HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Chu Đức Chính

NGHIÊN CỨU PHÁT TRIỂN HỆ THỐNG NGƯỜI ĐẠI DIỆN ẢO TRONG HOẠT ĐỘNG TƯ VẤN LỄ TÂN, CUNG CẤP THÔNG TIN CHO SINH VIÊN HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Chuyên ngành: Hệ thống thông tin

TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

HÀ NỘI - NĂM 2025

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học:

……TS. Trần Tiến Công……

(Ghi rõ học hàm, học vị)

Phản biện 1:

……

……

Phản biện 2:

……

…..

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ..... giờ ..... ngày ..... tháng .... năm .....

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

Trong bối cảnh công nghệ số phát triển mạnh mẽ và xu hướng chuyển đổi số diễn ra trên toàn cầu, việc ứng dụng trí tuệ nhân tạo (AI) vào các hoạt động dịch vụ khách hàng đã trở thành một xu thế tất yếu. Đặc biệt, hoạt động tư vấn lễ tân – một khâu quan trọng trong việc tạo ấn tượng đầu tiên với khách hàng – đang đối mặt với nhiều thách thức như: chi phí nhân sự cao, tính nhất quán trong chất lượng dịch vụ, khả năng phục vụ 24/7, và nhu cầu ngày càng tăng về trải nghiệm khách hàng cá nhân hóa.

Với phạm vi trong trường đại học như Học viện Công nghệ Bưu chính Viễn thông thì việc ứng dụng công nghệ số được phát triển rộng rãi, cùng với sự phát triển của công nghệ nhu cầu cung cấp thông tin cho sinh viên cũng là vấn đề cần chú ý tới làm sao có thể đáp ứng cung cấp thông tin được nhanh chóng và chính xác.

Hệ thống người đại diện ảo (Virtual Agent) với khả năng tương tác thông minh, xử lý ngôn ngữ tự nhiên và mô phỏng hành vi con người đang được coi là giải pháp hiệu quả để giải quyết những vấn đề trên. Việc nghiên cứu và phát triển một hệ thống như vậy không chỉ có ý nghĩa thực tiễn cao trong việc nâng cao chất lượng dịch vụ mà còn góp phần vào sự phát triển của lĩnh vực công nghệ AI tại Việt Nam.

Chính vì lý do đó tôi đã lựa chọn và nghiên cứu đề tài: "***Nghiên cứu phát triển hệ thống người đại diện ảo trong hoạt động tư vấn lễ tân, cung cấp thông tin cho sinh viên Học viện Công nghệ Bưu chính Viễn thông***". nhằm đóng góp vào việc ứng dụng công nghệ tiên tiến để cải thiện trải nghiệm nhỏ hơn là với sinh viên Học viện Công nghệ Bưu chính Viễn thông và tương lại rộng hơn là với khách hàng và tối ưu hóa hoạt động dịch vụ .

Nội dung đề án được chia thành 3 chương chính như sau:

Chương 1: Cơ sở lý thuyết và tổng quan

Giới thiệu tổng quan về hệ thống người đại diện ảo, cơ sở lý thuyết về trí tuệ nhân tạo (AI), và xử lý ngôn ngữ tự nhiên và ứng dụng vào trong hoạt động lễ tân. Nội dung chính của chương này sẽ giới thiệu về các khái niệm, các kiến thức mới liên quan đến người đại diện ảo, cơ sở lý thuyết về TTNT và xử lý ngôn ngữ tự nhiên. Ngoài ra, chương này cũng đưa ra phương thức ứng dụng trong hoạt động tư vấn lễ tân trong việc cung cấp thông tin tới sinh viên.

Chương 2: Phân tích thiết kế và xây dựng hệ thống

Nội dung chính của chương này sẽ tập trung vào việc phân tích, thiết kế và cài đặt triển khai hệ thống.

Chương 3: Thực nghiệm và đánh giá kết quả

Trong lĩnh vực trí tuệ nhân tạo nói chung và học sâu nói riêng, tập dữ liệu đóng vai trò quan trọng trong việc phát triển và đánh giá các thuật toán, ứng dụng. Việc phân tích và đánh giá chất lượng của tập dữ liệu cũng là một yếu tố quan trọng để đảm bảo tính đúng đắn và đáng tin cậy của các kết quả thu được. Trong chương 3 này, tsẽ trình bày chi tiết về thực nghiệm và đánh giá kết quả của hệ thống.

Chương 1. CƠ SỞ LÝ THUYẾT VÀ TỔNG QUAN

1.1. Tổng quan về hệ thống người đại diện ảo

1.1.1. Khái niệm và định nghĩa người đại diện ảo

Người đại diện ảo (Virtual Agent) là chương trình phần mềm sử dụng trí tuệ nhân tạo để tương tác với người dùng qua giao diện hội thoại, mô phỏng khả năng giao tiếp của nhân viên thực. Hệ thống tích hợp công nghệ xử lý ngôn ngữ tự nhiên (NLP), học máy và giọng nói tổng hợp, cho phép hiểu ngôn ngữ tự nhiên và phản hồi tự động 24/7 mà không cần sự can thiệp trực tiếp của con người.

Virtual Agent khác biệt với chatbot đơn giản ở khả năng học hỏi, cải thiện theo thời gian và xử lý tình huống phức tạp dựa trên ngữ cảnh và lịch sử tương tác, không chỉ phụ thuộc vào kịch bản có sẵn.

Tiêu chí	**Hỗ trợ trực tiếp**	**Chatbot đơn giản**	**Virtual Agent**
Thời gian phản hồi	Phụ thuộc nhân viên	Tức thì	Tức thì
Chi phí vận hành	Cao (Lương + đào tạo)	Thấp	Trung bình
Khả năng mở rộng	Hạn chế	Trung bình	Cao
Độ chính xác	Cao (Con người)	Trung bình	Cao (AI)
Khả năng học hỏi	✓ (Kinh nghiệm)	×	✓ (Machine Learning)
Hoạt động 24/7	×	✓	✓

1.1.2. Phân loại và đặc điểm

Theo hình thức thể hiện: Avatar 2D (thiết kế đơn giản, chi phí thấp, phù hợp web/mobile), Avatar 3D (chuyển động tự nhiên, yêu cầu Unity/Unreal, phù hợp kiosk), Digital Human (hình ảnh người thật + AI, độ tin cậy cao, chi phí cao).

Theo khả năng tương tác: Static Agent (hiển thị thông tin có sẵn), Reactive Agent (phản hồi theo kịch bản, keyword matching), Intelligent Agent (tích hợp AI/NLP, hiểu ngữ cảnh, học từ tương tác - lựa chọn tối ưu cho tư vấn phức tạp).

Theo chức năng: Agent tư vấn thông tin (chương trình đào tạo, học phí, thủ tục), Agent hỗ trợ thủ tục (đăng ký học phần, tra cứu kết quả), Agent giải trí/định hướng (hoạt động ngoại khóa, tư vấn nghề nghiệp, tour ảo).

Yêu cầu cho Học viện: Hỗ trợ tiếng Việt tốt (xử lý giọng địa phương), tích hợp CSDL hiện có, hoạt động đa nền tảng (web/mobile/kiosk), cập nhật liên tục với phân quyền truy cập, và chuyển sang tư vấn viên người khi vượt khả năng xử lý.

1.2. Cơ sở lý thuyết về trí tuệ nhân tạo và xử lý ngôn ngữ tự nhiên

1.2.1. Tổng quan về AI và học máy

Trí tuệ nhân tạo (AI) là khả năng máy tính mô phỏng tư duy và hành vi thông minh của con người, bao gồm Machine Learning, Deep Learning, NLP và Computer Vision. Trong Virtual Agent, AI đóng vai trò nhận dạng ý định người dùng, sinh phản hồi tự động và cá nhân hóa trải nghiệm.

Machine Learning gồm ba phương pháp: Supervised Learning (học với dữ liệu đã gán nhãn, phù hợp phân loại câu hỏi về học phí/lịch thi/thủ tục), Unsupervised Learning (phát hiện pattern ẩn, phân nhóm người dùng theo hành vi), Reinforcement Learning (học qua thử-sai, tối ưu chiến lược hội thoại).

Deep Learning sử dụng Transformer Architecture (nền tảng BERT/GPT với cơ chế attention) và Pre-trained Models (đã training trên corpus lớn, cho phép fine-tune với dữ liệu nhỏ - tiết kiệm chi phí, phù hợp ngân sách giáo dục).

1.2.2. Xử lý ngôn ngữ tự nhiên

Tác vụ NLP cơ bản: Tokenization (tách văn bản thành từ/cụm từ), POS Tagging (gán nhãn từ loại), NER (nhận diện tên ngành học/giảng viên/phòng ban), Dependency Parsing (phân tích mối quan hệ giữa các từ).

Tác vụ NLP nâng cao: Intent Recognition (xác định mục đích: hỏi thông tin/yêu cầu hỗ trợ/khiếu nại), Entity Extraction (trích xuất thông tin cụ thể như "Em muốn đăng ký Lập trình Java học kỳ 1" → môn học, hành động, thời gian), Sentiment Analysis (phát hiện cảm xúc để chuyển sang nhân viên khi cần), Text Generation (sinh câu trả lời đa dạng thay vì template), Question Answering (trả lời dựa trên cơ sở tri thức).

Mô hình ngôn ngữ hiện đại: Word Embeddings (Word2Vec/GloVe - từ gần nghĩa có vector gần nhau, hạn chế: mỗi từ một vector duy nhất), Contextual Embeddings (BERT/PhoBERT - vector phụ thuộc ngữ cảnh, PhoBERT tối ưu cho tiếng Việt), Large Language Models (GPT-3/4 - hàng trăm tỷ parameters, tạo nội dung tự nhiên).

1.2.3. Công nghệ xử lý tiếng Việt

Đặc thù tiếng Việt: Ngôn ngữ đơn lập không biến hình, từ có nhiều tiếng gây khó tách từ ("học sinh" vs "học sinh học"), dấu thanh điệu ảnh hưởng nghĩa (ma/mà/má/mã/mả/mạ), từ đồng âm nhiều (bàn: furniture/discuss).

Công cụ xử lý: VnCoreNLP (tách từ, POS tagging, NER, parsing - độ chính xác cao), PhoBERT (tạo biểu diễn ngữ cảnh, fine-tune cho domain cụ thể), VnTokenizer/PyVi/UndertheSea (tách từ và xử lý đa năng).

Thách thức: Đa dạng phương ngôn (Bắc-Trung-Nam), từ viết tắt và tiếng lóng sinh viên (mik/vs/dc/ko), thiếu dữ liệu chuyên ngành giáo dục CNTT, lỗi chính tả và thiếu dấu thanh điệu.

Giải pháp cho Học viện: Xây dựng corpus chuyên ngành PTIT (thuật ngữ, tên môn học, quy định), fine-tune PhoBERT với câu hỏi thực tế sinh viên (bao gồm viết tắt, lỗi), tạo từ điển domain-specific (PTIT/CTSV/ĐHCQ + module chuẩn hóa văn bản), quy trình cải thiện liên tục (ghi log, phân tích lỗi, bổ sung corpus).

1.2.4. Tương tác người-máy và trải nghiệm người dùng

HCI cơ bản: Nghiên cứu thiết kế, đánh giá hệ thống tương tác cho người dùng. Virtual Agent yêu cầu tương tác tự nhiên như nói chuyện với người thật. Hiện tượng "Computers Are Social Actors" (CASA) cho thấy người dùng tự nhiên áp dụng quy tắc xã hội khi tương tác với máy.

Nguyên tắc thiết kế hội thoại (4 nguyên tắc Grice): Số lượng (đủ thông tin), Chất lượng (đúng sự thật), Liên quan (đúng chủ đề), Cách thức (rõ ràng, tránh mơ hồ). Virtual Agent cần tính cách thân thiện, chuyên nghiệp, kiên nhẫn phù hợp môi trường giáo dục.

5 yếu tố usability (Nielsen): Learnability (dễ học), Efficiency (năng suất cao), Memorability (dễ nhớ), Errors (tỷ lệ lỗi thấp), Satisfaction (hài lòng).

UX Honeycomb (Morville): Useful, Usable, Desirable, Findable, Accessible, Credible, Valuable.

Đo lường UX: Khách quan (task completion rate, time on task, error rate), Chủ quan (System Usability Scale - SUS >68 = tốt, Chatbot Usability Questionnaire, Net Promoter Score - NPS).

Thiết kế đa kênh & accessibility: Omnichannel (chuyển liền mạch giữa web/mobile/kiosk), Mobile-first (ưu tiên tính năng quan trọng), WCAG 2.1 (Perceivable, Operable, Understandable, Robust - hỗ trợ screen readers, voice input, độ tương phản).

Niềm tin & minh bạch: Rõ ràng là hệ thống tự động, cung cấp nguồn thông tin ("Theo quy chế 2024..."), thừa nhận khi không chắc chắn, giải thích yêu cầu thông tin cá nhân.

Quy trình thiết kế: User-Centered Design (hiểu ngữ cảnh → xác định yêu cầu → tạo prototype → đánh giá). Testing với 5 người dùng phát hiện ~85% vấn đề usability.

1.3. Ứng dụng trong hoạt động tư vấn lễ tân

1.3.1. Tổng quan hệ thống

Phát triển Virtual Agent thông minh hỗ trợ tư vấn sinh viên PTIT 24/7 với 4 mục tiêu: (1) Giảm 70-80% câu hỏi lặp lại cho nhân viên, (2) Phản hồi <2 giây, (3) Nâng cao trải nghiệm sinh viên, (4) Thu thập insights người dùng.

1.3.2. Công nghệ và kiến trúc

Hệ thống theo mô hình Client-Server với 3 thành phần chính:

Frontend: Xây dựng bằng **ReactJS** và **JavaScript** cho giao diện web tương tác cao. Sử dụng **Electron.js** để đóng gói thành desktop app chạy trên Windows, macOS và Linux. Giao diện responsive, hoạt động mượt mà trên mọi thiết bị.

Backend: Phát triển hoàn toàn bằng **Python** với các thư viện AI/ML:

- PhoBERT (transformers, PyTorch): Nhận diện ý định và xử lý tiếng Việt
- spaCy, PyVi, UndertheSea: Tách từ, NER, xử lý ngôn ngữ tự nhiên
- SQLAlchemy: Kết nối database PostgreSQL
- Redis: Lưu trữ context và session
- Elasticsearch: Tìm kiếm knowledge base

Communication**:** Frontend và Backend giao tiếp real-time qua **WebSocket** sử dụng thư viện **Socket.IO** (cả JavaScript client và Python server). Đảm bảo tương tác tức thì, bidirectional communication và auto-reconnect.

Luồng xử lý (ví dụ "Học phí CNTT?"): Frontend gửi → Backend: Preprocessing (PyVi - chuẩn hóa, tách từ) → Intent Recognition (PhoBERT - "hỏi_học_phí") → Entity Extraction ("ngành CNTT") → Query DB (PostgreSQL) → Generate Response → Frontend hiển thị. Tổng: <1 giây.

1.3.3. Tính năng và khả năng

5 chức năng cốt lõi:

Tư vấn thông tin: Chương trình đào tạo, học phí/học bổng, lịch học/thi, quy chế, cơ sở vật chất

Hướng dẫn thủ tục: Đăng ký học phần, xin giấy tờ (bảng điểm, xác nhận SV), khiếu nại điểm, bảo lưu/chuyển trường, đăng ký ký túc xá

Tra cứu cá nhân (sau xác thực): Kết quả học tập, điểm thi, lịch thi cá nhân, công nợ học phí

Tư vấn định hướng: Chuyên ngành, học phần tự chọn, hướng nghiệp CNTT-Viễn thông, cơ hội thực tập/việc làm

Hỗ trợ kỹ thuật: Khôi phục mật khẩu, hướng dẫn hệ thống quản lý đào tạo, khắc phục sự cố đăng nhập

Tính năng mở rộng: Đa ngôn ngữ (Việt/Anh), đa phương thức (text/voice/multimedia), tích hợp hệ thống (SSO, real-time sync), cá nhân hóa (dựa trên profile/lịch sử), chuyển giao thông minh (sang nhân viên khi cần).

1.4. Kết chương

Chương 1 xây dựng nền tảng lý thuyết với 4 trụ cột:

Về Virtual Agent: Phân tích khái niệm, phân loại và xác định Intelligent Agent là lựa chọn phù hợp cho PTIT với khả năng xử lý NLP tiếng Việt, tích hợp đa nền tảng và cá nhân hóa.

Về công nghệ AI/NLP: Hệ thống hóa kiến thức về Machine Learning, Deep Learning và đặc biệt giải pháp xử lý tiếng Việt qua PhoBERT, VnCoreNLP, PyVi kết hợp corpus chuyên ngành PTIT.

Về HCI/UX: Trình bày nguyên lý thiết kế giao diện hội thoại, yếu tố tâm lý học nhận thức, phương pháp đánh giá UX - tạo cơ sở thiết kế tối ưu.

Về ứng dụng: Đề xuất kiến trúc ReactJS + Electron.js (Frontend), Python + PhoBERT + Vietnamese NLP (Backend), WebSocket (real-time), tích hợp CSDL PTIT với 5 nhóm chức năng cốt lõi.

Chương 1 chứng minh tính khả thi công nghệ qua việc xác định công cụ cụ thể (PhoBERT, PyVi) và yêu cầu xây dựng corpus chuyên ngành. Việc sử dụng các framework đã được kiểm chứng giúp tiết kiệm thời gian phát triển và giảm rủi ro.

Chương 2 sẽ phân tích yêu cầu hệ thống chi tiết qua khảo sát nhu cầu sinh viên/cán bộ, xác định use cases, yêu cầu chức năng và phi chức năng - tạo cầu nối giữa lý thuyết và thiết kế/triển khai.

Chương 2. PHÂN TÍCH THIẾT KẾ VÀ XÂY DỰNG HỆ THỐNG

2.1. Phân tích yêu cầu hệ thống

2.1.1. Phân tích các bên liên quan

Việc xác định đầy đủ các bên liên quan và nhu cầu của họ là bước quan trọng đầu tiên để đảm bảo hệ thống Virtual Agent được thiết kế đúng hướng. Phân tích này giúp hệ thống đáp ứng được kỳ vọng của tất cả đối tượng sử dụng tại Học viện.

Nhóm sinh viên đang học là đối tượng chính với khoảng 15,000+ người thuộc thế hệ Gen Z có kỹ năng công nghệ tốt và quen với smartphone. Nhu cầu của họ tập trung vào tra cứu nhanh lịch học, lịch thi, kết quả học tập, hướng dẫn thủ tục hành chính như đăng ký học phần và xin giấy tờ, thông tin học phí và học bổng, cùng giải đáp về quy chế. Họ kỳ vọng phản hồi tức thì dưới 3 giây với thông tin chính xác, giao diện mobile-friendly và hỗ trợ 24/7 đặc biệt trong cao điểm đăng ký học phần.

Nhóm sinh viên mới và thí sinh tuyển sinh chưa quen với Học viện và cần thông tin định hướng ban đầu về tuyển sinh, ngành học, điểm chuẩn, học phí, thủ tục nhập học, ký túc xá, cơ sở vật chất và tư vấn chọn ngành nghề nghiệp.

Sinh viên quốc tế tuy số lượng nhỏ nhưng đang tăng, đối mặt rào cản ngôn ngữ và văn hóa, cần hỗ trợ tiếng Anh và thông tin về visa, lưu trú cùng hướng dẫn thích nghi văn hóa Việt Nam.

2.1.2. Phân tích yêu cầu chức năng

Dựa trên phân tích các bên liên quan, hệ thống được xác định với hai nhóm chức năng chính.

Nhóm chức năng cốt lõi bao gồm bốn yêu cầu chính. Chức năng hỏi đáp thông tin tổng quát (CF01) cung cấp thông tin về Học viện, chương trình đào tạo, quy định thông qua tra cứu ngành học, cơ sở vật chất, câu lạc bộ, lịch học kỳ và quy chế đào tạo. Sinh viên đặt câu hỏi bằng ngôn ngữ tự nhiên (tiếng Việt hoặc

Anh), hệ thống phản hồi bằng text hoặc audio kèm link tham khảo và hình ảnh minh họa nếu có.

Chức năng hướng dẫn thủ tục hành chính (CF02) hướng dẫn sinh viên các quy trình đăng ký học phần, xin giấy tờ, khiếu nại điểm, bảo lưu/nghỉ học/chuyển trường và đăng ký ký túc xá. Khi được yêu cầu về thủ tục cụ thể, hệ thống cung cấp hướng dẫn từng bước chi tiết, danh sách giấy tờ, thời hạn, địa điểm nộp, form mẫu tải xuống và link hệ thống online.

Chức năng tư vấn học tập và định hướng (CF03) hỗ trợ quyết định học tập qua tư vấn chọn học phần tự chọn, chọn chuyên ngành, gợi ý lộ trình học tập, thông tin thực tập và việc làm, định hướng nghề nghiệp. Chức năng này yêu cầu thông tin ngữ cảnh cá nhân nhưng chỉ cung cấp thông tin để sinh viên tự quyết định chứ không đưa ra quyết định thay họ.

Chức năng hỗ trợ kỹ thuật cơ bản (CF04) giải quyết các vấn đề phổ biến như reset password, lỗi đăng nhập, hướng dẫn sử dụng portal và email, troubleshooting WiFi và cài đặt phần mềm học tập.

Nhóm chức năng hỗ trợ bao gồm quản lý ngữ cảnh hội thoại (CF06), theo dõi và duy trì ngữ cảnh qua nhiều lượt hội thoại bằng cách lưu trữ lịch sử 10 lượt gần nhất, hiểu đại từ tham chiếu, xử lý nhiều ý định trong một câu và yêu cầu làm rõ khi câu hỏi mơ hồ. Ví dụ, sau khi trả lời "Học phí năm 2024-2025 là 12 triệu đồng/năm", câu hỏi "Còn năm sau?" sẽ được hiểu là hỏi về năm 2025-2026 dựa trên ngữ cảnh.

2.1.3. Phân tích yêu cầu phi chức năng

Yêu cầu phi chức năng xác định cách thức hệ thống hoạt động về hiệu suất, bảo mật, khả năng sử dụng và bảo trì.

Về hiệu suất, hệ thống yêu cầu thời gian phản hồi trung bình không quá 2 giây, với 95% yêu cầu xử lý trong 3 giây và 99% trong 5 giây. Thời gian này được đo end-to-end bao gồm network và processing, có giám sát realtime với

cảnh báo khi vượt ngưỡng. Giải pháp bao gồm cache cho câu hỏi thường gặp, load balancing, tối ưu database queries và sử dụng CDN cho static content.

Độ sẵn sàng đạt uptime 99.5% tương đương 3.6 giờ downtime mỗi tháng. Bảo trì có kế hoạch được thông báo 24 giờ trước và thực hiện ngoài giờ cao điểm với Mean Time To Recovery không quá 30 phút. Biện pháp đảm bảo gồm redundancy với nhiều server, health checks tự động, cơ chế failover và backup hàng ngày.

Về bảo mật, dữ liệu nhạy cảm như điểm thi và thông tin cá nhân được mã hóa AES-256 trong database. Toàn bộ communication sử dụng HTTPS bắt buộc. Password được hash bằng bcrypt hoặc Argon2, không lưu plaintext. Thông tin cá nhân được ẩn danh trong logs. Hệ thống tuân thủ Nghị định 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân và OWASP Top 10.

Về khả năng sử dụng, giao diện responsive hoạt động tốt trên desktop, tablet, mobile với font size tối thiểu 14px, contrast ratio 4.5:1 theo WCAG AA, hỗ trợ dark mode và không yêu cầu plugin. Thiết kế đơn giản, rõ ràng với tối đa 3 cú nhấp đến thông tin quan trọng. Hệ thống hỗ trợ web app trên Chrome, Firefox, Safari, Edge (2 phiên bản gần nhất), mobile app cho iOS 14+ và Android 8+, tích hợp Facebook Messenger và Zalo trong Phase 2.

Về khả năng bảo trì, code quality đạt coverage tối thiểu 70%, tuân thủ PEP 8 cho Python và ESLint cho JavaScript, bắt buộc code review và có tài liệu đầy đủ. Khả năng debug được hỗ trợ qua error messages chi tiết, stack trace đầy đủ trong logs, monitoring dashboard và alerting khi bất thường. Deployment sử dụng CI/CD pipeline với automated testing, blue-green deployment cho zero downtime và rollback trong 5 phút. Kiến trúc theo mô hình microservices hoặc modular monolith với loose coupling, high cohesion, API-first design.

2.2. Thiết kế kiến trúc hệ thống

2.2.1. Kiến trúc tổng quan

Hệ thống được thiết kế theo mô hình Client-Server kết hợp ba lớp mô-đun AI chính là Speech-to-Text (ASR), Large Language Model (LLM) và Text-to-Speech (TTS), tạo thành một kiến trúc đa tầng linh hoạt và có khả năng mở rộng.

Tầng Frontend bao gồm giao diện web và desktop, chịu trách nhiệm tương tác với người dùng thông qua việc hiển thị giao diện, phát âm thanh, hiển thị avatar Live2D và gửi/nhận sự kiện qua WebSocket. Tầng WebSocketServer là backend ASGI nhận kết nối từ client, xử lý luồng dữ liệu thời gian thực và quản lý session cùng chat context cho từng người dùng.

Config Manager đảm nhiệm việc đọc và validate file cấu hình conf.yaml, đồng bộ và merge các cấu hình từ nhiều nguồn thông qua các script merge_configs.py và upgrade.py. Model Manager với Adapter Layer tạo ra interface chung để tải và chạy các mô-đun ASR, LLM và TTS, ẩn đi chi tiết triển khai cụ thể là local model hay API.

Tầng xử lý AI core gồm ba module chính. Module ASR chuyển đổi audio thành text, module LLM/Agent xử lý ngữ cảnh và prompt để trả về text kèm internal thoughts và actions, module TTS chuyển text thành audio với hỗ trợ voice cloning và chuyển đổi ngôn ngữ cho giọng nói. Live2D Controller ánh xạ cảm xúc và intent sang biểu cảm và animation của avatar Live2D dựa trên file mapping model_dict.json.

Visual Perception là module tùy chọn xử lý camera hoặc screen capture, trích xuất features và đưa vào context của LLM. Hệ thống Logging & Persistence lưu trữ chat logs, debug logs và health logs để phục vụ phân tích và cải thiện.

2.2.2. Luồng xử lý chính

- Luồng Khởi tạo kết nối & session
- Luồng: Ghi âm ở client -> ASR (streaming) -> partial + final transcript

- Luồng: From final transcript -> build prompt -> gọi LLM (streaming tokens)
- Luồng: TTS synthesis -> gửi audio tới client -> Live2D sync
- Luồng: Lưu trữ hội thoại & logging & backup config
- Luồng: Xử lý lỗi / fallback (ASR, LLM, TTS)
- Luồng: Proactive speaking (agent-triggered)

2.2.3. Thiết kế chi tiết các module

- Mô-đun ASR (STT): Module ASR đóng vai trò “tai” của hệ thống, chuyển đổi giọng nói sinh viên thành văn bản để xử lý. Mô đun gồm những tính năng chính sau:
    - Real-time Streaming
    - Phát hiện điểm kết thúc và im lặng
    - Điểm tin cậy và kết quả từng phần
    - Chế độ "AI không nghe TTS"
- Mô-đun LLM / Agent

Module LLM (Large Language Model) là "bộ não" trung tâm của hệ thống người đại diện ảo, đảm nhận vai trò chính trong việc tư vấn và cung cấp thông tin cho sinh viên.

Mô-đun gồm những vai trò chính sau:

- Hiểu và xử lý câu hỏi của sinh viên
- Tư vấn và cung cấp thông tin
- Ra quyết định và điều phối hệ thống
- Cá nhân hóa trải nghiệm
- Suy luận và giải quyết vấn đề

- Mô-đun TTS

Module TTS đóng vai trò **"giọng nói"** của hệ thống người đại diện ảo, chuyển đổi câu trả lời dạng văn bản thành giọng nói tự nhiên để giao tiếp với sinh viên

- Live 2D & Visual Perception

  Chức năng chính của mô-dun là để xử lý hình ảnh, hiển thị thông tin phản hồi lại người dùng qua hình ảnh 2D

2.3. Thiết kế thuật toán và mô hình AI

Streaming + partial results: hỗ trợ trả transcript/tokens từng phần để UX realtime.

Interrupt handling: khi TTS đang phát, microphone không gửi input tới ASR, hoặc hệ thống lọc audio do TTS để tránh vòng lặp.

Prompt management & context window: LLM adapter chịu trách nhiệm quản lý history, cắt giảm context theo token budget của backend.

Emotion mapping: LLM hoặc sentiment detector xuất thông tin cảm xúc -> ánh xạ motion/expr Live2D.

Config merge & upgrade: merge_configs.py đảm nhiệm merge an toàn user config với template new version, ghi backup, log các key mới/hết hạn.

2.4. Cài đặt và phát triển hệ thống

Hệ thống được xây dựng dựa trên open source **Open-LLM-Vtuber**, từ đó thiết kế hệ thống sao cho phù hợp với sinh viên Học viện Công nghệ Bưu chính Viễn thông. Hệ thống sử dụng các ngôn ngữ phát triển phổ biến như **Python** bởi vì Python là 1 trong những ngôn ngữ khá phổ biến và có thể kết hợp được nhiều thư viện AI hiện có. Ngoài ngôn ngữ Python để phát triển Backend thì tôi có sự dụng ngôn ngữ **Javascript** và các framework như **Reactjs**, **electronjs** để phát triển Frontend mang đến giao diện chuyên nghiệp và trực quan tới người dùng.

Ngoài ra để kiểm thử hệ thống trước khi tới người dùng tôi còn sử dụng phầm mềm kiểm thử **Jmeter** để thực hiện mô phỏng khoảng 1000 user cùng vào trang web để kiểm tra khả năng chịu tải của hệ thống.

2.5. Kết chương

Chương 2 trình bày quá trình phân tích, thiết kế và xây dựng hệ thống Virtual Agent qua ba giai đoạn chính:

Phân tích yêu cầu: Xác định các bên liên quan và nhu cầu cụ thể, bao gồm yêu cầu chức năng (chào hỏi, cung cấp thông tin, tư vấn) và phi chức năng (hiệu suất, độ tin cậy, bảo mật), tạo nền tảng cho thiết kế hệ thống.

Thiết kế kiến trúc: Triển khai mô hình 3-tier với các module chức năng rõ ràng (xử lý đầu vào, nhận diện ý định, quản lý hội thoại, sinh phản hồi, giao diện). Cơ sở dữ liệu được tối ưu hóa cho khả năng truy vấn và mở rộng.

Thiết kế AI và thuật toán: So sánh lựa chọn mô hình học máy, xây dựng giải pháp NLP cho tiếng Việt, và phát triển thuật toán quản lý hội thoại thông minh làm cơ sở cho cài đặt hệ thống.

Kết quả chương này cung cấp cơ sở kỹ thuật hoàn chỉnh cho triển khai thực tế. Chương 3 sẽ đánh giá hiệu năng và kết quả thử nghiệm hệ thống.

Chương 3. THỬ NGHIỆM VÀ ĐÁNH GIÁ

3.1. Tập dữ liệu

Tôi sử dụng tập dữ liệu là các câu hỏi và câu trả lời trong các thủ tục hành chính,…. Với tổng dữ liệu sấp sỉ 1600 dữ liệu bao gồm 424 dữ liệu về thông tin, 1120 dữ liệu về câu hỏi/trả lời, với các chủ đề chính như: Thông tin về học viện, các ngành học, tuyển sinh, các câu lạc bộ, các mẫu đơn,.. Với cách thức thu thập từ các trang chính thống của trưởng, sổ tay sinh viên.

3.2. Môi trường thử nghiệm

Hệ thống được chạy trên hệ điều hành Ubuntu vào được deploy lên hosting để chạy thử nghiệm.

Thông số	Giá trị
CPU	Xeon Silver 4309
GPU	Nvidia Quadro P4000
Bộ nhớ RAM	32 GB
Hệ điều hành	Ubuntu
Python	3.9

3.3. Xây dựng ứng dụng thử nghiệm

3.3.1. Giới thiệu về hệ thống

Ứng dụng được xây dựng dưới dạng trang web, với ngôn ngữ lập trình được lựa chọn là Python xử lý Backend và Reactjs và electronjs để xử lý FrontEnd. Tôi đã lựa chọn Python bởi ngôn ngữ này có nhiều thư viện hỗ trợ tốt cho việc xử lý âm thanh, hình ảnh hỗ trợ tốt cho ứng dụng.

3.3.2. Xây dựng hệ thống

Ứng dụng sử dụng Python bao gồm nhiều thư mục với các mục đích sử dụng khác nhau. Phần chính bao gồm các thư mục data bên trong chứa các tệp dữ liệu âm thanh mẫu. Ngoài ra, các thư mục khác chứa các tệp giao diện .html,

javascript, view.py, url.py…Ngoài ra tôi cũng xây dựng giao diện trên trang web để có thể tương thích với desktop và mobile.

3.4. Kết quả thử nghiệm và đánh giá

3.4.1. Thiết lập thử nghiệm

Mục tiêu thử nghiệm

- Đánh giá độ chính xác của hệ thống trong việc nhận diện ý định và trả lời câu hỏi của sinh viên
- Kiểm tra hiệu suất và thời gian phản hồi của hệ thống trong các tình huống khác nhau
- Đo lường mức độ hài lòng của người dùng đối với trải nghiệm tương tác
- Xác định các lỗi, hạn chế và điểm cần cải thiện của hệ thống
- Đánh giá khả năng xử lý đồng thời nhiều yêu cầu và tính ổn định của hệ thống

Phương pháp thử nghiệm

Thử nghiệm chức năng (Functional Testing):

- Kiểm tra từng chức năng theo use case đã thiết kế
- Sử dụng phương pháp black-box testing
- Ghi nhận các lỗi, bug và hành vi bất thường

Thử nghiệm hiệu năng (Performance Testing):

- Load testing: Mô phỏng [X] người dùng đồng thời
- Stress testing: Kiểm tra giới hạn chịu tải của hệ thống
- Đo lường thời gian phản hồi trung bình, thời gian xử lý tối đa

Kịch bản thử nghiệm

Kịch bản 1: Tương tác cơ bản

- Người dùng đặt câu hỏi đơn giản về thông tin cơ bản (giờ làm việc, địa chỉ phòng ban)
- Hệ thống phản hồi thông tin chính xác
- Đánh giá: Thời gian phản hồi, độ chính xác câu trả lời

Kịch bản 2: Tương tác phức tạp

- Người dùng đặt câu hỏi yêu cầu kết hợp nhiều thông tin (ví dụ: "Em muốn đăng ký học lại môn X, phải làm thế nào và mất bao nhiêu tiền?")
- Hệ thống cần hiểu đa ý định và trả lời đầy đủ
- Đánh giá: Khả năng hiểu ngữ cảnh, tính đầy đủ của câu trả lời

Kịch bản 3: Xử lý hội thoại nhiều lượt

- Người dùng đặt câu hỏi ban đầu, sau đó hỏi thêm các câu liên quan
- Hệ thống cần duy trì ngữ cảnh hội thoại
- Đánh giá: Context management, tính mạch lạc của cuộc hội thoại

Kịch bản 4: Xử lý tình huống ngoại lệ

- Người dùng đặt câu hỏi ngoài phạm vi kiến thức
- Người dùng nhập input không hợp lệ hoặc spam
- Đánh giá: Khả năng xử lý lỗi, thông báo phù hợp cho người dùng

Kịch bản 5: Kiểm tra tải đồng thời

- Mô phỏng nhiều người dùng truy cập đồng thời trong giờ cao điểm
- Đánh giá: Tốc độ phản hồi, tỷ lệ lỗi, độ ổn định hệ thống

3.4.2. Giao diện hệ thống

Hệ thống bao gồm các giao diện chính như: Giao diện trang chủ, giao diện khung chat, giao diện thanh công cụ, giao diện cài đặt, giao diện thanh công cụ biểu cảm.

Giao diện trang chủ: Cung cấp cái nhìn tổng quan của hệ thống

Giao diện khung chat: Giúp người dùng có thể nhập text thay vì nói chuyên trực tiếp trong trường hợp không thể nói được, và có thể tuỳ chỉnh bật tắt mic.

Giao diện thanh công cụ: Cung cấp một số công cụ tuỳ chỉnh hệ thống.

Giao diện cài đặt: Người dùng có thể cấu hình cơ bản cho hệ thống với tuỳ chỉnh cá nhân như ảnh nền.

Giao diện thanh công cụ biểu cảm: Mục đích để giúp tăng tương tác người dùng với hệ thống qua tương tác biểu cảm.

3.5. Kết chương

Chương 3 đã đi vào chi tiết thực nghiệm đề tài, việc lựa chọn tệp dữ liệu đào tạo sẽ quyết định sự thành công của mỗi mô hình. Chính vì vậy tôi đã lựa chọn kỹ lưỡng mô hình và bộ dữ liệu thử nghiệm. Trong chương này cũng đã đi sâu vào phân tích và đánh giá mô hình, các thức thiết lập thử nghiệm, kết quả đạt được và so sánh kết quả đó với các nghiên cứu khác. Thông qua đó sẽ định hướng tiếp theo cho sự mở rộng của đề án.

KẾT LUẬN

Đề án "***Nghiên cứu phát triển hệ thống người đại diện ảo trong hoạt động tư vấn lễ tân, cung cấp thông tin cho sinh viên Học viện Công nghệ Bưu chính Viễn thông***" đã hoàn thành các mục tiêu đề ra, bao gồm nghiên cứu cơ sở lý thuyết, phân tích thực trạng và thiết kế thành công hệ thống với các chức năng đáp ứng yêu cầu nghiệp vụ.

Hệ thống đã được xây dựng với kiến trúc hợp lý, tích hợp công nghệ AI hiện đại và giao diện thân thiện. Qua triển khai thử nghiệm, hệ thống đã hỗ trợ sinh viên tra cứu thông tin hiệu quả, giảm tải công việc cho bộ phận lễ tân và nâng cao chất lượng dịch vụ hỗ trợ sinh viên 24/7.

Về ý nghĩa, đề án góp phần làm phong phú nghiên cứu về ứng dụng AI trong giáo dục đại học Việt Nam, đồng thời chứng minh hiệu quả trong việc hiện đại hóa công tác quản lý sinh viên, cải thiện trải nghiệm người học và tạo tiền đề phát triển các hệ thống thông minh khác.

Hạn chế của hệ thống bao gồm cơ sở dữ liệu cần cập nhật thường xuyên, khả năng xử lý câu hỏi phức tạp còn hạn chế và chưa tích hợp đầy đủ với các hệ thống thông tin khác của Học viện.

Hướng phát triển tiếp theo bao gồm mở rộng cơ sở dữ liệu, bổ sung nhận dạng giọng nói và đa ngôn ngữ (ngắn hạn); kết nối với các nền tảng quản lý đào tạo và phát triển ứng dụng di động (trung hạn); ứng dụng công nghệ metaverse và phát triển khả năng học tự động (dài hạn).

Đề án đã chứng minh tính khả thi của việc ứng dụng công nghệ người đại diện ảo vào hoạt động tư vấn sinh viên, hứa hẹn trở thành công cụ quan trọng góp phần xây dựng môi trường giáo dục thông minh và hiện đại.