

GIẢI PHÁP ỨNG DỤNG CÔNG NGHỆ CHATBOT TRONG ĐÀO TẠO, BỒI DƯỠNG LÝ LUẬN CHÍNH TRỊ TRỰC TUYẾN TRONG THỜI ĐẠI CÁCH MẠNG CÔNG NGHIỆP 4.0

Văn Công Vũ, Lê Thị Ngọc Hoa¹

TÓM TẮT

Title: *Applying Chatbot technology in training and supporting online political theory in the 4.0 industrial revolution*

Từ khóa: *Chatbot, đào tạo trực tuyến, lý luận chính trị.*

Keywords: *Chatbot, Online Training, Political theory.*

Lịch sử bài báo:

Ngày nhận bài: 30/12/2019;

Ngày nhận kết quả bình duyệt: 5/01/2020;

Ngày chấp nhận đăng bài: 22/2/2020.

Tác giả:

¹ Trường Đại học Kinh tế - Đại học Đà Nẵng

Email:

vuvc@due.edu.vn

Sự phát triển của khoa học và công nghệ, đặc biệt là công nghệ thông tin đã đem lại nhiều ứng dụng quan trọng trong lĩnh vực giáo dục và đào tạo. Trong phạm vi bài báo này, tác giả phân tích sâu về ứng dụng Chatbot và vai trò của Chatbot trong đào tạo, bồi dưỡng lý luận chính trị trực tuyến. Kết quả nghiên cứu cho thấy, Chatbot khi đưa vào sử dụng, sẽ mang lại hiệu quả cao cho các cơ sở đào tạo bởi nó giúp tương tác với người dùng một cách tự động, khả năng tương tác nhanh ở mọi lúc, mọi nơi, giúp các cơ sở đào tạo tăng hiệu suất tương tác và cắt giảm chi phí nhân lực đào tạo, bồi dưỡng lý luận chính trị. Như vậy, người dùng thông qua ứng dụng Chatbot có thể tham khảo những câu trả lời về vấn đề lý luận chính trị theo bộ câu hỏi đã được lập trình sẵn.

ABSTRACT

The development of science and technology, especially information technology has brought to us many important applications in the field of education and training. Within the scope of this article, the researchers have an in-depth analysis of the Chatbot application and the role of Chatbot in training political theory online. Research results show that, when Chatbot is put into use, it will bring high efficiency to training institutions because it helps interact with users automatically. The ability to interact quickly anytime, anywhere, helps training institutions to increase interaction efficiency and cuts the cost of human resources. Thus, users through the Chatbot application can consult the answers about political theory according to the programmed questions.

1. Đặt vấn đề

Hàng năm, các doanh nghiệp phải bỏ ra một khoản chi phí khá lớn cho việc chăm sóc và tư vấn khách hàng, nhất là đối với các doanh nghiệp cung cấp dịch vụ như du lịch,

hàng không, giáo dục, bán hàng online,... Việc ứng dụng công nghệ thông tin (CNTT) vào hoạt động sản xuất - kinh doanh là vấn đề đang được quan tâm bởi lẽ CNTT có vai trò rất lớn trong các hoạt động kinh tế, sản

xuất kinh doanh, cung cấp dịch vụ, xúc tiến thương mại và quản trị doanh nghiệp. Đối với lĩnh vực giáo dục, trong cuộc cách mạng công nghiệp 4.0, sự lên ngôi của trí tuệ nhân tạo (AI) đã đưa Chatbot (hay còn gọi là Trợ lý ảo) trở thành một khái niệm quen thuộc với người dùng. Tại Việt Nam hàng loạt Chatbot đã và đang được nghiên cứu và đưa vào sử dụng trong thời gian gần đây. Chatbot khi đưa vào sử dụng, sẽ mang lại hiệu quả cao cho các cơ sở đào tạo bởi nó giúp tương tác với người dùng một cách tự động, khả năng tương tác 24/7 nên giúp các cơ sở đào tạo tăng hiệu suất tương tác và cắt giảm chi phí nhân lực đào tạo, trong tham luận này, tác giả đề cập đến vấn đề đào tạo, bồi dưỡng lý luận chính trị.

Mục tiêu nghiên cứu của bài viết là việc xây dựng hệ phần mềm mô phỏng một hệ thống trợ lý ảo, hoạt động online trên Internet, hỗ trợ việc tiếp nhận và trả lời một cách tự động các câu hỏi của học sinh, sinh viên và phụ huynh về các thông tin liên quan đến chương trình học các môn giáo dục chính trị, lý luận chính trị của nhà trường hiện nay.

2. Nội dung

2.1. Đào tạo, bồi dưỡng lý luận chính trị trực tuyến

Trong một số định nghĩa cho rằng đào tạo trực tuyến bao gồm nhiều hơn việc cung cấp các khóa học trực tuyến hoàn toàn. Oblinger và Hawkins (2005) cho rằng đào tạo trực tuyến đã chuyển đổi từ một khóa học trực tuyến hoàn toàn sang sử dụng công nghệ để phân phối một phần hoặc toàn bộ khóa học độc lập với thời gian và địa điểm thường trực. Ủy ban châu Âu (EC) (2001) mô tả, đào tạo trực tuyến là việc sử dụng các công nghệ đa phương tiện mới và Internet

để nâng cao chất lượng học tập bằng cách nói lỏng các cơ sở vật chất và dịch vụ cũng như sự trao đổi và hợp tác từ xa.

Đào tạo trực tuyến đề cập đến việc sử dụng các công nghệ thông tin và truyền thông cho phép truy cập vào tài nguyên học tập hoặc giảng dạy trực tuyến. Theo nghĩa rộng nhất của nó, Abbad và cộng sự (2009) đã xác định Giáo dục điện tử có nghĩa là hình thức học được kích hoạt bằng điện tử.

Một số nhà nghiên cứu khác cũng đã xác định, đào tạo trực tuyến là một cách tiếp cận giáo dục mang tính cách mạng (Jennex, 2005; Twigg, 2002) cho phép tạo ra một lực lượng lao động có kiến thức và kỹ năng cần thiết phù hợp với sự phát triển của thời đại (Jennex, 2005). Twigg (2002) mô tả cách tiếp cận đào tạo trực tuyến tập trung vào người học cũng như thiết kế chương trình học liên quan đến một hệ thống tương tác, lặp đi lặp lại, tự lập trình và có thể tùy chỉnh. Welsh và cộng sự (2003) cũng gọi thuật ngữ này là sử dụng công nghệ mạng máy tính, chủ yếu thông qua Internet, để cung cấp thông tin và hướng dẫn cho các cá nhân.

Như vậy, đào tạo, bồi dưỡng lý luận chính trị trực tuyến là việc sử dụng các công nghệ đa phương tiện mới và Internet để nâng cao chất lượng học tập lý luận chính trị bằng cách nói lỏng các cơ sở vật chất và dịch vụ cũng như sự trao đổi và hợp tác từ xa, với cách tiếp cận tập trung vào người học cũng như thiết kế chương trình học liên quan đến một hệ thống tương tác, lặp đi lặp lại, tự lập trình và có thể tùy chỉnh. Đào tạo, bồi dưỡng lý luận chính trị trực tuyến là sử dụng công nghệ mạng máy tính, chủ yếu thông qua Internet, để cung cấp thông tin và hướng dẫn cho các cá nhân về việc đào tạo, bồi dưỡng lý luận chính trị.

2.2. Công nghệ Chatbot

Trong cuộc cách mạng công nghiệp 4.0, sự lên ngôi của trí tuệ nhân tạo (AI) đã đưa Chatbot (hay còn gọi là trợ lý ảo) trở thành một khái niệm quen thuộc với người dùng. Tại Việt Nam hàng loạt Chatbot đã và đang được nghiên cứu và đưa vào sử dụng trong thời gian gần đây. Việc sử dụng Chatbot sẽ mang lại hiệu quả cao cho doanh nghiệp bởi nó giúp tương tác với người dùng một cách tự động. Bên cạnh đó, Chatbot có khả năng trả lời tự động 24/7 nên giúp doanh nghiệp tăng hiệu suất kinh doanh và cắt giảm chi phí nhân lực.

Chatbot là một phần mềm được thiết kế để thực hiện các tác vụ tự động. Nói cách khác, Chatbot là một chương trình máy tính được thiết kế để giao tiếp với người dùng thông qua Internet. Chatbot đơn thuần là một chương trình máy tính mô phỏng cơ bản các cuộc hội thoại của con người. Nó cho phép một hình thức tương tác giữa con người và máy giao tiếp, được thực hiện thông qua tin nhắn hoặc lệnh thoại. Một Chatbot được lập trình để hoạt động độc lập với người vận hành. Nó có thể trả lời các câu hỏi được xây dựng theo ngôn ngữ tự nhiên và trả lời như một người thực sự. Nó cung cấp các câu trả lời dựa trên sự kết hợp của các tập lệnh được xác định trước và các ứng dụng học máy.

Khi được hỏi một câu hỏi, Chatbot sẽ trả lời dựa trên cơ sở dữ liệu kiến thức có sẵn tại thời điểm đó. Nếu cuộc hội thoại giới thiệu một khái niệm mà nó không được lập trình để hiểu, nó có thể làm chệch hướng cuộc trò chuyện hoặc có khả năng truyền thông tin cho một nhà điều hành con người. Tuy nhiên, dù bằng cách nào, nó cũng sẽ học hỏi từ sự tương tác đó cũng như từ các tương tác trong tương lai. Do đó, Chatbot sẽ dần dần phát

triển về phạm vi, trở nên thông minh hơn và đạt được sự chính xác cao.

Hiện nay đã có nhiều doanh nghiệp, cơ sở giáo dục, chính phủ... đã và đang dần áp dụng và chuẩn bị các điều kiện hạ tầng để áp dụng công nghệ Chatbot. Các chuyên gia của Business Insider dự đoán đến năm 2020, 80% doanh nghiệp sẽ sử dụng Chatbot. Theo Lauren Foye, đến năm 2022, các ngân hàng có thể tự động hóa tới 90% tương tác khách hàng của họ bằng cách sử dụng Chatbot. Một cuộc khảo sát do Spiceworks thực hiện cho thấy 40% các công ty lớn sử dụng hơn 500 người có kế hoạch triển khai một hoặc nhiều trợ lý thông minh hoặc robot trò chuyện dựa trên AI trên các thiết bị di động của công ty vào năm 2019.

Các hệ thống Chatbot giao tiếp với con người bằng giọng nói (*như Siri*) hoặc bằng văn bản (*như các Chatbot phát triển trên nền Facebook Messenger*). Dù giao tiếp bằng hình thức nào, Chatbot cũng cần phải hiểu văn bản để có thể đưa ra những câu trả lời phù hợp cho khách hàng. Thành phần đảm nhiệm công việc này trong hệ thống Chatbot được gọi là NLU (*Natural Language Understanding*), trong đó có rất nhiều các kỹ thuật xử lý ngôn ngữ tự nhiên (*NLP - Natural Language Processing*) được áp dụng. Chatbots xử lý ngôn ngữ tự nhiên bằng cách chuyển đổi lời nói hoặc văn bản của người dùng thành dữ liệu có cấu trúc. Và sau đó, chọn một câu trả lời có liên quan để phản hồi.

Xử lý ngôn ngữ tự nhiên NLP bao gồm các bước:

Mã thông báo: NLP chia một loạt các từ thành mã thông báo hoặc các phần đại diện theo ngôn ngữ.

Phân tích ý định/cảm xúc của người dùng: Nó sẽ nghiên cứu và học hỏi từ trải

nghiệm người dùng và chuyển yêu cầu cho con người khi cần thiết.

Chuẩn hóa: Xử lý văn bản để tìm lỗi chính tả có thể dẫn đến thay đổi ý định của người dùng.

Định danh thực thể được đặt tên: Mô hình chương trình của Chatbot tìm kiếm các loại từ khác nhau, tương tự như tên của một sản phẩm, địa chỉ hoặc tên người dùng cụ thể, tùy thuộc vào thông tin nào được yêu cầu.

Phân tích cú pháp phụ thuộc: Chatbot tìm kiếm các chủ đề chung, động từ, chủ đề, cụm từ và danh từ trong văn bản người dùng để khám phá các cụm từ liên quan mà người dùng muốn truyền đạt.

Các hệ thống Chatbot được triển khai trong thực tế phần lớn tuân theo mô hình truy xuất thông tin và được áp dụng trong những miền ứng dụng nhất định. Đối với các Chatbot được sử dụng trong miền ứng dụng đóng (*closed domain*) và trả lời theo mô hình truy xuất thông tin (*retrieval-based model*). Mô hình truy xuất thông tin là mô hình trong đó, Chatbot đưa ra những phản hồi được chuẩn bị trước hoặc tuân theo những mô thức nhất định. Mô hình này khác với mô hình tự động sinh câu trả lời (*generative*), trong đó câu trả lời của Chatbot được tự động sinh ra bằng việc học từ một tập dữ liệu các đoạn hội thoại.

2.3. Vai trò của công nghệ Chatbot trong đào tạo, bồi dưỡng lý luận chính trị trực tuyến

Hiện nay, vấn đề đào tạo, bồi dưỡng lý luận chính trị có vai trò quan trọng trong việc đào tạo đội ngũ tri thức xã hội chủ nghĩa ở nước ta. Việc trang bị cho sinh viên thế giới quan, nhân sinh quan khoa học của chủ nghĩa Mác – Lênin, bồi dưỡng phẩm chất đạo đức cách mạng, lòng yêu nước, niềm tự hào dân tộc, rèn luyện bản lĩnh chính trị,

kiên định mục tiêu độc lập dân tộc và chủ nghĩa xã hội mà Đảng và Chủ tịch Hồ Chí Minh đã lựa chọn là hết sức cần thiết. Tuy nhiên, với một số cơ sở đào tạo, khung chương trình và thực tiễn việc đào tạo trực tiếp tại giảng đường có phần chênh lệch; bên cạnh đó, việc chậm trễ hay thiếu hụt thông tin về các vấn đề của bài giảng là không thể tránh khỏi. Việc ứng dụng công nghệ Chatbot có thể giải quyết được vấn đề này.

Sử dụng công nghệ Chatbot có ý nghĩa lớn trong việc đào tạo lý luận chính trị trực tuyến. Với các môn học lý luận chính trị, việc tìm kiếm tài liệu thủ công liên quan đến vấn đề nghiên cứu có phần khó khăn hơn các môn học thông thường, do đó tốn kém nhiều thời gian. Trong khi đó, Chatbot luôn sẵn sàng hỗ trợ người học lý luận chính trị mọi lúc mọi nơi, phù hợp với lịch trình bận rộn của người học, đây là cách nhanh nhất để tiếp cận tri thức với chi phí thấp hơn. Khi giảng dạy lý luận chính trị với hình thức dạy học trực tuyến (E-learning), giảng viên có thể sử dụng hai phương pháp, cụ thể: *Một là*, sử dụng các bài giảng video, audio, file word,... gửi đến cho sinh viên, từ đó sinh viên theo dõi và thực hiện các nhiệm vụ mà giảng viên đưa ra (môi trường tự học). *Hai là*, giảng viên sử dụng mạng xã hội, vclass, hoặc phần mềm online trực tiếp phát hình ảnh, video, trao đổi trực tiếp giữa giảng viên và sinh viên.

Đối với phương pháp học E-Learning cung cấp lý thuyết của các môn học lý luận chính trị thông qua môi trường tự học. Trong môi trường này, phần lớn giảng viên, người hướng dẫn sẽ không xuất hiện trực tiếp, người học sẽ tự học, tự nghiên cứu thông qua các bài giảng được đăng tải trên hệ thống với các nội dung lý thuyết được số hóa theo nhiều định dạng khác nhau như:

Bài giảng đa phương tiện, Ebook, Audio, Video... Chatbot khi đó sẽ đóng vai trò như một trợ lý giảng dạy ảo (hoặc người bạn cùng học) nhằm hỗ trợ người học trong việc tra cứu các thông tin, cung cấp các dữ liệu có trong tài liệu của môn học lý luận chính trị hoặc tham khảo nội dung bên ngoài tài liệu. Bên cạnh đó, với khả năng tự học nhằm tăng tính thông minh, Chatbot sẽ sắp xếp, tự điều chỉnh, định hướng và gợi mở cho người học lý luận trực tuyến đối với những lần tương tác tiếp theo.

Đối với một số bài học, giảng viên và người học trực tiếp tham gia theo lịch học đã được bố trí trước, giảng viên thường bị quá tải với các công việc trên lớp học online của mình. Đặc biệt, sẽ càng khó khăn hơn với giảng viên giảng dạy lý luận chính trị trực tuyến, giảng viên vừa giảng bài, vừa trả lời các câu hỏi của người học, các câu hỏi có thể đến liên tục, dồn dập, vì là môn lý luận nên để trả lời các câu hỏi lý luận, cần lượng thông tin, kiến thức lớn. Những câu hỏi này thường tìm thấy trong các tài liệu nội dung của khóa học hoặc có sự liên quan nhưng người học vẫn hỏi như: Khái niệm, định nghĩa (*Tư tưởng Hồ Chí Minh, giá trị thặng dư, hàng hóa, tiền tệ...*); nguyên nhân, ý nghĩa của sự kiện lịch sử (*Cách mạng tháng Tám năm 1945, Cuộc kháng chiến chống Mỹ cứu nước...*); nội dung của một số vấn đề (*các nguyên lý, các quy luật, các cấp phạm trù của phép biện chứng duy vật; quá trình hình thành các chế độ, các thành phần trong các chế độ đó; quá trình giàu lên của chủ nghĩa tư bản; lý giải sứ mệnh của giai cấp công nhân...*) Ở đây, Chatbot đóng vai trò là người trợ lý sẽ phân loại và tự động trả lời các câu hỏi có sẵn trong danh mục đã được thiết lập hoặc các câu hỏi gợi mở có liên quan. Do đó, Chatbot giúp giảm tải áp lực cho giảng viên với các câu hỏi phổ biến. Như

vậy, người dạy sẽ có thêm nhiều thời gian quan tâm hơn đến người học, và ngược lại, người học lý luận chính trị trực tuyến cũng cảm thấy mình luôn được quan tâm và được hỗ trợ một cách kịp thời, tạo động lực để hoàn thành chương trình học hiệu quả hơn.

Môi trường E-learning cung cấp cho người học hệ thống tự luyện tập, tự đánh giá. Hệ thống này cung cấp bộ ngân hàng câu hỏi tự học, tự luyện nhằm cá nhân hóa quá trình luyện tập, tự kiểm tra, đánh giá của người học. Do vậy, Chatbot sẽ hỗ trợ người học hoàn thành bài luyện tập từng bước một, giúp người học củng cố lại lý thuyết hoặc tiếp cận lý thuyết mới theo cách hiệu quả hơn. Điều này rất phù hợp với các môn học lý luận chính trị, bởi với những môn lý luận chính trị, cần phải giải thích một cách tuần tự, logic, thậm chí giải thích bằng phương pháp truyền thống nếu giảng một lần học viên vẫn chưa hiểu hết vấn đề... thì ở đây, Chatbot có thể giải quyết được vấn đề này. Bên cạnh đó, các thông tin được cung cấp thông qua các trải nghiệm mới, thu hút và duy trì người học vào các cuộc trò chuyện, tương tác sống động, thực tế. Từ đó, góp phần khắc phục tâm lý ngại học các môn học lý luận chính trị của học sinh, sinh viên trong giai đoạn hiện nay.

2.4. Xây dựng mô hình Chatbot trong đào tạo, bồi dưỡng lý luận chính trị trực tuyến

Thông thường, lượng thông tin cần giải đáp lại xoay quanh một vấn đề chính nhưng lại mang tính chất lặp lại nhiều lần, trùng lặp trong câu trả lời dẫn đến lãng phí tài nguyên, nguồn lực con người, thời gian cũng như tiền bạc. Trong khi đó, thời gian làm việc hành chính chỉ có 8 giờ/ngày và 5 ngày trong tuần nhưng nhu cầu được tiếp cận thông tin từ học sinh, sinh viên, học viên

là 24/24, dẫn đến có những thông tin không được giải đáp kịp thời, phản hồi kém hiệu quả và công việc bị tích tụ theo thời gian. Việc này gây ra nhiều phiền toái, áp lực không cần thiết cho giảng viên đào tạo, bồi dưỡng lý luận chính trị. Ở đây, tác giả đề xuất giải pháp xây dựng ứng dụng chatbot, để giúp bot hiểu và giao tiếp được với con người thông qua đàm thoại văn bản, kỹ thuật xác định độ tương đồng ngữ nghĩa giữa câu hỏi đầu vào với tập không gian câu hỏi – câu trả lời chuẩn, là bộ dữ liệu được xây dựng để huấn luyện bot.

Theo đó, công nghệ Chatbot sẽ đưa ra bài toán và hướng xử lý như sau:

* **Bài toán:** Đề xuất câu trả lời phù hợp nhất cho một câu hỏi ở dạng văn bản đúng văn phạm bởi người dùng, trong đó:

Input gồm:

- Câu hỏi q cần tìm lời giải đáp;
- Tập QA có khuôn mẫu gồm các cặp Câu hỏi – Câu trả lời $\{q_k, a_k\}$, được biên soạn theo nội dung của quy chế 43/2007/QĐ-BGDĐT.

Output gồm:

Câu trả lời phù hợp được đề xuất cho Câu hỏi q ;

Độ phức tạp của bài toán tập trung chủ yếu vào vấn đề là *làm sao Bot có thể hiểu được ý định (Intents) của con người thông qua một câu đàm thoại ở dạng văn bản (text)?* Sau khi hiểu được ý định của con người thì hệ thống sẽ đề xuất câu trả lời phù hợp. Để trả lời cho câu hỏi trên, nhóm tác giả đã đề xuất giải pháp như sau:

- Xây dựng bộ dữ liệu chuẩn bao gồm các cặp: {Câu hỏi - Câu trả lời}, $QA = \{q_k, a_k | k \geq 1\}$, thông tin được biên soạn nằm trong giáo trình các môn lý luận chính trị của Bộ Giáo dục và Đào tạo hoặc các dữ liệu trên các

website chính thức của cơ quan Trung ương và địa phương (*dữ liệu tri thức*).

- Một câu hỏi q được đặt ra, sẽ được ánh xạ với câu trả lời a_j nằm trong cặp $\{q_j, a_j\}$, khi q có độ tương đồng về mặt ngữ nghĩa là lớn nhất với q_j .

Để xây dựng được bộ dữ liệu chuẩn cũng là một vấn đề cần được quan tâm, với ngữ pháp tiếng Việt, văn hóa vùng miền, từ ngữ địa phương, cũng như cách diễn đạt văn bản của mỗi người. Sau khi xây dựng được bộ QA ban đầu, trong quá trình vận hành Bot cần có khả năng tự học (*học tăng cường*) và tự hoàn thiện bộ QA này.

Các bước để xây dựng bộ dữ liệu chuẩn QA:

- **Bước 1:** Xây dựng tập dữ liệu QA, có khuôn dạng gồm các cặp Câu hỏi – Câu trả lời $\{q_k, a_k\}$, ban đầu chỉ có các câu trả lời a_k được biên soạn tóm tắt dựa theo nội dung của Quy chế 43;

- **Bước 2:** Xây dựng tập $Q' = \{Q_k | k \geq 1\}$, bằng cách thu thập tập các câu hỏi Q_k dựa theo câu trả lời a_k , bằng hình thức khảo sát các đối tượng là nhóm người dùng trong tương lai, một cách ngẫu nhiên, không phân biệt độ tuổi, giới tính và vùng miền;

- **Bước 3:** Sử dụng mô hình được đề xuất bởi *Samujjwal Ghosh, Maunendra Sankar Desarkar (2018) [8]*, để tính độ tương đồng giữa tập các câu hỏi trong Q_k với câu trả lời a_k , câu hỏi có độ tương đồng cao nhất q_k sẽ được thêm vào tập QA để tạo thành cặp $\{q_k, a_k\}$

*** Phương pháp nghiên cứu**

Thứ nhất, chuẩn bị dữ liệu:

+ Cấu trúc File .txt chứa tập QA:

Bộ QA được biên soạn dựa trên các dữ liệu tri thức, có cấu trúc:

[STT Câu] [Câu hỏi] ? [Câu trả lời]

Trong đó, STT câu được dùng để kiểm tra số lượng, tính chính xác của câu trả lời trong quá trình thử nghiệm.

1- Cho em hỏi, chúng ta có thể khái quát những nguồn gốc chính hình thành tư tưởng Hồ Chí Minh được không ạ?

Về nguồn gốc hình thành tư tưởng Hồ Chí Minh, em có thể khái quát thành những nguồn gốc chính sau đây:

- Truyền thống lịch sử - văn hóa của dân tộc Việt Nam
- Tinh hoa văn hóa nhân loại
- Chủ nghĩa Mác - Lênin: Cơ sở thế giới quan và phương pháp luận của tư tưởng Hồ Chí Minh
- Những nhân tố nội tại thuộc về phẩm chất con người Nguyễn Ái Quốc.

Em có thể tham khảo thêm Giáo trình Tư tưởng Hồ Chí Minh, NXB Chính trị quốc gia, Hà Nội, 2011 hoặc Tìm hiểu Tư tưởng Hồ Chí Minh, NXB Lý luận chính trị, Hà Nội, 2004.

2- Quan điểm về xây dựng một nền văn hóa mới của Hồ Chí Minh là như thế nào ạ?

Cùng với định nghĩa về văn hóa, Hồ Chí Minh còn đưa ra Năm điểm lớn định hướng cho việc xây dựng nền văn hóa dân tộc:

- Xây dựng tâm lý: tinh thần độc lập tự cường
- Xây dựng luân lý: biết hy sinh mình, làm lợi cho quần chúng
- Xây dựng xã hội: mọi sự nghiệp có liên quan đến phúc lợi của nhân dân trong xã hội
- Xây dựng chính trị: dân quyền
- Xây dựng kinh tế

Như vậy, ngay từ rất sớm, Hồ Chí Minh đã quan tâm đến văn hóa, đã thấy rõ vai trò, vị trí của văn hóa trong đời sống xã hội. Điều này rất đáng chú ý vì ngay sau khi giành được độc lập, Hồ Chí Minh đã bắt tay vào việc xây dựng, kiến tạo một nền văn hóa mới ở Việt Nam trên tất cả mọi lĩnh vực, từ kinh tế, chính trị, xã hội, đạo đức đến tâm lý con người, đã sớm đưa văn hóa vào chiến lược phát triển đất nước.

Em có thể tham khảo thêm Giáo trình Tư tưởng Hồ Chí Minh, NXB Chính trị quốc gia, Hà Nội, 2011 hoặc Tìm hiểu Tư tưởng Hồ Chí Minh, NXB Lý luận chính trị, Hà Nội, 2004.

Hình 3. Ví dụ về các cặp Câu hỏi – Câu trả lời trong tập QA.

+ File .csv chứa tập Q’:

Tập Q’={Q_k | k>=1}, là bao gồm tập các câu hỏi Q_k gắn với từng câu trả lời {a_k}

Q’ được làm sạch dữ liệu bằng cách loại bỏ, điều chỉnh những câu hỏi/từ không phù hợp với câu trả lời, không hướng vào câu trả lời, lỗi soạn thảo, trình bày khi làm khảo sát,... Tiếp theo, thực hiện tính độ tương đồng giữa các câu hỏi trong tập Q_k với câu trả lời a_k bằng phương pháp TF-IDF (đo tần số suất hiện của 1 từ trong câu hỏi). Câu hỏi có độ tương đồng cao nhất q_k là câu hỏi chuẩn, sẽ được rút khỏi Q’ và ghép vào bộ

dữ liệu chính QA để tạo thành cặp {q_k, a_k}. Các câu hỏi còn lại trong Q’ có độ tương đồng thấp quá thấp sẽ bị loại bỏ, chỉ giữ lại những câu hỏi có độ tương đồng từ 40% trở lên. Q’ được sử dụng để bổ trợ cho bộ QA, khi gọi ý định của người dùng thông qua câu hỏi.

Trong quá trình tương tác người – máy, Bot sẽ tự động lưu lại những câu hỏi có độ tương đồng thấp vào 1 file JSON, có thể sử dụng những câu hỏi này để huấn luyện tăng cường cho bot về sau.

+ Tổ chức dữ liệu cho Chatbot

Dựa vào bộ dữ liệu QA và Q’, chúng ta sẽ sử dụng thư viện Pandas để tổ chức dữ liệu cho Chatbot. Sử dụng chức năng cơ bản của thư viện Pandas, tách dòng dựa trên bố cục trình bày với sep = **None**. Trong trường hợp chúng ta muốn tách dựa trên dấu chấm câu thì sử dụng sep = “.”. Tách Câu hỏi – Câu trả lời ra hai cột khác nhau và bổ sung Câu hỏi bổ sung dựa trên STT đã sắp xếp. Tiếp theo, ta xác định số cột tương ứng với số câu hỏi – trả lời QA, tiến hành nhập dữ liệu từ file .csv vào bảng dữ liệu của Chatbot.

Với cách tổ chức này, Chatbot có thể thực hiện ánh xạ như sau:

$$\{Q_i, Q_{ij}\} \rightarrow A_i$$

Q: Câu hỏi

A: Câu trả lời

i: Số STT dòng

j: Số STT của câu hỏi trong tập câu hỏi bổ sung (ứng với STT dòng trong bảng .csv)

Thứ hai, phân tích ngữ nghĩa câu

Ví dụ 1: “Cho em hỏi, chúng ta có thể khái quát những nguồn gốc chính hình thành tư tưởng Hồ Chí Minh được không ạ?”

Ví dụ 2: “cho Em hỏi chúng ta có thể khái quát những nguồn gốc chính hình thành TT HCM được không ạ?”

Theo ngữ nghĩa thì ví dụ 1 và ví dụ 2 tương đồng với nhau, nếu tách ví dụ 1 và ví dụ 2 thành các cặp câu phân tách như sau:

[Cho em hỏi chúng ta có thể khái quát] (A) + [những nguồn gốc chính] (B) + [hình thành tư tưởng Hồ Chí Minh được không ạ?] (C).

[cho Em hỏi chúng ta có thể khái quát](A2) + [những nguồn gốc chính](B2) + [hình thành TT HCM được không ạ?] (C2)

Vậy ta có bảng sự khác biệt giữa cách hiểu của con người và cách hiểu của máy móc như sau:

Cặp câu phân tách	Con người	Máy móc
A – A2	Giống nhau	Khác nhau
B – B2	Giống nhau	Giống nhau
C – C2	Giống nhau về mặt tương đồng ngữ nghĩa đối với câu hỏi	Khác nhau hoàn toàn

Đến đây ta nhận thấy cặp A – A2 đang nhận kết quả False về sự tương quan, bởi kí tự “e” trong “em” - “Em” và kí tự “C” trong “Cho” - “cho” khiến cho 2 cặp từ này trở nên không giống nhau. Hướng xử lý cho vấn đề không đồng nhất về kí tự hoa hay kí tự thường sẽ được xử lý bằng cách chuyển tất cả các ký tự hoa về ký tự thường: Sử dụng lệnh lower sẽ đưa cặp A – A2 về sự đồng nhất. Tiến hành tách từ:

- Tách từ theo phương pháp thông thường, dựa trên khoảng cách trống được biểu thị “ ”

- Tách từ theo loại từ.

Tính chính xác về loại từ, kích thước từ (từ đơn, từ ghép) được xác định khi sử dụng cách thứ 2

Để đảm bảo sự đa dạng trong ngôn từ hỏi, cần tiến hành khảo sát người dùng và bổ sung vào dữ liệu những cụm từ phổ biến. Việc khảo sát thu thập dữ liệu sẽ hỗ trợ cho việc học hỏi của Chatbot, gia tăng dữ liệu và mở rộng tính chính xác.

Khi tiến hành tách từ kết quả sẽ được lưu ở dạng list, mỗi phần tử ứng với từng index trong câu đều được xem là 1 khóa – key, các key này khi gộp lại sẽ đại diện cho câu trong công thức tính TF-IDF. Tuy nhiên, vẫn còn những key không cần thiết để đại diện cho câu, như là dấu hỏi, chấm phẩy,... và những từ thừa, nên loại bỏ bớt để tránh làm loãng giá trị thực tế của câu.

Thứ ba, tiến hành so sánh giữa câu hỏi đầu vào và dữ liệu:

+ Sử dụng phương pháp TF-IDF

Tính TF – tần số của 1 từ xuất hiện trong 1 văn bản:

$$tf(t, d) = \frac{f(t, d)}{\max\{f(w, d) : w \in d\}}$$

Tính IDF – inverse document frequency. Tần số nghịch của 1 từ trong tập văn bản (corpus). Tính IDF để giảm giá trị của những từ phổ biến. Mỗi từ chỉ có 1 giá trị IDF duy nhất trong tập văn bản:

$$idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$$

+ Giá trị TF-IDF:

$$tfidf(t, d, D) = tf(t, d) \times idf(t, D)$$

Để thực hiện kỹ thuật tính TF-IDF, cần xác định rõ D bao gồm những đối tượng nào trong dữ liệu đã chuẩn bị, D – “Tổng hợp” bao gồm số câu hỏi trong bộ QA và số câu hỏi bổ sung trong bộ Q’.

+ Tính độ tương đồng giữa các câu văn bản **Error! Reference source not found.**

$$TF = \frac{\text{Số lần từ đó xuất hiện trong câu đang xét}}{\text{Độ dài câu đang xét}}$$

$$IDF = \log \frac{\text{Độ dài tập D}}{\text{Từ đó xuất hiện bao nhiêu lần trong D}}$$

+ Lựa chọn mô hình so sánh độ tương đồng

Để khớp một câu hỏi q với một câu trả lời phù hợp nhất trong tập QA, phải tìm ra phương án để tìm lời giải tối ưu, 2 phương án tác giả đề xuất là:

Phương án 1: Tìm độ tương đồng giữa câu hỏi đầu vào q với tập các câu hỏi trong QA, chọn ra a_k là câu trả lời cần tìm nếu a_k có độ tương đồng với q là cao nhất;

Phương án 2: Tìm độ tương đồng giữa câu hỏi đầu vào q với các câu hỏi trong tập dữ liệu chuẩn QA, trong cặp $\{a_k, q_k\}$ chọn ra a_k là câu trả lời cần tìm nếu q_k có độ tương đồng với q là cao nhất;

Như vậy, thực hiện theo tuần tự quy trình trên chúng ta sẽ triển khai thành công hệ thống mô phỏng ChatBot trên Bot Framework Dialogflow, hoạt động đàm thoại văn bản với con người theo mô hình QnA, hoạt động trên đa nền tảng nhắn tin, như: Skype, Facebook, Website, Slack, Viber, ... Với cách tiếp cận của bài viết là:

Đề xuất câu trả lời bằng cách tìm độ tương đồng giữa câu hỏi đầu vào với tập câu hỏi nằm trong tập dữ liệu QA được thiết kế sẵn, nên độ chính xác phụ thuộc lớn vào khâu thiết kế tập QA; Phương pháp TF-IDF dựa trên mô hình xác suất, phụ thuộc lớn vào mạng ngữ nghĩa (WordNet), corpus các từ đồng nghĩa tiếng Việt; các kỹ thuật tiền xử lý văn bản tiếng Việt như tách từ, xác định loại từ,...

3. Kết luận

Việc áp dụng công nghệ vào công tác đào tạo, bồi dưỡng lý luận chính trị trong thời đại cách mạng công nghiệp 4.0 là hết sức cần thiết. Chatbot là một trong những công nghệ đơn giản, dễ áp dụng, nhưng mang lại hiệu quả khả quan. Chatbot kết hợp hoàn hảo giữa phương thức học tập truyền thống và việc tích hợp ứng dụng công nghệ thông tin trong đào tạo, qua đó, nâng cao tính linh động, chủ động trong việc lĩnh hội kiến thức của người học cũng như tiết kiệm chi phí, rút ngắn không gian, khoảng cách địa lý giữa giảng viên và sinh viên. Ứng dụng công nghệ Chatbot trong đào tạo, bồi dưỡng lý luận chính trị trực tuyến góp phần đáp ứng nhu cầu thông tin lý luận kịp thời, ngoài ra còn tạo hứng thú hơn đối với sinh viên khi học tập các môn lý luận chính trị. Tuy nhiên, khoa học, công nghệ không ngừng phát triển, cần có thêm những nghiên cứu góp phần hoàn thiện, phát triển hệ thống Chatbot với những tính năng vượt trội hơn trong tương lai.

TÀI LIỆU THAM KHẢO

- Dương Thăng Long (2014), *Nghiên cứu đo độ tương đồng văn bản trong tiếng Việt và ứng dụng*, Viện Đại học Mở Hà Nội;
- Đỗ Thị Thanh Nga (2010), *Tính toán độ tương tự ngữ nghĩa văn bản dựa vào độ tương tự giữa từ với từ*, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội;
- Phạm Quang Nhật Minh-FTRI, *3 vấn đề NLP cơ bản khi phát hiện một hệ thống chatbot và một số phương pháp giải quyết điển hình*, Techinsight, ngày 13/04/2017
- Abbad, M. M., Morris, D., & de Nahlik, C. (2009). Looking under the Bonnet: Factors Affecting Student Adoption of E-Learning Systems in Jordan. *The International Review of Research in Open and Distance Learning*
- Brain [BRN.AI] code for equity, *Chatbot Report 2019: Global Trends and Analysis*, Chatbot Magazine, 19/04/2019
- Jennex, M.E. (2005). *Case Studies in Knowledge Management*. Idea Group Publishing: Hersley.
- Oblinger, D.G., and Hawkins, B.L. "The Myth about students," *Educause Review*, 2005, July/August.
- Samujjwal Ghosh, Maunendra Sankar Desarkar (2018), *Class Specific TF-IDF Boosting for Short-text Classification: Application to Short-texts Generated During Disasters*. Companion Proceedings of the The Web Conference 2018;
- Twigg C. (2002). Quality, cost and access: the case for redesign. In *The Wired Tower*. Pittinsky MS (ed.). Prentice-Hall: New Jersey
- Vishnu Elupula, *How do Chatbots work? An overview of the architecture of Chatbots, Bigdata-madesimple*, 15/05/2019
- Welsh ET, Wanberg CR, Brown EG, Simmering M.J. (2003). E-learning: emerging uses, empirical results and future directions. *International Journal of Training and Development* 2003(7): 245-258