

Nghiên cứu, xây dựng công cụ hướng dẫn thủ tục hành chính về khoa học và công nghệ

Trần Văn Vinh*

*Khoa Cơ khí, Học viện Kỹ thuật Quân sự

Received: 8/02/2024; Accepted: 16/02/2024; Published: 20/02/2024

Abstract: Chatbot is a typical tool in the field of communication, answering and guiding procedures and regulations. The article presents research on building an Chatbot tool to guide administrative procedures in Science & Technology that can communicate and answer users' questions about administrative procedures in Science & Technology. Chatbot built with the Python language installer and capable of accurately and quickly answering questions about Science & Technology administrative procedures. Chatbots can be widely applied on platforms such as applications, websites and can be applied in other fields such as finance, law, services, etc., reaching

Keywords: Chatbot, Natural Language Processing, Artificial Intelligence

1. Đặt vấn đề

Ngày nay, trong thời đại của cuộc cách mạng công nghệ lần thứ 4, sự phát triển nhanh chóng của lĩnh vực trí tuệ nhân tạo (Artificial Intelligent – AI) đã tạo ra những công cụ có khả năng thay thế con người trong nhiều lĩnh vực như sản xuất, công nghiệp và dịch vụ. Chatbot là một ứng dụng hoặc hệ thống phần mềm được thiết kế để tương tác với con người thông qua cuộc trò chuyện hoặc giao tiếp văn bản. Chatbot công nghệ trí tuệ nhân tạo (AI) và xử lý ngôn ngữ tự nhiên (NLP) để hiểu các câu hỏi và tự động trả lời. Khái niệm hoá Chatbot đầu tiên được quy cho Alan Turning, người đã hỏi “Máy móc có thể suy nghĩ không?” vào năm 1950 (Turning A, 1950). Tuy nhiên, giai đoạn quan trọng nhất đến vào những năm 1960 - 1970 với sự xuất hiện của ELIZA – chatbot đầu tiên trên thế giới, là một chương trình máy tính của Joseph Weizenbaum (Viện công nghệ Massachusetts, Mỹ) [1]. Chatbot ngày nay sử dụng NLP để hiểu ngôn ngữ tự nhiên, xử lý ngữ cảnh và cung cấp trải nghiệm giao tiếp ngày càng tự nhiên. Xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP) là một lĩnh vực thuộc Trí tuệ nhân tạo với mục tiêu giúp máy tính có thể hiểu được ngôn ngữ của con người. Xử lý ngôn ngữ tự nhiên (NLP) là một nhánh cực kỳ quan trọng của trí tuệ nhân tạo (AI), là giao điểm của Ngôn ngữ học, Khoa học Máy tính và AI. Chatbot được xây dựng trên cơ sở trí tuệ nhân tạo (AI) và có khả năng tự động phản hồi và hiểu ngôn ngữ tự nhiên của con người. Chatbot ứng dụng sử dụng trí tuệ nhân tạo AI đang là một công cụ quan trọng trong nhiều lĩnh vực kinh doanh và đời sống thường ngày trên các nền

tảng số hoá. Điều này mở ra tương lai triển vọng khi chatbot có thể trở thành một phần quan trọng trong nhiều lĩnh vực như: dịch vụ công, y tế, giáo dục, doanh nghiệp và có khả năng tương tác với con người một cách thông minh và linh hoạt.

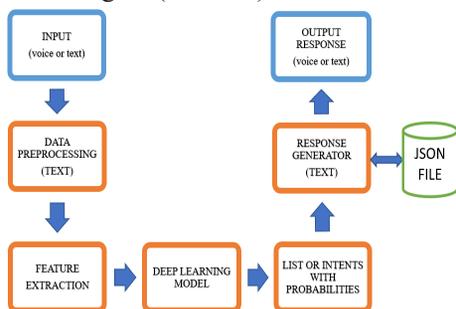
Những năm gần đây, sự ra đời của các nền tảng Chatbot như Bing chat, ChatGPT, Google Bard đã tạo ra những làn sóng mới trong sự phát triển của công nghệ AI. Interactive Tax Assistant (ITA) là một trợ lý ảo ở Mỹ được sử dụng để cung cấp câu trả lời cho một số câu hỏi về luật thuế cụ thể, ITA hỗ trợ người dân và doanh nghiệp hiểu rõ về các quy định thuế tại Mỹ [2]. Ở Việt Nam, sau đại dịch Covid-19, Chatbot MED (CHAM) đã được xây dựng và ứng dụng vào cuộc sống nhằm theo dõi cá thể mang bệnh nền và mô hình hoá thời gian thực của di chứng hậu Covid-19 nhằm trợ giúp bác sĩ trong quá trình chuẩn đoán, điều trị. Sự phát triển và ứng dụng rộng rãi của công nghệ Chatbot đã cho thấy rằng chúng sẽ còn trở nên mạnh mẽ hơn, phổ biến hơn trong tương lai. Trong các cơ quan hành chính công, Chatbot có thể hỗ trợ ra quyết định và tăng khả năng phản hồi dẫn đến sự hài lòng của người dân cao hơn. Chatbot hướng dẫn thủ tục hành chính có thể giải đáp các thắc mắc, hướng dẫn xử lý các thủ tục. Tuy nhiên, dù có đột phá đến đâu, hiện đại đến đâu thì công nghệ vẫn phải hướng đến sự chấp nhận của người dùng. Mặc dù AI đã có những bước tiến phát triển mạnh mẽ trên thế giới nhưng việc ứng dụng vào trong lĩnh vực dịch vụ công vẫn còn ở giai đoạn sơ khai. Việc hướng dẫn các thủ tục hành chính hiện nay còn nhiều khó khăn, bất cập và tốn kém. Vì vậy, ứng dụng Chatbot

để thay thế con người trong hướng dẫn thủ tục hành chính là rất cần thiết, mang lại những lợi ích và hiệu quả cao đối với hệ thống hành chính của tổ chức. Trong đó, Khoa học và Công nghệ là một lĩnh vực quan trọng góp phần vào sự phát triển Kinh tế - Xã hội của đất nước. Việc đưa Chatbot vào việc hướng dẫn thủ tục hành chính Khoa học – Công nghệ giúp tiết kiệm thời gian, nhận được thông tin cần thiết bất cứ lúc nào, quảng bá công nghệ và sáng tạo, tăng cường minh bạch và tuân thủ quy định quy trình, góp phần nâng cao được hiệu quả và sự hài lòng của người sử dụng trong quá trình thực hiện thủ tục hành chính, thúc đẩy hoạt động Khoa học – Công nghệ trong nước ngày càng phát triển.

2. Nội dung nghiên cứu

2.1. Cách thức hoạt động

Chatbot là sự kết hợp của các kịch bản có trước và khả năng tự học trong quá trình tương tác. Ta sẽ tương tác với Chatbot qua nền tảng tin nhắn. Trước hết, để có thể hoạt động hiệu quả; một Chatbot cần được huấn luyện trên một lượng lớn dữ liệu để hiểu và phản hồi phù hợp với các câu hỏi và yêu cầu của người dùng. Dữ liệu thu thập được cần tiền xử lý để chuyển đổi văn bản thành định dạng mà mô hình học máy (Machine Learning) có thể hiểu được. Sau đó, mô hình sẽ được huấn luyện bằng cách đưa vào dữ liệu đào tạo và điều chỉnh trọng số sao cho nó có thể dự đoán câu trả lời chính xác dựa trên câu hỏi đầu vào. Chatbot sử dụng xử lý ngôn ngữ tự nhiên (NLP) để hiểu và tạo ra văn bản theo cách mà con người có thể hiểu thông qua các quá trình phân tích cú pháp (syntax), phân tích ý nghĩa (semantic analysis) và phân loại ý nghĩa (intent classification). Sau khi mô hình đã được huấn luyện, Chatbot được triển khai để tương tác với người dùng. Chatbot sẽ đưa ra những câu trả lời dựa trên những dữ liệu đã được huấn luyện và được cập nhật theo thời gian (hình 2.1).



Hình 2.1: Cách thức hoạt động của chatbot

Một trong các yếu tố làm nên sức mạnh của Chatbot là khả năng tự học hỏi. Càng được sử dụng, tương tác với người dùng càng nhiều, nền tảng chatbot càng “thông minh”. Chatbot có khả năng tự học hỏi dựa trên các dữ liệu đưa vào mà không cần phải được lập trình cụ thể (đó được gọi là phương pháp học máy – Machine Learning). Chính điều này làm cho các nhà phát triển dễ dàng tạo các chương trình trò chuyện và tự động hoá cuộc trò chuyện với người dùng. Đối với dữ liệu là các thủ tục hành chính có yêu cầu khắt khe về kết quả đầu ra. Do đó, các Chatbot hướng dẫn về thủ tục hành chính phải có khả năng nhận diện tốt các dữ liệu đầu vào và phải đưa ra một kết quả đầu ra có độ chính xác cực cao.

2.2. Xây dựng hệ thống

2.2.1. Xây dựng cơ sở dữ liệu

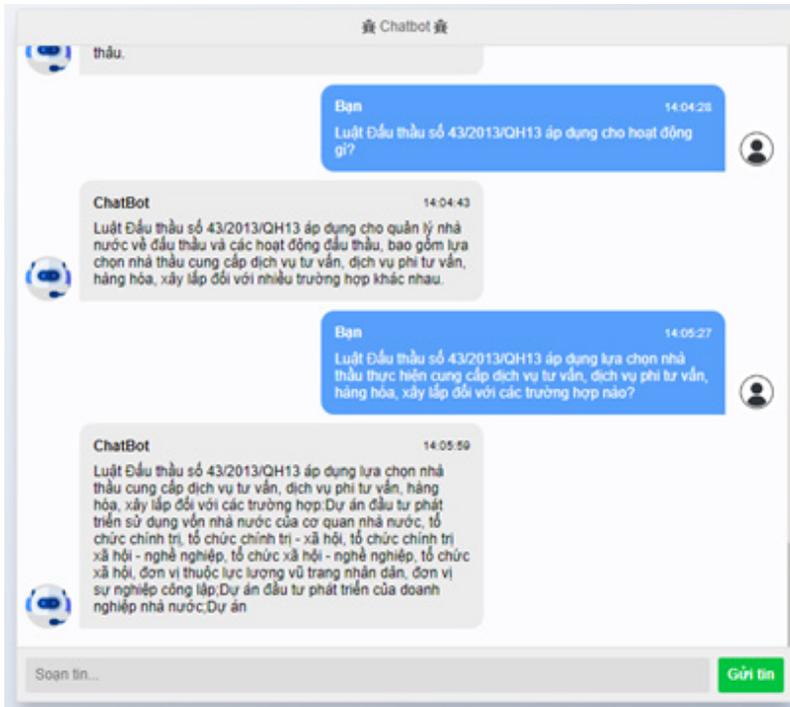
Cơ sở dữ liệu của Chatbot được huấn luyện trên 2 bộ dữ liệu được xây dựng: bộ dữ liệu giao tiếp đời thường và bộ dữ liệu về thủ tục hành chính Khoa học & Công nghệ. Bộ dữ liệu giao tiếp đời thường được lấy từ các dữ liệu có sẵn trên cộng đồng làm tăng độ thân thiện của Chatbot trong quá trình sử dụng. Bộ dữ liệu về thủ tục hành chính Khoa học & Công nghệ. Dữ liệu để huấn luyện Chatbot được thu thập từ hơn 13 loại thủ tục hành chính lấy từ website của Bộ Khoa học & Công nghệ. Dựa vào những dữ liệu từ các văn bản thu thập được; sử dụng Copilot, ChatGPT để tạo ra các cuộc hội thoại và thu được 150MB dữ liệu là các cặp câu hỏi – câu trả lời về thủ tục hành chính Khoa học & Công nghệ như sau (bảng 2.1).

Bảng 2.1: Ví dụ cơ sở dữ liệu của Chatbot

Luật Đấu thầu số 43/2013/QH13 áp dụng cho hoạt động gì?
Luật Đấu thầu số 43/2013/QH13 áp dụng cho quản lý nhà nước về đấu thầu và các hoạt động đấu thầu, bao gồm lựa chọn nhà thầu cung cấp dịch vụ tư vấn, dịch vụ phi tư vấn, hàng hóa, xây lắp đối với nhiều trường hợp khác nhau.
Luật Đấu thầu số 43/2013/QH13 áp dụng cho lựa chọn nhà thầu trong những trường hợp nào?
Luật Đấu thầu số 43/2013/QH13 áp dụng cho lựa chọn nhà thầu trong các trường hợp bao gồm dự án đầu tư phát triển, mua sắm sử dụng vốn nhà nước, mua hàng dự trữ quốc gia, và nhiều trường hợp khác.

2.2.2. Xây dựng hệ thống Chatbot

Hệ thống Chatbot được lập trình trên PyCharm với ngôn ngữ lập trình Python. Chatbot sử dụng các thư viện hỗ trợ như: os, json, fire, torch, datasets, bitsandbytes, loralib. Chatbot có thể được sử dụng như một ứng dụng trên các nền tảng hay có thể được tích hợp trên các trang mạng xã hội, trên website (hình 2.2: Chatbot ứng dụng trên website).



Hình 2.2: Chatbot ứng dụng trên website

Hệ thống Chatbot được thiết lập môi trường với Pytorch và Huggingface Transformers package. Các mô hình fine – turning được xây dựng dựa trên mô hình PEFT (Parameter - Efficient Fine-Tuning). PEFT nhằm mục đích tối ưu hoá quá trình fine – turning bằng cách sử dụng hiệu quả các tài nguyên tính toán có sẵn và giảm số lượng các tham số cần thiết được cập nhật. Cách tiếp cận này trở nên đặc biệt phù hợp khi làm việc với dữ liệu được dán nhãn hạn chế cho một nhiệm vụ cụ thể. Các phương pháp tiếp cận PEFT chỉ tinh chỉnh một số lượng nhỏ các tham số mô hình (bổ sung) trong khi đóng băng hầu hết các tham số của LLM được huấn luyện trước, do đó giảm đáng kể chi phí tính toán và lưu trữ. Điều này cũng khắc phục được các vấn đề về sự quên lãng nghiêm trọng, một hành vi được quan sát thấy trong quá trình hoàn thiện LLM. Các phương pháp tiếp cận PEFT cũng cho thấy hiệu quả hơn so với việc tinh chỉnh trong chế độ dữ liệu thấp và khái quát hóa tốt hơn cho các kịch bản ngoài miền [6]. Nghiên cứu sử dụng PERT LoRA - Phương pháp thích ứng cấp thấp để xây dựng Chatbot. LoRA cho phép chúng ta huấn luyện một vài dense layer trong mạng nơ ron một cách gián tiếp bằng cách tối ưu hóa ma trận phân rã thứ hạng của các dense layer thay đổi trong quá trình thích nghi, trong khi vẫn giữ cố định trọng số đã được huấn luyện trước. Chatbot được thiết lập

với độ dài tối đa của chuỗi đầu vào là 512, độ dài tối đa của chuỗi đầu ra là 1024, thứ hạng k trong LoRA đến 8, tốc độ học (learning rate) là 3e-4; các tham số khác: Temperature: 0.5, batch_size: int = 128, num_epochs: int = 500.

Sau khi huấn luyện mô hình dữ liệu được lưu vào file định dạng “Json”. Quá trình phản hồi dữ liệu sẽ dựa trên các chỉ số checkpoint và dữ liệu mà Chatbot đã được đào tạo. Sau đó, phản hồi lại người dùng qua việc hiển thị trên màn hình.

2.3. Đánh giá kết quả

Thực hiện đánh giá hiệu suất của chatbot dựa trên tiêu chí về độ chính xác của Chatbot. Tiến hành thử nghiệm giao tiếp với chatbot theo hai trường hợp: Trường hợp 1: Giao tiếp ngẫu nhiên với Chatbot,

thực hiện giao tiếp 5 lần, mỗi lần 50 câu hỏi khác nhau với dung lượng cơ sở dữ liệu khác nhau; trường hợp 2: Giao tiếp với Chatbot với điều kiện thay đổi từ ngữ mà ý nghĩa câu hỏi không đổi, thực hiện 5 lần.

Bảng 2.2. Kết quả đánh giá với trường hợp 1

Lần thử nghiệm	Dung lượng cơ sở dữ liệu (Mb)	Số câu đúng	Độ chính xác (%)
1	30	26/50	52
2	50	32/50	64
3	70	38/50	76
4	100	43/50	86
5	150	47/50	94

Bảng 2.3. Kết quả đánh giá với trường hợp 2

Lần thử nghiệm	Số lần biến đổi câu	Số câu đúng	Độ chính xác (%)
1	2	2/2	100
2	3	2/3	66,67
3	4	3/4	75
4	5	3/5	60
5	6	4/6	66,67

Từ bảng 2.2. và 2.3 có thể thấy: Ở trường hợp 1: Khi dung lượng cơ sở dữ liệu càng lớn sẽ cho độ chính xác của phản hồi càng cao. Do đó, với mô hình đã được xây dựng, độ chính xác của Chatbot phụ thuộc rất lớn vào dung lượng cơ sở dữ liệu. Đối với trường hợp 2: Khi thay đổi câu từ trong câu hỏi với số lần khác nhau sẽ cho độ chính xác khác nhau. Điều này cho thấy về khả năng nhận diện câu hỏi

của chatbot, cụ thể hơn là thuật toán xây dựng của Chatbot có ảnh hưởng lớn đến độ chính xác của Chatbot được xây dựng. Từ phân tích, đánh giá về độ chính xác, tốc độ phản hồi của chatbot cho thấy chatbot được xây dựng có hiệu quả cao, độ chính xác hơn 90%. Tuy nhiên, chatbot được xây dựng để hoạt động trong lượng cơ sở dữ liệu nhất định và phản hồi theo dữ liệu đã được xây dựng từ trước, khả năng tự sinh văn bản bị hạn chế. Do đó, Chatbot vẫn được tiếp tục cập nhật, cải thiện.

3. Kết luận

Bài báo đã trình bày một nghiên cứu xây dựng công cụ Chatbot hướng dẫn thủ tục hành chính Khoa học & Công nghệ có khả năng giao tiếp, giải đáp thắc mắc của người hỏi về các thủ tục hành chính Khoa học & Công nghệ. Kết quả nghiên cứu đã xây dựng được Chatbot có hiệu suất cao và có khả năng ứng dụng được trong thực tế nhằm nâng cao hiệu quả công tác Khoa học & Công nghệ. Trong tương lai chatbot có thể phát triển thành trợ lý ảo, có khả năng giao tiếp tự nhiên với con người và thay thế hoàn toàn con người trong việc hướng dẫn các loại thủ tục hành chính, cung cấp thông tin cho người dùng về các công trình nghiên cứu khoa học.

Tài liệu tham khảo

1. O. Trofymenko , Y. Prokop , N. Loginova and A. Zadereyko, Taxonomy of Chatbots, International Scientific And Practical conference, Odesa, Ukraine 09/2021.

2. Cigdem Akkays, Helmut Kremer, Potential Use of Digital Assistants by Governments for Citizen Services: The Case of Germany.

3. Mgr. Tomáš ZEMČÍK, A Brief History of Chatbots, Department of Social Sciences, VŠB - Technical University of Ostrava, Czech republic (2019)

4. Richard Michael Dreyling III, Tanel Tammet and Ingrid Pappel, Artificial Intelligence Use in e-Government Services: A Systematic Interdisciplinary Literature Review (2022), https://doi.org/10.1007/978-981-19-8069-5_37

5. Vu-Thuan Doan, Quoc-Truong Truong, Duc-Vu Nguyen, Vinh-Tiep Nguyen, and Thuy-Ngan Nguyen Luu, Efficient Finetuning Large Language Models For Vietnamese Chatbot, arXiv:2309.04646 [cs.CL] 09 September 2023, <https://doi.org/10.48550/arXiv.2309.04646>.

6. Edward Hu*, Yelong Shen*, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, LORA: LOW-RANK ADAPTATION OF LARGE LANGUAGE MODELS, arXiv:2106.09685v2 [cs.CL] 16 Oct 2021;

7. DopikAI Joint Stock Company, ViGPT™: Large Language Model for Vietnamese Chatbot;

8. Nguyễn Thiện Nhơn, Phạm Văn Hiến, Nghiên cứu sự phát triển của Chatbot 10 năm trở lại đây, Tạp chí Khoa học và kinh tế phát triển số 12 – 2021.

Vận dụng một số phần mềm công nghệ... (tiếp theo trang 243)

Chính vì thế các tác phẩm của người nghệ sĩ luôn thấm đẫm cảm xúc, có thể làm lay động trái tim của con người.

Với SV chuyên ngành Giáo dục Mầm non thì học phần Mĩ thuật còn là học phần cơ sở để bồi đắp kiến thức cho các em bước vào những học phần tiếp theo để sau khi hoàn thành chương trình các em có thể hướng dẫn, tổ chức các hoạt động tạo hình cho trẻ ở trường Mầm non, là những người sẽ bồi đắp kiến thức, phát triển quá trình cảm thụ và sáng tạo nghệ thuật cho trẻ Mầm non sau khi ra trường.

Hiện nay với xu thế phát triển của thời đại, người ta đã thay thế một số công đoạn, hoặc hoàn toàn trong quá trình sáng tạo mĩ thuật bằng máy móc công nghệ bởi tính năng vượt trội: Nhanh, chính xác, kinh tế. Nhưng những sản phẩm do công nghệ thực hiện theo phương thức mô phỏng, sao chép chính xác từ: bố cục, hình hài, đường nét, hình khối, màu sắc,

ánh sáng, nhịp điệu, bút pháp... hết như những tác phẩm đã có hoặc đúng như hiện thực khách quan. Đó cũng chỉ là những sản phẩm của công nghệ, máy móc thiếu vắng hoàn toàn cảm xúc của con người, và chắc chắn không thể so sánh được với những giá trị độc đáo từ xúc cảm thẩm mĩ do chính con người thể hiện.

Tài liệu tham khảo

1. Đỗ Mạnh Cường, “Giáo trình ứng dụng công nghệ thông tin trong dạy học”, Nxb Đại học Quốc gia thành phố Hồ Chí Minh; 2007

2. Ngô Bá Công, “Giáo trình Mĩ thuật cơ bản”; Nxb Đại học Sư phạm; 2011

3. Triệu Thế Hùng, “Ứng dụng tin học trong dạy học trang trí Mĩ thuật”, Nxb Đại học Sư phạm Hà Nội; 2013

4. Lê Thanh Thủy, “Phương pháp tổ chức hoạt động tạo hình cho trẻ mầm non”; Nxb Đại học Sư phạm; 2008