, HTML, CSS, JavaScript, Tailwind CSS), cơ sở dữ liệu (SQLite, PostgreSQL/MySQL), xử lý tài liệu (Python-docx), AI và ML (OpenAI API (GPT-4)), xác thực (Flask-Login, OAuth (Google)), containerization (Docker, Docker Compose) và quản lý phiên bản (Git) trong bối cảnh xây dựng hệ thống điền biểu mẫu tự động.

Phân tích các giải pháp và hệ thống điền biểu mẫu tự động hiện có, đánh giá ưu nhược điểm và khả năng ứng dụng.

❖ Thiết kế và xây dựng mô hình:

Đề xuất kiến trúc tổng thể cho hệ thống điền biểu mẫu tự động dựa trên lịch sử người dùng, tận dụng các công nghệ đã nghiên cứu.

Thiết kế cơ sở dữ liệu để lưu trữ lịch sử điền biểu mẫu của người dùng một cách hiệu quả và bảo mật.

Xây dựng mô hình backend sử dụng Python và Flask để quản lý logic nghiệp vụ, tương tác với cơ sở dữ liệu và tích hợp với OpenAI API.

Phát triển giao diện người dùng (frontend) thân thiện và dễ sử dụng bằng HTML, CSS, JavaScript và Tailwind CSS, cho phép người dùng tương tác với hệ thống.

Nghiên cứu và triển khai phương pháp tích hợp OpenAI API (GPT-4) để cung cấp các gợi ý điền thông tin thông minh và ngữ cảnh.

Triển khai các cơ chế xác thực người dùng (Flask-Login, Google OAuth) để đảm bảo an toàn và quản lý truy cập.

Triển khai hệ thống bằng Docker, đảm bảo có thể chạy ổn định trên nhiều môi trường khác nhau.

#### Phương pháp nghiên cứu

❖ Phân tích tài liệu

Nghiên cứu các tài liệu khoa học, sách, báo cáo liên quan đến phát triển web (backend, frontend), quản lý cơ sở dữ liệu, trí tuệ nhân tạo (đặc biệt là các ứng dụng của mô hình ngôn ngữ lớn như GPT-4), các kỹ thuật xác thực và quản lý phiên, cũng như các hệ thống điền biểu mẫu tự động hiện có.

❖ Phân tích và thiết kế hệ thống

Phân tích yêu cầu của người dùng và các quy trình nghiệp vụ liên quan đến việc điền biểu mẫu.

Sử dụng các phương pháp thiết kế phần mềm hướng đối tượng (Object-Oriented Design) để xây dựng kiến trúc hệ thống và mô hình dữ liệu.

Thiết kế giao diện người dùng dựa trên các nguyên tắc về tính dễ sử dụng (Usability).

Phát triển phần mềm theo mô hình Agile: Phân chia dự án thành các giai đoạn nhỏ để phát triển, kiểm thử và điều chỉnh linh hoạt.

Sử dụng các công nghệ đã lựa chọn (Python, Flask, HTML, CSS, JavaScript, Tailwind CSS, SQLite/PostgreSQL/MySQL, OpenAI API, Flask-Login, OAuth (Google), Docker, Git) để xây dựng và triển khai hệ thống.

Áp dụng các phương pháp phát triển phần mềm linh hoạt (Agile) để quản lý quá trình phát triển một cách hiệu quả.

# CHƯƠNG 4

# KẾT QUẢ WEBSITE

## Chức năng hệ thống

Việc hoàn thành các chức năng cốt lõi như quản lý người dùng, quản lý biểu mẫu, xử lý tài liệu và tích hợp AI cho thấy việc thực hiện thành công các nhiệm vụ đã lên kế hoạch. Tuyên bố của thực tập sinh về một hệ thống ổn định và đáp ứng cần được hỗ trợ bởi các số liệu hiệu suất tiềm năng trong báo cáo. Các đoạn trích dẫn thảo luận về các kỹ thuật tối ưu hóa hiệu suất cho các ứng dụng Flask. Một hệ thống nhập liệu thông minh hoạt động tốt sẽ xử lý hiệu quả các tương tác của người dùng và xử lý dữ liệu. Tính ổn định và khả năng phản hồi là rất quan trọng để có trải nghiệm người dùng tích cực.

Việc đề cập đến "gợi ý thông minh dựa trên ngữ cảnh biểu mẫu" và "phân tích nội dung biểu mẫu để hiểu ngữ cảnh" cho thấy việc sử dụng tinh vi OpenAI API. Báo cáo nên cung cấp chi tiết về cách ngữ cảnh được trích xuất và sử dụng để tạo ra các gợi ý có liên quan. Các gợi ý được hỗ trợ bởi AI hiệu quả nên nhận biết ngữ cảnh để thực sự hữu ích cho người dùng. Điều này đòi hỏi phải thiết kế cẩn thận sự tương tác với mô hình AI.

Việc phát triển một API RESTful cho tất cả các chức năng là một lựa chọn kiến trúc tốt, cho phép tích hợp tiềm năng với các hệ thống hoặc nền tảng khác. API RESTful là một cách tiêu chuẩn để cung cấp các chức năng ứng dụng web, thúc đẩy khả năng tương tác và khả năng mở rộng. Việc sử dụng Docker để đóng gói và triển khai phù hợp với các thông lệ DevOps hiện đại, tạo điều kiện triển khai nhất quán và có thể tái sản xuất trên các môi trường khác nhau. Docker đơn giản hóa quy trình triển khai bằng cách container hóa ứng dụng và các phần phụ thuộc của nó, đảm bảo một môi trường thời gian chạy nhất quán. Việc bao gồm tài liệu API và hướng dẫn sử dụng là những chỉ số tích cực về một dự án hoàn chỉnh. Các đoạn trích dẫn làm nổi bật tầm quan trọng của tài liệu API và thảo luận về các thông lệ tốt nhất để viết hướng dẫn sử dụng.

❖ Các chức năng chính:

❖ Admin

Quản lý người dùng

Quản lý biểu mẫu

Quản lý lịch sử biểu mẫu

Cấu hình website

Tạo tài liệu docx

Chức năng điền tự động

Lưu và tạo tài liệu docx

❖ Users

Chỉnh sửa thông tin cá nhân

Tạo tài liệu docx

Chức năng điền tự động

Lưu và tạo tài liệu docx

#### 4.1.1 Chức năng quản lý người dùng

Hệ thống hỗ trợ một loạt các chức năng quản lý người dùng nhằm đảm bảo bảo mật, cá nhân hóa trải nghiệm và kiểm soát quyền truy cập:

❖ Đăng ký tài khoản

Người dùng có thể tạo tài khoản mới bằng cách cung cấp email, tên người dùng và mật khẩu.

Mật khẩu được mã hóa trước khi lưu vào cơ sở dữ liệu để đảm bảo an toàn.

❖ Đăng nhập

Người dùng có thể đăng nhập bằng email và mật khẩu hợp lệ.

Sau khi đăng nhập thành công, phiên làm việc sẽ được thiết lập để nhận diện người dùng trong suốt quá trình sử dụng hệ thống.

❖ Đăng xuất

Cho phép người dùng thoát khỏi hệ thống, kết thúc phiên làm việc một cách an toàn.

Dữ liệu phiên sẽ được xóa để đảm bảo bảo mật.

❖ Xác thực OAuth (Google Login)

Hệ thống tích hợp xác thực bằng tài khoản Google thông qua giao thức OAuth 2.0.

Người dùng có thể đăng nhập nhanh mà không cần tạo mật khẩu mới.

Giảm thiểu rủi ro bảo mật liên quan đến mật khẩu yếu hoặc bị rò rỉ.

❖ Phân quyền người dùng

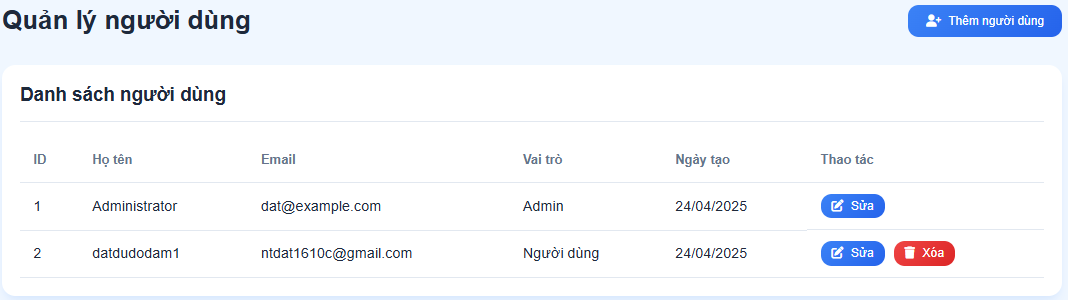
Hệ thống có thể phân loại người dùng thành các vai trò như: người dùng thông thường, quản trị viên.

Mỗi vai trò có các quyền truy cập khác nhau: quản trị viên có quyền quản lý người dùng, chỉnh sửa biểu mẫu, xem thống kê...

❖ Quản lý hồ sơ cá nhân

Người dùng có thể cập nhật thông tin cá nhân như: tên, ảnh đại diện, thông tin liên hệ.

Hồ sơ cũng lưu trữ lịch sử các biểu mẫu đã điền để phục vụ gợi ý sau này.



*Hình 4. 1: Kết quả, giao diện quản lý người dùng*

#### 4.1.2 Chức năng quản lý biểu mẫu

Hệ thống hỗ trợ quy trình làm việc linh hoạt với các tệp biểu mẫu .docx, kết hợp trí tuệ nhân tạo (AI) để tự động hóa và tối ưu hóa việc điền thông tin:

❖ Tải lên tệp biểu mẫu (.docx)

Cho phép người dùng tải lên các mẫu tài liệu Word (.docx) cần điền thông tin.

Hệ thống xử lý nội dung tệp để chuẩn bị cho các bước trích xuất tiếp theo.

❖ Trích xuất trường dữ liệu

Hệ thống tự động nhận diện và trích xuất các trường cần điền trong biểu mẫu (ví dụ: [Tên], [Địa chỉ], [Ngày sinh]...).

Các trường được hiển thị dưới dạng form nhập liệu trên giao diện web.

❖ Nhập liệu với gợi ý AI

Hệ thống sử dụng mô hình ngôn ngữ (GPT-4) để gợi ý thông tin dựa trên:

Lịch sử biểu mẫu đã điền của người dùng.

Ngữ cảnh của biểu mẫu hiện tại.

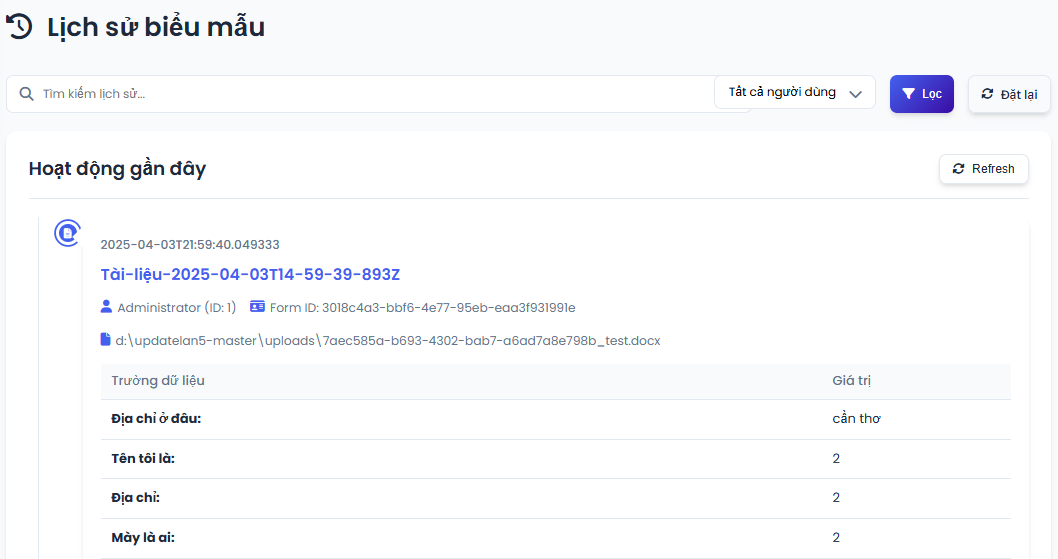
Gợi ý có thể chỉnh sửa trước khi người dùng xác nhận để hoàn tất biểu mẫu.

❖ Lưu trữ lịch sử biểu mẫu

Mỗi biểu mẫu đã điền được lưu vào cơ sở dữ liệu gắn liền với tài khoản người dùng.

Cho phép truy xuất, chỉnh sửa hoặc tải lại biểu mẫu cũ bất kỳ lúc nào.

Hệ thống sử dụng lịch sử này để cải thiện gợi ý AI trong tương lai.



*Hình 4. 2: Kết quả, giao diện quản lý lịch sử biểu mẫu*

#### 4.1.3 Chức năng tạo tài liệu docx

Hệ thống cung cấp khả năng tạo và quản lý tài liệu đầu ra (.docx) dựa trên biểu mẫu đã xử lý và dữ liệu người dùng nhập vào, giúp người dùng hoàn thiện quy trình làm việc một cách nhanh chóng và hiệu quả.

❖ Tạo file DOCX từ biểu mẫu và dữ liệu

Hệ thống kết hợp biểu mẫu gốc với dữ liệu đã điền để sinh ra tệp .docx hoàn chỉnh.

Các trường được thay thế chính xác theo vị trí trong biểu mẫu gốc.

Hỗ trợ định dạng giữ nguyên (font, cỡ chữ, màu sắc...) giúp đảm bảo tính chuyên nghiệp.

❖ Tải xuống tài liệu

Người dùng có thể tải xuống tài liệu đã hoàn chỉnh dưới định dạng .docx.

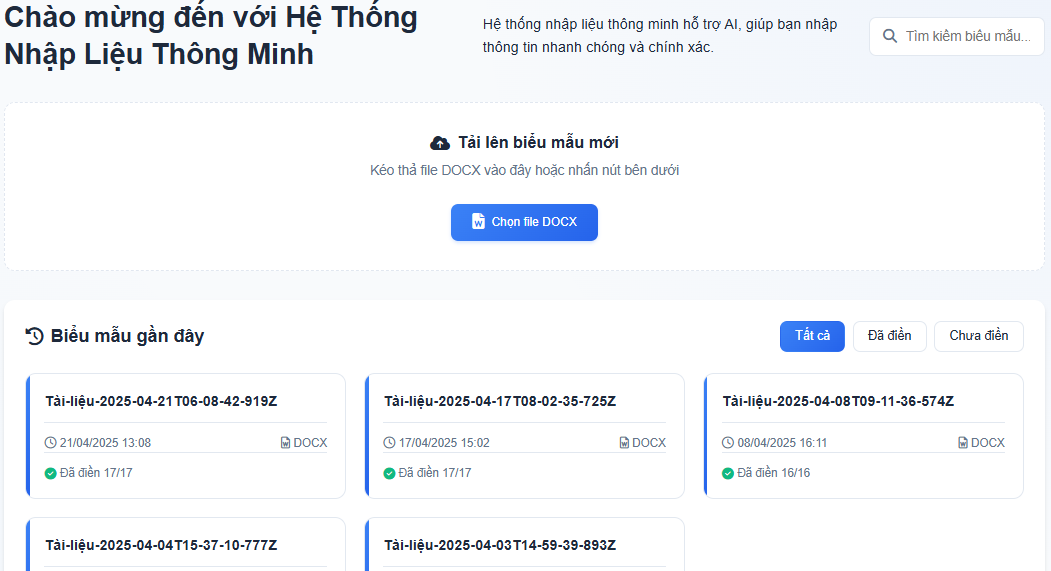
Hệ thống tự động gán tên file theo thời gian hoặc nội dung biểu mẫu để dễ quản lý.

❖ Quản lý lượt tải xuống

Hệ thống lưu lại số lần và thời gian tải xuống của mỗi biểu mẫu.

Cho phép người dùng xem lại các tệp đã từng tải và tải lại nếu cần.

Tính năng này hữu ích để kiểm tra quá trình làm việc hoặc khôi phục khi mất tệp gốc.



*Hình 4. 3: Chức năng tạo tài liệu docx*

#### 4.1.4 Chức năng tự động điền

Hệ thống ứng dụng trí tuệ nhân tạo (AI), cụ thể là mô hình ngôn ngữ GPT-4, nhằm tự động hóa quá trình điền biểu mẫu bằng cách đưa ra gợi ý nhập liệu thông minh, giúp tiết kiệm thời gian và nâng cao độ chính xác.

❖ Tích hợp AI (OpenAI GPT-4)

Hệ thống sử dụng API của OpenAI để tích hợp mô hình ngôn ngữ mạnh mẽ GPT-4.

AI có khả năng hiểu ngữ cảnh của biểu mẫu và nội dung đã có để hỗ trợ người dùng hiệu quả.

❖ Gợi ý thông minh dựa trên ngữ cảnh

Phân tích nội dung biểu mẫu hiện tại, nhận diện các trường cần điền.

Đưa ra gợi ý phù hợp với chủ đề và nội dung cụ thể của từng biểu mẫu.

❖ Phân tích nội dung biểu mẫu

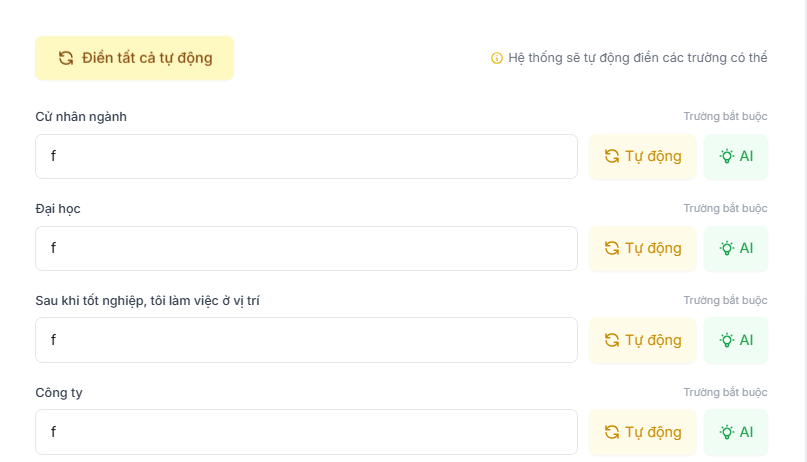
AI trích xuất thông tin từ văn bản, nhận diện mối quan hệ giữa các trường (ví dụ: họ tên – giới tính – ngày sinh).

Giúp đảm bảo dữ liệu điền vào phù hợp logic và mạch lạc.

❖ Cá nhân hóa gợi ý dựa trên người dùng

Hệ thống học hỏi từ lịch sử nhập liệu của từng người dùng để đưa ra gợi ý mang tính cá nhân hóa.

Gợi ý có thể thay đổi linh hoạt tùy theo từng người dùng và từng biểu mẫu.



*Hình 4. 4: Chức năng tự động điền*

#### 4.1.5 Thư viện hỗ trợ phần mềm

Việc sử dụng các thư viện trong website yêu cầu kiến thức đa dạng và sâu rộng về ngôn ngữ lập trình Python. Các thư viện này đa phần được cài đặt và tích hợp vào website để cung cấp các tính năng và chức năng phong phú và đa dạng. Tuy nhiên, việc làm này đòi hỏi nhà lập trình phải có khả năng tìm hiểu và áp dụng các thư viện một cách hiệu quả.

Ngoài ra, việc quản lý và bảo trì các thư viện cũng là một phần quan trọng trong quá trình phát triển website. Nhà lập trình cần phải biết cách cài đặt và quản lý các thư viện một cách hiệu quả để đảm bảo tính ổn định và bảo mật của hệ thống. Tóm lại, việc sử dụng các thư viện trong phát triển website đòi hỏi sự am hiểu sâu rộng về ngôn ngữ lập trình và kỹ năng sử dụng các công cụ và tài nguyên có sẵn một cách linh hoạt và hiệu quả.

*Bảng 4. 1: Thư viện hỗ trợ hệ thống được cài đặt*

|  |  |
| --- | --- |
| flask | Framework web nhẹ, dùng để xây dựng API và giao diện người dùng. |
| flask-login | Hỗ trợ quản lý đăng nhập, đăng xuất và phiên làm việc người dùng. |
| flask-sqlalchemy | ORM để tương tác dễ dàng với cơ sở dữ liệu quan hệ như SQLite, MySQL, PostgreSQL. |
| werkzeug | Thư viện hỗ trợ WSGI và các chức năng như routing, bảo mật, và xử lý HTTP. |
| python-docx | Dùng để tạo, chỉnh sửa và lưu trữ file tài liệu Microsoft Word (.docx). |
| openai | Giao tiếp với OpenAI API (GPT-4) để sinh văn bản và tạo gợi ý điền biểu mẫu. |
| numpy | Thư viện tính toán khoa học với mảng đa chiều, hỗ trợ các thao tác hiệu suất cao. |
| pandas | Xử lý dữ liệu dạng bảng (DataFrame), dùng để lưu trữ và phân tích lịch sử người dùng. |
| scikit-learn | Cung cấp các thuật toán học máy cơ bản như phân loại, hồi quy, phân cụm. |
| nltk | Thư viện xử lý ngôn ngữ tự nhiên, hỗ trợ phân tích cú pháp, tách từ, gắn thẻ từ loại. |
| scipy | Bổ trợ cho NumPy trong các phép toán khoa học, tối ưu hóa, thống kê. |
| gensim | Xử lý mô hình chủ đề, biểu diễn văn bản bằng vector (word2vec, LDA). |
| python-dotenv | Đọc và quản lý biến môi trường từ file .env, đảm bảo an toàn cấu hình hệ thống. |
| authlib | Cung cấp OAuth và OpenID để xác thực người dùng qua bên thứ ba như Google. |
| requests | Thư viện HTTP để gửi yêu cầu và lấy dữ liệu từ các API. |
| flask-oauthlib | Hỗ trợ xác thực OAuth1/OAuth2 trong Flask (kế thừa từ OAuthlib). |
| Flask-Session | Quản lý phiên làm việc người dùng bằng cách lưu session phía server. |
| sentence-transformers | Biến câu hoặc đoạn văn thành vector ngữ nghĩa, hỗ trợ so khớp và gợi ý thông minh. |
| huggingface\_hub[hf\_xet] | Kết nối tới mô hình học sâu từ HuggingFace, sử dụng các mô hình xử lý ngôn ngữ nâng cao. |

# CHƯƠNG 5

# KẾT LUẬN NGHIÊN CỨU

Hệ thống điền biểu mẫu tự động là một dự án đầy tham vọng và phức tạp liên quan đến việc phát triển một hệ thống nhập liệu thông minh hỗ trợ AI. Thực tập sinh đã thể hiện sự tham gia vào nhiều khía cạnh của vòng đời phát triển phần mềm, từ phân tích và thiết kế đến phát triển backend và frontend, tích hợp AI, triển khai và tài liệu. Việc lựa chọn công nghệ là hiện đại và phù hợp với yêu cầu của dự án. Các kết quả đạt được, như được mô tả trong báo cáo, cho thấy một hệ thống hoạt động ổn định và đáp ứng các yêu cầu chính.

# CHƯƠNG 6

# KẾT QUẢ ĐẠT ĐƯỢC QUA THỰC TẬP

Trong quá trình thực tập đã mang lại nhiều kết quả tích cực, đóng góp vào sự phát triển cả về kiến thức và kỹ năng trong lĩnh vực Trí tuệ nhân tạo (AI), tôi có cơ hội học hỏi và phát triển một loạt các kỹ năng và kiến thức quan trọng:

❖ Kiến thức lý thuyết được củng cố/bổ sung:

Phạm vi kiến thức lý thuyết được củng cố và bổ sung bao gồm phát triển web, cơ sở dữ liệu, bảo mật, AI và ML, và DevOps, cho thấy một trải nghiệm học tập toàn diện. Việc sử dụng Flask và mô hình MVC chứng minh sự hiểu biết sâu sắc về các nguyên tắc phát triển web. Việc lựa chọn cơ sở dữ liệu và đề cập đến ORM (Object-Relational Mapping) với SQLAlchemy cho thấy sự nắm bắt vững chắc các khái niệm về cơ sở dữ liệu. Việc triển khai các phương pháp xác thực và đề cập đến việc bảo vệ chống lại các lỗ hổng bảo mật phổ biến cho thấy sự hiểu biết về bảo mật web. Việc tích hợp OpenAI API và đề cập đến xử lý ngôn ngữ tự nhiên (NLP) cơ bản cho thấy sự hiểu biết về AI/ML. Cuối cùng, việc sử dụng Docker cho container hóa cho thấy kiến thức về các thông lệ DevOps.

❖ Kỹ năng thực hành học được:

Kỹ năng lập trình Python thực tế đã được cải thiện đáng kể thông qua việc viết code cho toàn bộ hệ thống. Việc đối mặt với các vấn đề thực tế và tìm cách giải quyết chúng đã giúp nâng cao khả năng lập trình.

Kỹ năng xây dựng giao diện form động bằng Flask đã được phát triển, bao gồm việc tạo các template HTML, xử lý dữ liệu form, và hiển thị thông tin một cách linh hoạt.

Các kỹ năng thực hành đã học bao gồm phát triển backend và frontend, sử dụng thư viện xử lý tài liệu, tối ưu hóa code, thiết kế UI/API/cơ sở dữ liệu/kiến trúc hệ thống, tích hợp API bên thứ ba và OAuth, kiểm thử đơn vị và tích hợp, gỡ lỗi và quản lý dự án theo phương pháp Agile. Phạm vi kỹ năng này rất quan trọng đối với một nhà phát triển full-stack, và sự tham gia của thực tập sinh vào các lĩnh vực này cho thấy một trải nghiệm học tập toàn diện. Việc đề cập đến phương pháp Agile cho thấy sự tiếp xúc với các thông lệ phát triển phần mềm hiện đại.

❖ Kinh nghiệm thực tiễn tích lũy:

Kinh nghiệm thực tiễn thu được bao gồm phát triển sản phẩm thực tế, làm việc với công nghệ mới, tối ưu hóa hiệu suất, xử lý vấn đề bảo mật và triển khai và bảo trì. Việc xây dựng một sản phẩm hoàn chỉnh từ đầu đến cuối là một kinh nghiệm vô giá cho một thực tập sinh. Báo cáo nên làm nổi bật những thách thức chính gặp phải và các phương pháp giải quyết vấn đề đã được áp dụng. Phát triển phần mềm thực tế liên quan đến việc điều hướng các phức tạp và sự không chắc chắn thường không có trong các dự án học thuật. Kinh nghiệm này giúp phát triển tư duy phản biện và kỹ năng giải quyết vấn đề.