

**XÂY DỰNG HỆ THỐNG AI CHATBOT
KẾT HỢP PHÂN TÍCH DỰ BÁO DỮ LIỆU CÔNG TRỰC TUYẾN:
NGHIÊN CỨU TẠI TỈNH THỪA THIÊN HUẾ**

Bùi Quang Vũ^{1*}, Phạm Quang Trí², Lê Văn Mẫn², Nguyễn Kim Tùng³

¹ Khoa Toán, Trường Đại học Khoa học, Đại học Huế

² Văn phòng Ủy ban, tỉnh Thừa Thiên Huế

³ Sở Khoa học Công nghệ, tỉnh Thừa Thiên Huế

* Email: buiquangvu@husc.edu.vn

Ngày nhận bài: 21/9/2024; ngày hoàn thành phần biên: 17/10/2024; ngày duyệt đăng: 01/11/2024

TÓM TẮT

Dịch vụ công trực tuyến (DVCTT) đóng vai trò quan trọng trong tiến trình xây dựng chính quyền số, góp phần nâng cao hiệu quả quản lý nhà nước và chất lượng phục vụ người dân. Mặc dù các nền tảng DVCTT đã được triển khai rộng rãi, việc hỗ trợ người dân tiếp cận, khai thác và sử dụng hiệu quả các dịch vụ này vẫn còn nhiều hạn chế, đặc biệt trong khâu hướng dẫn thủ tục và tương tác. Bài báo này đề xuất và đánh giá một hệ thống hỗ trợ người dân dựa trên công nghệ trí tuệ nhân tạo (AI), bao gồm hai hợp phần chính: (i) hệ thống AI Chatbot hỗ trợ hướng dẫn thủ tục, tra cứu thông tin và tương tác với người dân; (ii) hệ thống phân tích và dự báo dữ liệu DVCTT nhằm hỗ trợ công tác quản lý và ra quyết định. Chatbot được xây dựng trên nền tảng RASA với mô hình xử lý ngôn ngữ tự nhiên phù hợp cho bối cảnh hành chính công tiếng Việt, trong khi hệ thống phân tích dữ liệu áp dụng các mô hình ARIMA và LSTM cho bài toán dự báo chuỗi thời gian. Kết quả đánh giá trong môi trường thực tế cho thấy hệ thống hoạt động ổn định, góp phần cải thiện trải nghiệm người dân và hỗ trợ hiệu quả công tác quản lý DVCTT.

Từ khóa: AI Chatbot, dịch vụ công trực tuyến, xử lý ngôn ngữ tự nhiên, phân tích dữ liệu, ARIMA, LSTM.

1. MỞ ĐẦU

Chuyển đổi số đang trở thành xu thế tất yếu trong hoạt động quản lý nhà nước, trong đó xây dựng chính phủ số và chính quyền số là một trong những nhiệm vụ trọng tâm của Việt Nam trong giai đoạn hiện nay. Dịch vụ công trực tuyến (DVCTT) được xem là một trong những thành phần cốt lõi của chính quyền số, cho phép người dân và doanh

nghiệp thực hiện các thủ tục hành chính một cách thuận tiện, minh bạch và hiệu quả [1],[3]. Việc tích hợp và cung cấp dịch vụ công trực tuyến thông qua Cổng Dịch vụ công quốc gia đã góp phần chuẩn hóa quy trình và tạo điều kiện thuận lợi cho người dân tiếp cận các dịch vụ hành chính [2]. Tuy nhiên, trên thực tế, việc tiếp cận và sử dụng DVCTT vẫn gặp nhiều khó khăn do quy trình thủ tục phức tạp, thiếu hướng dẫn trực quan và hạn chế trong tương tác giữa người dân với hệ thống.

Tại tỉnh Thừa Thiên Huế, hệ thống Cổng dịch vụ công trực tuyến đã được triển khai đồng bộ và đạt được nhiều kết quả tích cực. Tuy nhiên, thực tế cho thấy người dân vẫn gặp khó khăn trong việc tiếp cận thông tin thủ tục hành chính, tra cứu hồ sơ và sử dụng các dịch vụ trực tuyến, đặc biệt đối với các thủ tục phức tạp. Bên cạnh đó, hệ thống hiện tại mới dừng lại ở mức thống kê mô tả dữ liệu, chưa khai thác hiệu quả các kỹ thuật phân tích và dự báo nhằm hỗ trợ công tác quản lý và ra quyết định.

Nhiều nghiên cứu gần đây cho thấy trí tuệ nhân tạo, đặc biệt là các hệ thống chatbot và phân tích dữ liệu thông minh, có tiềm năng lớn trong việc cải thiện khả năng tương tác và khai thác dữ liệu trong các hệ thống dịch vụ công. Tuy vậy, phần lớn các nghiên cứu hiện nay tập trung vào các giải pháp riêng lẻ, chưa xem xét đồng thời việc hỗ trợ người dùng và khai thác dữ liệu DVCTT cho mục tiêu quản lý và ra quyết định.

Mặc dù chatbot đã được nghiên cứu và ứng dụng trong nhiều lĩnh vực công trên thế giới, phần lớn các nghiên cứu tập trung vào hỗ trợ tương tác người dùng, trong khi việc kết hợp chatbot với phân tích và dự báo dữ liệu dịch vụ công nhằm hỗ trợ quản lý vẫn còn hạn chế, đặc biệt trong bối cảnh các nước đang phát triển. Các báo cáo quốc tế cũng chỉ ra rằng, mặc dù các nền tảng dịch vụ công số đã được triển khai rộng rãi, việc hỗ trợ tương tác và khai thác dữ liệu phục vụ quản lý vẫn còn là thách thức đối với nhiều quốc gia [10]. Xuất phát từ khoảng trống nghiên cứu này, bài báo đề xuất một hệ thống tích hợp AI Chatbot và mô hình phân tích – dự báo dữ liệu DVCTT, được đánh giá thông qua một nghiên cứu trường hợp tại tỉnh Thừa Thiên Huế, một địa phương tiên phong trong chuyển đổi số tại Việt Nam. Đóng góp chính của nghiên cứu bao gồm:

- (i) Xây dựng AI Chatbot dựa trên nền tảng RASA phù hợp với tiếng Việt hành chính công;
- (ii) Xây dựng bộ dữ liệu hỏi – đáp và kịch bản hội thoại phục vụ DVCTT;
- (iii) Ứng dụng các mô hình ARIMA và LSTM để phân tích và dự báo dữ liệu DVCTT;
- (iv) Triển khai và đánh giá hệ thống trong môi trường thực tế.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Phương pháp nghiên cứu được xây dựng nhằm đánh giá khả năng ứng dụng trí tuệ nhân tạo trong việc hỗ trợ người dân sử dụng dịch vụ công trực tuyến, đồng thời khai thác dữ liệu phát sinh để phục vụ công tác quản lý và ra quyết định. Nghiên cứu áp dụng cách tiếp cận nghiên cứu ứng dụng kết hợp nghiên cứu trường hợp (applied research & case study), trong đó các giải pháp được thiết kế, triển khai và đánh giá trong một môi trường thực tế.

2.1. Cách tiếp cận nghiên cứu và phạm vi nghiên cứu

Nghiên cứu này tiếp cận bài toán theo hướng tích hợp, kết hợp giữa hỗ trợ người dùng cuối và khai thác dữ liệu nghiệp vụ. Cụ thể, nghiên cứu tập trung vào hai câu hỏi chính:

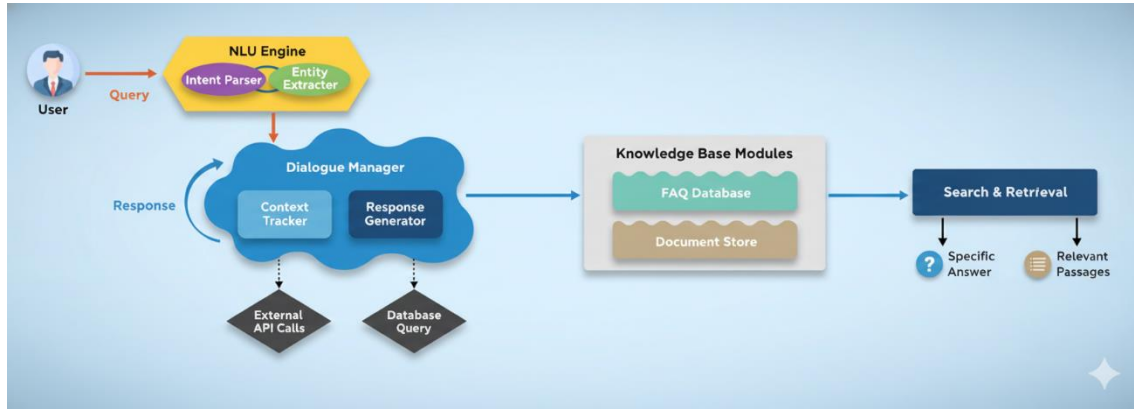
- i) AI Chatbot có thể hỗ trợ người dân tiếp cận và sử dụng dịch vụ công trực tuyến hiệu quả đến mức nào?
- ii) Dữ liệu phát sinh từ quá trình cung cấp dịch vụ công trực tuyến có thể được khai thác như thế nào để hỗ trợ công tác quản lý và ra quyết định?

Phạm vi nghiên cứu được giới hạn trong hệ thống dịch vụ công trực tuyến tại tỉnh Thừa Thiên Huế, với dữ liệu và quy trình nghiệp vụ thực tế. Việc lựa chọn nghiên cứu trường hợp cho phép đánh giá trực tiếp hiệu quả của giải pháp trong bối cảnh triển khai cụ thể, đồng thời rút ra các nhận xét có khả năng khái quát cho các địa phương khác.

2.2. Kiến trúc tổng thể của hệ thống

Hệ thống được thiết kế theo kiến trúc nhiều lớp, bao gồm: (i) lớp giao diện người dùng, (ii) lớp xử lý ngôn ngữ tự nhiên và tương tác, (iii) lớp xử lý nghiệp vụ và phân tích dữ liệu, và (iv) lớp dữ liệu. Cách tổ chức này nhằm đảm bảo tính mô-đun, khả năng mở rộng và dễ dàng tích hợp với các nền tảng dịch vụ công trực tuyến hiện có.

Lớp giao diện người dùng cho phép người dân tương tác với hệ thống thông qua văn bản hoặc giọng nói trên nền tảng web và ứng dụng di động. Lớp xử lý ngôn ngữ tự nhiên chịu trách nhiệm phân tích truy vấn, xác định ý định và trích xuất thông tin cần thiết. Lớp xử lý nghiệp vụ thực hiện các tác vụ truy xuất dữ liệu, xử lý logic và kết nối với các hệ thống dịch vụ công. Cuối cùng, lớp dữ liệu lưu trữ và quản lý dữ liệu hội thoại, dữ liệu hồ sơ dịch vụ công và dữ liệu phục vụ phân tích. Kiến trúc này cho phép hệ thống dễ dàng mở rộng và tích hợp với các nền tảng DVCTT hiện có. Hình 1 minh họa kiến trúc tổng quan của hệ thống AI Chatbot được đề xuất. Các kỹ thuật xử lý ngôn ngữ tự nhiên được áp dụng trong nghiên cứu này dựa trên các nguyên lý và mô hình cơ bản của lĩnh vực xử lý ngôn ngữ tự nhiên hiện đại [4], [5]



Hình 1. Kiến trúc tổng thể của hệ thống AI Chatbot.

2.3. Phương pháp xây dựng và huấn luyện AI Chatbot

Hệ thống AI Chatbot được xây dựng trên nền tảng RASA, một framework mã nguồn mở cho xử lý ngôn ngữ tự nhiên và quản lý hội thoại, đã được áp dụng trong nhiều nghiên cứu và ứng dụng thực tế [6]. Trong nghiên cứu này, chatbot được thiết kế để xử lý các truy vấn tiếng Việt trong bối cảnh hành chính công, nơi ngôn ngữ sử dụng mang tính đặc thù và thường không đồng nhất.

Dữ liệu huấn luyện chatbot được xây dựng từ tập hợp các câu hỏi thường gặp và nội dung hướng dẫn thủ tục hành chính. Các câu hỏi được chuẩn hóa và phân loại theo nhóm ý định, phản ánh các nhu cầu phổ biến của người dân khi sử dụng dịch vụ công trực tuyến. Quá trình huấn luyện tập trung vào việc tối ưu khả năng nhận dạng ý định và trích xuất thông tin quan trọng từ truy vấn của người dùng, nhằm đảm bảo chatbot có thể phản hồi chính xác và phù hợp với ngữ cảnh. Trong quá trình huấn luyện, mô hình chatbot được kiểm tra định kỳ thông qua tập dữ liệu kiểm thử nhằm đánh giá khả năng nhận dạng ý định và mức độ phù hợp của phản hồi.

Bảng 1. Thống kê dữ liệu huấn luyện AI Chatbot

Nhóm dữ liệu	Số lượng câu hỏi	Ghi chú
Hướng dẫn thủ tục	16647	Theo nhóm lĩnh vực và 14 nhóm thủ tục hành chính
Tra cứu hồ sơ	50	Chuẩn hóa mẫu
Tài khoản, Thông tin liên hệ, phản ánh	400	Theo các mục
Câu hỏi thường gặp (FAQ)	112	Diễn đạt đa dạng
Tổng cộng	17209	

2.4. Xây dựng kịch bản AI Chatbot trong dịch vụ hành chính công

Các kịch bản AI Chatbot được xây dựng dựa trên khảo sát thực tế hành vi người dùng và nội dung tương tác phổ biến trên hệ thống dịch vụ công trực tuyến. Kịch bản tập trung vào bốn nhóm chức năng chính:

- i) hướng dẫn thủ tục hành chính,
- ii) tra cứu trạng thái hồ sơ
- iii) hỗ trợ các câu hỏi thường gặp
- iv) tài khoản, thông tin liên hệ, phản ánh

2.5. Phương pháp thu thập và tiền xử lý dữ liệu dịch vụ công trực tuyến

Dữ liệu dịch vụ công trực tuyến được thu thập từ hệ thống nghiệp vụ trong một khoảng thời gian đủ dài để phản ánh xu hướng và biến động theo thời gian. Dữ liệu bao gồm số lượng hồ sơ tiếp nhận, số lượng hồ sơ đã giải quyết, hồ sơ trễ hạn và các chỉ số liên quan đến mức độ hài lòng của người dân.

Trước khi tiến hành phân tích, dữ liệu được tiền xử lý nhằm loại bỏ các giá trị thiếu, chuẩn hóa định dạng thời gian và xử lý các điểm bất thường. Quá trình này giúp đảm bảo chất lượng dữ liệu đầu vào và độ tin cậy của các kết quả phân tích, dự báo

2.6. Phương pháp phân tích và dự báo dữ liệu

Nghiên cứu áp dụng các phương pháp phân tích chuỗi thời gian nhằm khai thác dữ liệu dịch vụ công trực tuyến cho mục tiêu hỗ trợ quản lý. Mô hình ARIMA được lựa chọn để phân tích và dự báo các chuỗi dữ liệu có xu hướng ổn định, dựa trên các phương pháp phân tích chuỗi thời gian truyền thống [8], [9]. Đối với các chuỗi dữ liệu có tính phi tuyến và biến động phức tạp, mô hình mạng nơ-ron hồi tiếp LSTM được áp dụng nhờ khả năng học các quan hệ phụ thuộc dài hạn trong dữ liệu [7].

Việc lựa chọn và so sánh hai mô hình cho phép đánh giá khả năng áp dụng các kỹ thuật phân tích truyền thống và học sâu trong bối cảnh dữ liệu hành chính công. Các kết quả dự báo được sử dụng để phân tích xu hướng tiếp nhận và giải quyết hồ sơ, từ đó hỗ trợ công tác lập kế hoạch và phân bổ nguồn lực.

2.7. Phương pháp đánh giá hiệu quả hệ thống

Hiệu quả của hệ thống được đánh giá thông qua hai nhóm tiêu chí chính. Nhóm tiêu chí thứ nhất tập trung vào hiệu quả hỗ trợ người dùng, bao gồm khả năng trả lời đúng các truy vấn phổ biến, thời gian phản hồi và mức độ ổn định của chatbot. Nhóm tiêu chí thứ hai tập trung vào hiệu quả phân tích dữ liệu, bao gồm độ phù hợp của mô hình dự báo và khả năng phản ánh xu hướng thực tế của dữ liệu dịch vụ công trực tuyến.

Việc đánh giá được thực hiện trong môi trường triển khai thực tế, cho phép xem xét toàn diện cả khía cạnh kỹ thuật và khía cạnh ứng dụng của hệ thống.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kết quả đánh giá hệ thống AI Chatbot

Hiệu quả của hệ thống AI Chatbot được đánh giá dựa trên tập kịch bản và dữ liệu tương tác thực tế được xây dựng trong quá trình triển khai. Các tiêu chí đánh giá tập trung vào khả năng nhận dạng đúng nhu cầu người dùng, mức độ bao phủ kịch bản và hiệu năng phản hồi. Kết quả đánh giá chi tiết theo từng nhóm kịch bản được trình bày trong Bảng 2 và Bảng 3, cho thấy chatbot đạt hiệu quả cao đối với các kịch bản đã được chuẩn hóa.

Bảng 2. Kết quả đánh giá khả năng hỗ trợ của AI Chatbot theo nhóm kịch bản

Nhóm kịch bản	Tỷ lệ trả lời đúng (%)	Ghi chú
Hướng dẫn thủ tục hành chính	~88–92	Độ chính xác cao với thủ tục phổ biến
Tra cứu trạng thái hồ sơ	~90–95	Phụ thuộc tính hợp lệ của mã hồ sơ
Hỏi – đáp câu hỏi thường gặp	~85–88	Giảm dần với câu hỏi diễn đạt tự do
Thông tin liên hệ, hỗ trợ	~95	Nội dung tĩnh, ít sai lệch

Kết quả cho thấy chatbot hoạt động hiệu quả nhất đối với các kịch bản đã được chuẩn hóa rõ ràng, đặc biệt là tra cứu hồ sơ và cung cấp thông tin liên hệ. Đối với các câu hỏi có cách diễn đạt đa dạng, độ chính xác có xu hướng giảm nhẹ. Các kết quả này phù hợp với đặc điểm của dữ liệu huấn luyện và cách xây dựng kịch bản trong Mục 2, cho thấy vai trò quyết định của việc chuẩn hóa nội dung và cấu trúc kịch bản đối với hiệu quả của chatbot.

Bảng 3. Đánh giá hiệu năng phản hồi của AI Chatbot

Tiêu chí	Giá trị trung bình
Thời gian phản hồi	< 1 giây
Số lượt tương tác liên tục	≥ 3 lượt
Tỷ lệ yêu cầu chuyển hỗ trợ thủ công	< 10%

Kết quả cho thấy hệ thống đáp ứng tốt yêu cầu tương tác thời gian thực, góp phần cải thiện trải nghiệm người dùng so với hình thức tra cứu thủ công hoặc liên hệ trực tiếp. Kết quả này phù hợp với các nghiên cứu trước đây về ứng dụng chatbot trong hỗ trợ dịch vụ công và các hệ thống thông tin công, cho thấy vai trò ngày càng rõ nét của chatbot trong cải thiện tương tác người dùng [4], [6].

3.2. Kết quả phân tích dữ liệu dịch vụ công trực tuyến

Dữ liệu dịch vụ công trực tuyến được phân tích nhằm làm rõ xu hướng tiếp nhận và giải quyết hồ sơ theo thời gian, đồng thời đánh giá tiềm năng khai thác dữ liệu phục vụ công tác quản lý. Bảng 4 trình bày thống kê mô tả dữ liệu dịch vụ công trực tuyến.

Bảng 4. Thống kê mô tả dữ liệu dịch vụ công trực tuyến

Chỉ tiêu	Giá trị điển hình
Số lượng hồ sơ tiếp nhận trung bình/tháng	Hàng chục nghìn
Tỷ lệ hồ sơ giải quyết đúng hạn	> 90%
Tỷ lệ hồ sơ trễ hạn	< 10%
Mức độ hài lòng trung bình	Cao

Các kết quả thống kê cho thấy hệ thống dịch vụ công trực tuyến đã được vận hành ổn định, đồng thời cung cấp nguồn dữ liệu có giá trị cho các phân tích sâu hơn. Việc phân tích dữ liệu DVCTT không chỉ phục vụ báo cáo định kỳ mà còn mở ra khả năng giám sát xu hướng và phát hiện sớm các điểm nghẽn trong quá trình giải quyết thủ tục.

3.3. Kết quả dự báo dữ liệu bằng mô hình ARIMA và LSTM

Khả năng dự báo xu hướng dữ liệu DVCTT được đánh giá thông qua việc so sánh kết quả dự báo của hai mô hình ARIMA và LSTM với dữ liệu thực tế.

Bảng 5. So sánh kết quả dự báo của ARIMA và LSTM

Tiêu chí	ARIMA	LSTM
Khả năng bám xu hướng	Trung bình	Cao
Phù hợp chuỗi ổn định	Tốt	Tốt
Phù hợp chuỗi biến động	Hạn chế	Tốt
Khả năng diễn giải	Cao	Trung bình

Kết quả cho thấy ARIMA phù hợp với các chuỗi dữ liệu có xu hướng ổn định, trong khi LSTM thể hiện ưu thế rõ rệt đối với các chuỗi dữ liệu biến động mạnh và phi tuyến. Sự khác biệt này củng cố lựa chọn phương pháp kết hợp đã trình bày ở Mục 2, cho thấy việc sử dụng linh hoạt các mô hình dự báo là cần thiết trong bối cảnh dữ liệu hành chính công đa dạng. Các kết quả dự báo có thể được sử dụng như một công cụ hỗ trợ lập kế hoạch tiếp nhận và giải quyết hồ sơ, đặc biệt trong các giai đoạn cao điểm. Sự khác biệt về hiệu quả dự báo giữa ARIMA và LSTM trong nghiên cứu này cũng được ghi nhận trong nhiều nghiên cứu về phân tích chuỗi thời gian, cho thấy cần lựa chọn mô hình phù hợp với đặc điểm cụ thể của dữ liệu [7]–[9].

3.4. Đánh giá tổng hợp hiệu quả hệ thống

Để đánh giá tổng thể hiệu quả của hệ thống, các kết quả từ AI Chatbot và hệ thống phân tích – dự báo dữ liệu được tổng hợp và xem xét trong bối cảnh ứng dụng thực tế.

Bảng 6. Tổng hợp hiệu quả của hệ thống đề xuất

Thành phần	Kết quả chính	Ý nghĩa ứng dụng
AI Chatbot	Tỷ lệ trả lời đúng cao, phản hồi nhanh	Nâng cao trải nghiệm người dân
Phân tích dữ liệu	Nhận diện xu hướng, điểm nghẽn	Hỗ trợ quản lý, giám sát
Dự báo dữ liệu	Dự báo xu hướng hiệu quả	Hỗ trợ lập kế hoạch, phân bổ nguồn lực

3.5. Thảo luận

Các kết quả thu được từ việc đánh giá hệ thống AI Chatbot và hệ thống phân tích – dự báo dữ liệu dịch vụ công trực tuyến cho thấy cách tiếp cận tích hợp được đề xuất trong nghiên cứu là phù hợp và có tính khả thi cao trong bối cảnh ứng dụng thực tế. Việc kết hợp hỗ trợ tương tác và khai thác dữ liệu trong cùng một hệ thống phù hợp với định hướng phát triển chính phủ số và quản trị dựa trên dữ liệu được khuyến nghị trong các báo cáo quốc tế về chính phủ điện tử [10]. Việc kết hợp hỗ trợ người dùng thông qua AI Chatbot và khai thác dữ liệu nghiệp vụ thông qua các mô hình phân tích – dự báo không chỉ giải quyết các vấn đề tương tác thường gặp, mà còn mở rộng giá trị sử dụng của dữ liệu dịch vụ công trực tuyến.

Thứ nhất, xét trên góc độ hỗ trợ người dân, AI Chatbot cho thấy hiệu quả rõ rệt trong việc cung cấp thông tin và hướng dẫn sử dụng dịch vụ công trực tuyến. Các kết quả đánh giá cho thấy chatbot hoạt động ổn định đối với các kịch bản đã được chuẩn hóa, đặc biệt là các kịch bản hướng dẫn thủ tục và tra cứu trạng thái hồ sơ. Điều này khẳng định rằng việc áp dụng chatbot dựa trên xử lý ngôn ngữ tự nhiên có thể góp phần giảm rào cản tiếp cận dịch vụ công, nhất là đối với người dân chưa quen với các quy trình trực tuyến. Tuy nhiên, kết quả cũng cho thấy hiệu quả của chatbot phụ thuộc đáng kể vào chất lượng dữ liệu huấn luyện và mức độ chuẩn hóa kịch bản, gợi ý rằng việc duy trì và cập nhật nội dung là yếu tố then chốt để đảm bảo hiệu quả lâu dài.

Thứ hai, xét trên góc độ khai thác dữ liệu và hỗ trợ quản lý, các kết quả phân tích và dự báo dữ liệu dịch vụ công trực tuyến cho thấy tiềm năng lớn của việc sử dụng dữ liệu hành chính cho các mục tiêu ngoài báo cáo thống kê truyền thống. Việc áp dụng các mô hình ARIMA và LSTM cho phép nhận diện xu hướng biến động theo thời gian và dự báo nhu cầu xử lý hồ sơ trong các giai đoạn tiếp theo. Kết quả so sánh giữa hai mô hình cho thấy không tồn tại một phương pháp dự báo tối ưu cho mọi loại dữ liệu, mà cần lựa chọn mô hình phù hợp với đặc điểm cụ thể của từng chuỗi dữ liệu. Nhận xét này

có ý nghĩa thực tiễn đối với các cơ quan quản lý khi triển khai các công cụ phân tích dữ liệu trong hệ thống dịch vụ công.

Thứ ba, từ góc độ tính tích hợp của hệ thống, nghiên cứu cho thấy việc kết hợp AI Chatbot và hệ thống phân tích – dự báo dữ liệu trong cùng một kiến trúc mang lại lợi ích bổ sung lẫn nhau. Trong khi chatbot tập trung vào tương tác và hỗ trợ người dùng cuối, hệ thống phân tích dữ liệu cung cấp các thông tin tổng hợp và dự báo phục vụ công tác quản lý. Cách tiếp cận này giúp tận dụng tối đa dữ liệu phát sinh trong quá trình cung cấp dịch vụ công, đồng thời tạo ra một hệ sinh thái hỗ trợ toàn diện hơn so với các giải pháp đơn lẻ.

Bên cạnh các kết quả đạt được, nghiên cứu cũng chỉ ra một số hạn chế cần được xem xét. Thứ nhất, sự đa dạng trong cách diễn đạt của người dùng đặt ra thách thức đối với khả năng xử lý ngôn ngữ tự nhiên của chatbot, đặc biệt trong bối cảnh tiếng Việt hành chính công. Thứ hai, chất lượng và tính đồng nhất của dữ liệu dịch vụ công trực tuyến giữa các đơn vị có thể ảnh hưởng đến độ chính xác của các mô hình phân tích và dự báo. Những hạn chế này cho thấy nhu cầu tiếp tục hoàn thiện dữ liệu, cải thiện mô hình xử lý ngôn ngữ và mở rộng phạm vi đánh giá trong các nghiên cứu tiếp theo.

Cuối cùng, từ góc độ ý nghĩa khoa học và ứng dụng, nghiên cứu góp phần làm rõ vai trò của trí tuệ nhân tạo trong việc nâng cao hiệu quả cung cấp và quản lý dịch vụ công trực tuyến. Kết quả nghiên cứu không chỉ có giá trị trong phạm vi nghiên cứu trường hợp được lựa chọn, mà còn cung cấp các gợi ý phương pháp luận cho việc triển khai các hệ thống tương tự tại các địa phương khác. Điều này cho thấy tiềm năng của việc ứng dụng AI như một công cụ hỗ trợ thiết thực trong tiến trình xây dựng chính quyền số và cải cách hành chính.

4. KẾT LUẬN

Bài báo đã đề xuất và đánh giá một hệ thống hỗ trợ dịch vụ công trực tuyến dựa trên công nghệ trí tuệ nhân tạo, kết hợp giữa AI Chatbot và các mô hình phân tích – dự báo dữ liệu. Cách tiếp cận tích hợp này nhằm giải quyết đồng thời hai vấn đề trọng tâm trong cung cấp dịch vụ công trực tuyến: hỗ trợ người dân tiếp cận và sử dụng dịch vụ một cách hiệu quả, và khai thác dữ liệu nghiệp vụ để phục vụ công tác quản lý và ra quyết định.

Trên cơ sở phương pháp nghiên cứu được trình bày, hệ thống AI Chatbot được xây dựng với các kịch bản bám sát nghiệp vụ hành chính công, cho thấy khả năng hỗ trợ tốt các nhu cầu tương tác phổ biến như hướng dẫn thủ tục, tra cứu trạng thái hồ sơ và giải đáp các câu hỏi thường gặp. Các kết quả đánh giá trong môi trường thực tế cho thấy chatbot hoạt động ổn định, phản hồi nhanh và đạt độ chính xác cao đối với các kịch bản

đã được chuẩn hóa, qua đó góp phần cải thiện trải nghiệm người dùng và giảm tải cho bộ phận hỗ trợ trực tiếp.

Bên cạnh đó, nghiên cứu đã chỉ ra tiềm năng khai thác dữ liệu dịch vụ công trực tuyến thông qua các phương pháp phân tích và dự báo chuỗi thời gian. Việc áp dụng và so sánh các mô hình ARIMA và LSTM cho thấy dữ liệu hành chính công không chỉ có giá trị thống kê mà còn có thể được sử dụng để nhận diện xu hướng, dự báo nhu cầu và hỗ trợ công tác quản lý. Kết quả này khẳng định tính khả thi của việc ứng dụng các kỹ thuật phân tích dữ liệu và học sâu trong lĩnh vực hành chính công.

Từ góc độ khoa học, nghiên cứu đóng góp một cách tiếp cận tích hợp giữa hỗ trợ tương tác và phân tích dữ liệu trong cùng một hệ thống dịch vụ công trực tuyến. Cách tiếp cận này giúp làm rõ vai trò của trí tuệ nhân tạo không chỉ như một công cụ giao tiếp, mà còn như một thành phần hỗ trợ ra quyết định dựa trên dữ liệu. Về mặt ứng dụng, các kết quả nghiên cứu cho thấy hệ thống đề xuất có khả năng triển khai và mở rộng trong thực tiễn, đặc biệt đối với các địa phương đang đẩy mạnh chuyển đổi số trong lĩnh vực hành chính công.

Mặc dù đạt được các kết quả tích cực, nghiên cứu vẫn còn một số hạn chế, bao gồm sự phụ thuộc vào chất lượng dữ liệu đầu vào và những thách thức trong xử lý ngôn ngữ tự nhiên tiếng Việt với nhiều cách diễn đạt khác nhau. Đây cũng là hướng mở cho các nghiên cứu tiếp theo, với việc mở rộng dữ liệu huấn luyện, tích hợp các mô hình ngôn ngữ tiên tiến và đánh giá hệ thống trên phạm vi rộng hơn.

Tóm lại, nghiên cứu đã chứng minh tiềm năng của việc ứng dụng trí tuệ nhân tạo trong nâng cao hiệu quả cung cấp và quản lý dịch vụ công trực tuyến. Các kết quả nghiên cứu góp phần cung cấp cơ sở khoa học cho việc triển khai các giải pháp trí tuệ nhân tạo trong dịch vụ công trực tuyến, phù hợp với định hướng chuyển đổi số quốc gia và xu hướng phát triển chính phủ số hiện nay [1], [10].

LỜI CẢM ƠN

Đây là kết quả của đề tài khoa học và công nghệ cấp tỉnh, mã số TTH.2021-KC.18 được ngân sách nhà nước tỉnh Thừa Thiên Huế đầu tư.

TÀI LIỆU THAM KHẢO

- [1] Bộ Thông tin và Truyền thông (2020), Chương trình Chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030, Quyết định của Thủ tướng Chính phủ, Hà Nội.
- [2] Văn phòng Chính phủ (2021), Cổng Dịch vụ công quốc gia – Định hướng phát triển và tích hợp dịch vụ, Báo cáo chuyên đề, Hà Nội.
- [3] Bộ Nội vụ (2021), Cải cách thủ tục hành chính và xây dựng chính quyền số ở Việt Nam, Nhà xuất bản Chính trị Quốc gia Sự thật, Hà Nội.
- [4] Adamopoulou, E. and Moussiades, L. (2020), “An overview of chatbot technology”, *Artificial Intelligence Applications and Innovations, IFIP Advances in Information and Communication Technology*, vol. 584, pp. 373–383.
- [5] Jurafsky, D. and Martin, J.H. (2023), *Speech and Language Processing*, 3rd ed., Pearson, New York.
- [6] Bocklisch, T., Faulkner, J., Pawlowski, N. and Nichol, A. (2017), “Rasa: Open source language understanding and dialogue management”, *Proceedings of the NIPS Workshop on Conversational AI*, Long Beach, USA.
- [7] Hochreiter, S. and Schmidhuber, J. (1997), “Long short-term memory”, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780.
- [8] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015), *Time Series Analysis: Forecasting and Control*, 5th ed., Wiley, New Jersey.
- [9] Chatfield, C. (2016), *The Analysis of Time Series: An Introduction*, 6th ed., CRC Press, Boca Raton.
- [10] United Nations (2022), *E-Government Survey 2022: The Future of Digital Government*, United Nations Department of Economic and Social Affairs, New York.

AN AI-BASED CHATBOT AND DATA ANALYSIS SYSTEM FOR ONLINE PUBLIC SERVICES: A CASE STUYDY IN THUA THIEN HUE PROVINCE

Bui Quang Vu^{1*}, Pham Quang Tri², Le Van Man², Nguyen Kim Tung³

1 Mathematics Department, University of Sciences, Hue University

2 Office of the People's Committee, Thua Thien Hue Province

3 Department of Science and Technology, Thua Thien Hue Province

* Email: buiquangvu@husc.edu.vn

ABSTRACT

Online public services (OPS) play a crucial role in the development of digital government, contributing to improved public administration efficiency and enhanced service quality for citizens. Although OPS platforms have been widely deployed, supporting citizens in effectively accessing, utilizing, and benefiting from these services remains challenging, particularly in procedural guidance and user interaction. This paper proposes and evaluates an AI-based support system for online public services, consisting of two main components: (i) an AI chatbot designed to assist citizens with procedural guidance, information retrieval, and interactive support; and (ii) a data analysis and forecasting system aimed at supporting management and decision-making processes. The chatbot is developed using the RASA framework with a natural language processing model tailored to the Vietnamese public administration context, while the data analysis component applies ARIMA and LSTM models for time series forecasting. Evaluation results in a real-world deployment environment indicate that the proposed system operates stably, improves user experience, and effectively supports the management of online public services.

Keywords: AI chatbot; online public services; natural language processing; data analysis; ARIMA; LSTM.



Bùi Quang Vũ sinh ngày 28/08/1980 tại Thừa Thiên Huế. Năm 2002, ông tốt nghiệp cử nhân chuyên ngành Toán học tại trường Đại học Khoa học, Đại học Huế. Năm 2007, ông nhận bằng thạc sĩ chuyên ngành Xác suất thống kê tại trường Đại học Khoa học, Đại học Huế. Năm 2018, ông nhận học vị tiến sĩ chuyên ngành Khoa học máy tính, Thống kê và nhận thức tại trường Paris Sciences et Lettres, Cộng hòa Pháp. Hiện nay, ông đang công tác tại Khoa Toán, trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: Thống kê, Trí tuệ nhân tạo, học máy, thị giác máy tính.

Phạm Quang Trí sinh năm 1974 tại Thừa Thiên Huế. Năm 1997, ông tốt nghiệp cử nhân chuyên ngành Sinh học tại trường Đại học Khoa học, Đại học Huế. Năm 2000, ông tốt nghiệp cử nhân chuyên ngành CNTT tại trường Đại học Khoa học, Đại học Huế. Năm 2020, ông nhận bằng thạc sĩ chuyên ngành Quản lý kinh tế tại trường Đại học Kinh tế, Đại học Huế. Hiện nay, ông đang công tác tại Văn phòng ủy ban tỉnh Thừa Thiên Huế.

Lĩnh vực nghiên cứu: Hành chính công, Chính quyền điện tử.



Lê Văn Mẫn sinh ngày 26/11/1987 tại Thừa Thiên Huế. Năm 2011, ông tốt nghiệp kỹ sư chuyên ngành Điện tử viễn thông tại trường Đại học Khoa học, Đại học Huế. Năm 2017, ông tốt nghiệp cử nhân Luật tại trường Đại học Luật, Đại học Huế. Hiện nay, ông đang công tác tại Văn phòng ủy ban tỉnh Thừa Thiên Huế.

Lĩnh vực nghiên cứu: Hành chính công, Chính quyền điện tử



Nguyễn Kim Tùng sinh ngày 11/03/2079 tại Thừa Thiên Huế. Năm 2001, ông tốt nghiệp cử nhân chuyên ngành Công Nghệ Thông tin tại trường Đại học Sư phạm, Đại học Huế. Năm 2024, ông nhận bằng thạc sĩ chuyên ngành Quản lý Công nghệ Thông tin tại trường Đại học Khoa học, Đại học Huế. Hiện nay, ông đang là công tác tại Sở Khoa học Công nghệ, tỉnh Thừa Thiên Huế.

Lĩnh vực nghiên cứu: Hành chính công, Chính quyền điện tử, Quản lý CNTT, Học máy.