

**NGHIÊN CỨU, THIẾT KẾ TRỢ LÝ ẢO CÁ NHÂN TRÊN NỀN TẢNG  
WEB KẾT HỢP HỌC MÁY (MACHINE LEARNING) VÀ XỬ LÝ  
NGÔN NGỮ TỰ NHIÊN (NATURAL LANGUAGE PROCESSING)**  
RESEARCH, DESIGN, WEB-BASED PERSONAL VIRTUAL ASSISTANCE  
COMBINING MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING

*SV. Tăng Xuân Biên<sup>1</sup>, SV. Trần Minh Chiến<sup>1</sup>, SV. Lê Thanh Nga<sup>1</sup>,  
SV. Trần Bá Hiến<sup>1</sup>, SV. Nguyễn Đức Nam Bình<sup>1</sup>,  
TS. Hà Thị Kim Duyên<sup>1</sup>, TS. Ngô Mạnh Tiến<sup>2</sup>*

<sup>1</sup>*Khoa Điện tử, Trường Đại học Công nghiệp Hà Nội*

<sup>2</sup>*Viện Vật lý, Viện hàn lâm Khoa học và Công nghệ Việt Nam*

*Email: [ha.duyen@hau.edu.vn](mailto:ha.duyen@hau.edu.vn)*

*Số điện thoại: 0988901420*

## **TÓM TẮT**

Bài báo trình bày về ứng dụng trí tuệ nhân tạo, máy học, xử lý ngôn ngữ tự nhiên xây dựng giải pháp chatbot giao tiếp giữa robot với người - IVASTChatbot. Hệ thống IVASTChatbot cho phép robot có khả năng nhận dạng được thông tin mà con người truyền đạt qua giọng nói, thực hiện các bước xử lý và cuối cùng trả về thông tin chính xác mà con người cần biết qua văn bản, giọng nói. Một phương pháp xây dựng hệ thống chatbot dựa trên kiến trúc Task-oriented Dialogue Systems (TODs) và phát triển trên nền tảng mã nguồn mở Rasa Framework được trình bày, hệ thống còn thực hiện tích hợp thêm các dịch vụ Speech To Text (STT) và Text To Speech (TTS) để cung cấp tương tác bằng giọng nói cho robot. Quá trình tương tác giọng nói giữa người và robot được thông qua loa, micro và hiển thị bằng màn hình LED được trang bị trên robot, con người có thể dễ dàng hiểu được. Các kết quả mô phỏng và quá trình chạy thử nghiệm đã cho thấy khả năng giao tiếp linh hoạt, chính xác được thực hiện trên tập kiểm tra. Kết quả của bài báo là một ứng dụng có thể được áp dụng trong dịch vụ, y tế, bán hàng,...

***Từ khóa:** Trí tuệ nhân tạo (AI), Chuyển giọng nói thành văn bản (STT), Chuyển văn bản thành giọng nói (TTS), Quản lý hội thoại (DM), Trình theo dõi trạng thái đối thoại (DST), Xử lý ngôn ngữ tự nhiên (NLP), Sinh ngôn ngữ tự nhiên (NLG), Hiểu ngôn ngữ tự nhiên (NLU).*

## **ABSTRACT**

The article presents the application of artificial intelligence, machine learning and natural language processing to build a chatbot solution to communicate between robots and

humans - IVASTChatbot. The IVASTChatbot system allows the robot to be able to recognize the information communicated by humans through voice, perform processing steps, and finally return the correct information that humans need to know through voice. A method of building a chatbot system based on the Task-oriented Dialogue Systems (TODs) architecture and developed on the open source Rasa Framework platform is presented, the system also integrates the Speech To Text services (STT) and Text To Speech (TTS) to provide voice interaction for robots. The process of voice interaction between human and robot is through speaker and microphone and displayed by LED screen equipped on robot, human can easily understand. The simulation results and the test run have shown that flexible and accurate communication is performed on the test set. The result of the article is an application that can be applied in service, healthcare, sales, etc

**Keywords:** *Artificial Intelligence (AI), Speech to Text (STT), Text to Speech (TTS), Dialog Management (DM), Dialogue State Tracker (DST), Natural Language Processing (NLP), Natural Language Generation (NLG), Natural Language Understanding (NLU).*

### CHỮ VIẾT TẮT

<i>NLP</i>	<i>Natural Language Processing</i>
<i>NLG</i>	<i>Natural Language Generation</i>
<i>NLU</i>	<i>Natural Language Understanding</i>
<i>ROS</i>	<i>Robot Operating System</i>
<i>AI</i>	<i>Artificial Intelligence</i>
<i>DM</i>	<i>Dialog Management</i>
<i>DST</i>	<i>Dialogue State Tracker</i>
<i>TTS</i>	<i>Text to Speech</i>
<i>STT</i>	<i>Speech to Text</i>

### 1. GIỚI THIỆU

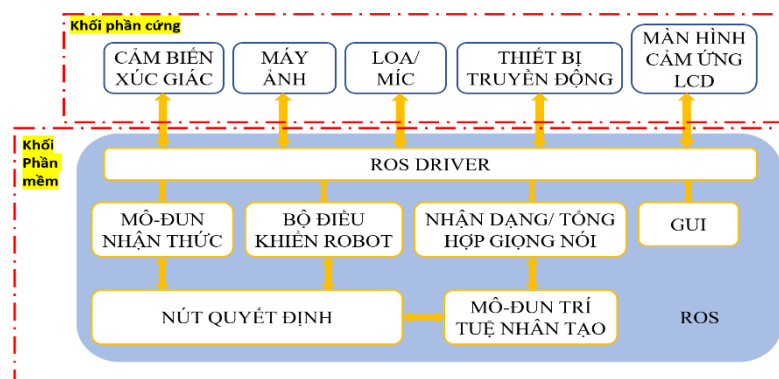
Trí tuệ nhân tạo, máy học và học sâu đang trở nên rất phổ biến ngày nay. Tất cả các công nghệ này được liên kết với nhau và mục tiêu chung là bắt chước trí thông minh của con người. Có rất nhiều ứng dụng cho các lĩnh vực này như suy luận logic (Logical Reasoning), biểu diễn tri thức (knowledge representation), lập kế hoạch (Planning), học tập (learning), xử lý ngôn ngữ tự nhiên (natural language processing), nhận thức (perception), trong đó đặc biệt là trí tuệ xã hội (social intelligence), đây cũng chính là một trong những lĩnh vực ứng dụng của AI. Sử dụng AI, chúng ta có thể xây dựng trí thông minh xã hội trong một cỗ máy hoặc robot – robot xã hội (social robot). Nói một cách dễ hiểu, robot xã hội (social robot) là bạn đồng hành cá nhân hoặc robot hỗ trợ có thể tương tác với con người bằng giọng nói, thị giác và cử chỉ. Những con robot này hoạt động giống như một

con người có thể hiểu ngôn ngữ của con người và có thể truyền đạt trao đổi với chúng bằng ngôn ngữ giọng nói.

Một hệ thống chatbot dựa trên sự tương tác giữa người và robot qua ngôn ngữ giọng nói (IVASTChatbot) được trình bày trong bài báo là một hệ thống con của “Robot dạng người thông minh IVASTBot ứng dụng trong giao tiếp, phục vụ con người”[9]. Hệ thống IVASTChatbot được thiết kế chủ yếu cho ba mục tiêu: một là khả năng của robot nhận dạng ra ngôn ngữ giọng nói của con người, hai là khả năng robot có thể xử lý và đưa ra kết quả, ba là robot có thể thực hiện giao tiếp truyền đạt ngôn ngữ giọng nói.

## 2. KIẾN TRÚC TỔNG QUAN ROBOT DẠNG NGƯỜI THÔNG MINH IVASTBot

Robot thông minh IVASTBot có khả năng nhận thức thế giới bằng camera, tương tác với con người bằng giọng nói, cử chỉ và đưa ra các quyết định bằng thuật toán trí tuệ nhân tạo. Robot này có thiết kế như sơ đồ khối sau[9]:



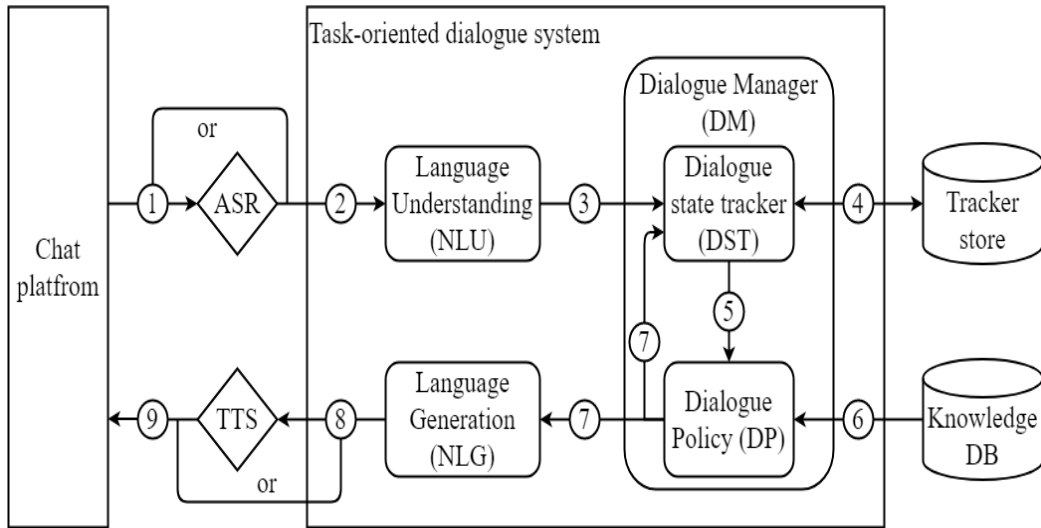
**Hình 2. Sơ đồ khối của IVASTBot**

Khối phần cứng gồm máy ảnh có độ phân giải cao được sử dụng để chụp các bức ảnh thời gian thực về nét mặt và cử chỉ cơ thể của người dùng; micro được sử dụng để thu thập tín hiệu lời nói; thiết bị truyền động giúp robot chuyển động đầu, cơ thể, cánh tay; động cơ và bánh đa hướng omni được tích hợp để điều hướng; màn hình LCD hiển thị giao diện và tương tác với người dùng.

Bên trong khối phần mềm, gồm các mô-đun nhận thức giúp xử lý dữ liệu camera và tìm kiếm các đối tượng cần thiết từ hiện trường; mô-đun nhận dạng / tổng hợp giọng nói giúp giao tiếp với người dùng; mô-đun trí tuệ nhân tạo; mô-đun điều khiển robot để điều khiển thiết bị truyền động; nút quyết định kết hợp tất cả dữ liệu từ các cảm biến và đưa ra quyết định cuối cùng về việc phải làm tiếp theo Hệ điều hành ROS kết nối tới các cảm biến, bộ truyền động. Khối GUI giúp giao tiếp giữa người dùng với robot thông qua thao tác với hình ảnh trên bảng điều khiển LCD.

## 3. Giải pháp xây dựng

### A. Tổng quan các thành phần xử lý trong chatbot



**Hình 3. Kiến trúc mô hình tổng quan các thành phần xử lý trong Chatbot**

Kiến trúc Task-oriented Dialogue Systems (TODs) [5] được sử dụng để xây dựng Chatbot và phát triển trên nền tảng mã nguồn mở Rasa framework, bên cạnh đó hệ thống Chatbot còn thực hiện tích hợp thêm các dịch vụ Automatic Speech Recognition (ASR) và Text to Speed (TTS) để cung cấp tương tác bằng giọng nói cho IVASTChatbot.

Mỗi thành phần trong Chatbot đều có vai trò riêng biệt

**NLU:** Chịu trách nhiệm chuyển đổi tin nhắn văn bản của người dùng thành dạng dữ liệu có cấu trúc đã được định nghĩa từ trước. Dạng dữ liệu có cấu trúc này chính là các Intents, Entities.

**DST:** Chịu trách nhiệm theo dõi và cập nhật trạng thái của cuộc hội thoại. Có 2 luồng xử lý riêng biệt trong module này bao gồm: luồng 1 cập nhật trạng thái được kích hoạt bởi module NLU, luồng 2 cập nhật trạng thái kích hoạt bởi Dialogue Policy.

**DB:** Thực hiện dự đoán hành động kế tiếp mà Chatbot cần thực hiện dựa trên trạng thái cuộc hội thoại được gửi tới từ DST.

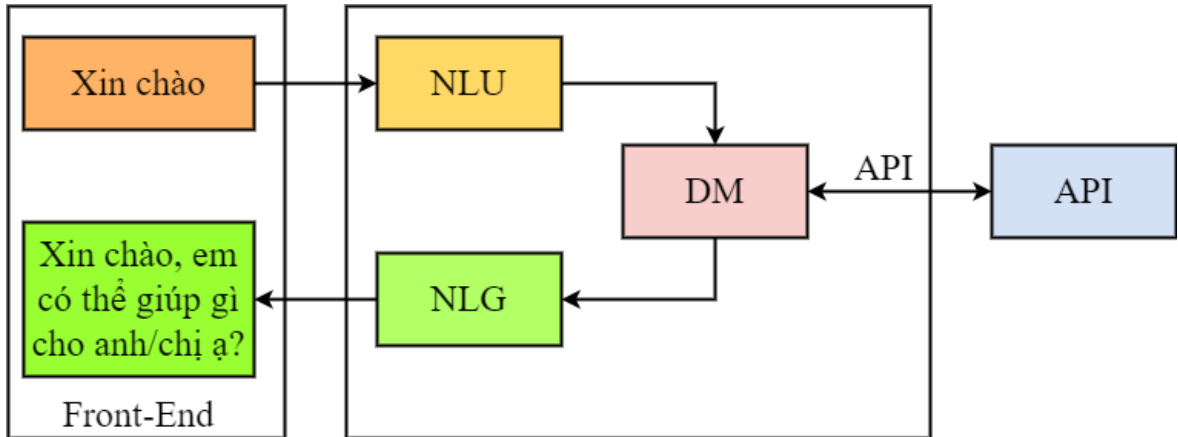
**NLG:** Chịu trách nhiệm tạo ra câu trả lời bằng ngôn ngữ tự nhiên từ kết quả của module DP. Phương pháp truyền thống là sử dụng một bộ các mẫu câu có sẵn kết hợp với kết quả từ DP để tạo ra câu trả lời.

### B Môi trường thực nghiệm

Chương trình thử nghiệm được thiết kế, xây dựng và thực hiện trên môi trường hệ điều hành Windows với nền tảng framework Rasa, dựa trên ngôn ngữ lập

trình python. Giao diện người dùng sử dụng nền tảng web/ứng dụng chat.

### C. Thiết kế giao diện người dùng tương tác



**Hình 4. Kiến trúc chung của hệ thống**

Front-end sử dụng giao diện web hoặc các trình nhắn tin phổ biến. Với mục tiêu minh họa, ở đây nhóm nghiên cứu sử dụng giao diện web messenger.

- Mỗi khi có một người dùng gửi tin nhắn cho chatbot thì nội dung tin nhắn này sẽ gửi một POST request để webhook được sử dụng để lắng nghe sự kiện. Webhook này sẽ chuyển đến bộ NLU của RASA.

- RASA nhận diện ý định, sau khi đã thu được message của người dùng thì sử dụng RASA để hiểu ý định của người dùng cùng các thông tin thực thể

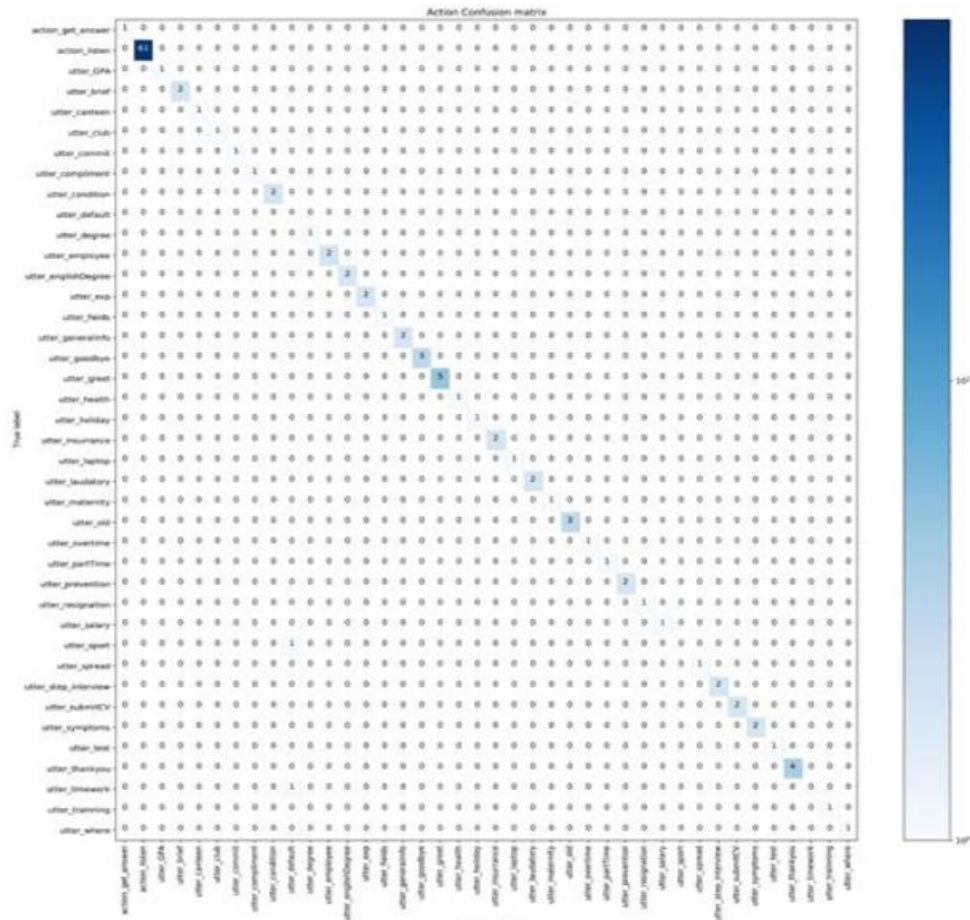
- Thông tin này tiếp tục chuyển đến DM của RASA, tại đây tùy theo ý định và thông tin thực thể cùng với các thông tin theo dõi của cuộc trò chuyện đã xảy ra cho đến nay, để dự đoán một phản ứng thích hợp, bao gồm cả việc gọi API để lấy thông tin trả lời.

- NLG sinh ra câu trả lời dựa vào dữ liệu từ thành phần DM theo các mẫu câu template đã được xây dựng trước hoặc là kết quả của API.

- Gửi tin nhắn qua phản hồi trả về cho người dùng.

### D. Kết quả thực nghiệm

Kết quả đánh giá NLU model và Rasa Core sau khi thực hiện đào tạo chatbot và kiểm tra trên dữ liệu test, dữ liệu người dùng nhập vào:



**Hình 5. Intent Confusion matrix**

**Bảng 1. Đánh giá trích chọn thông tin thực thể (entity)**

	Precision	Recall	F1-score	support
Room_type	0.85	0.75	0.79	8
location	1.00	0.58	0.74	17

Tính chung, kết quả test trên tập dữ liệu cho độ chính xác khoảng 81%

Correct: 131/214

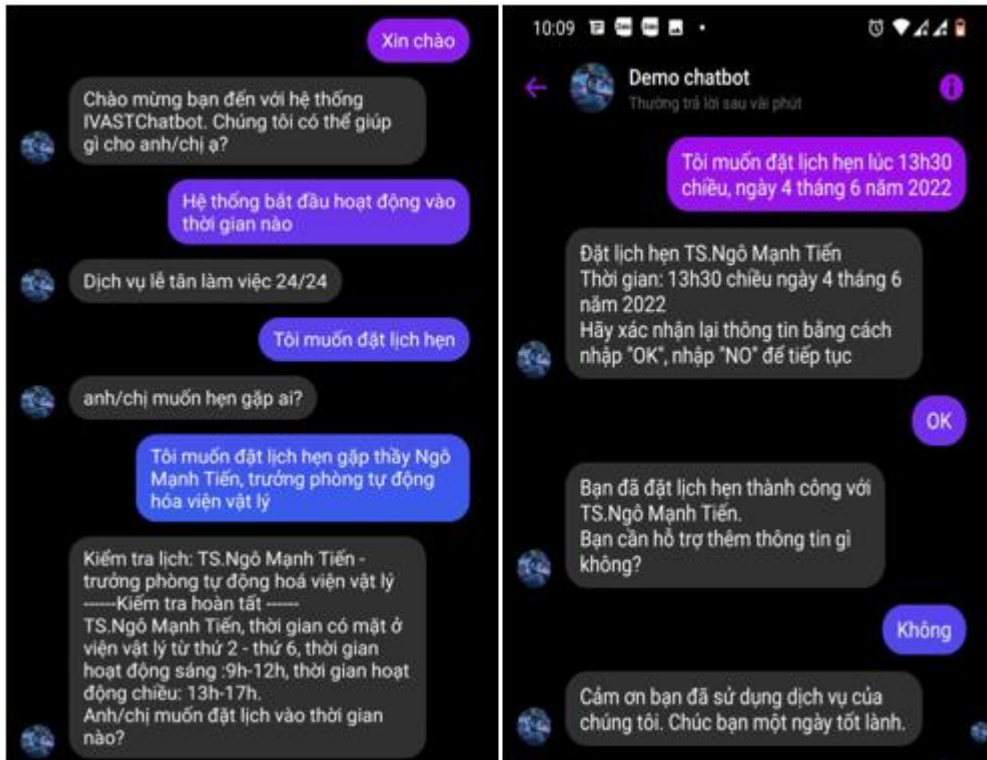
F1-Score: 0.906

Precision: 0.926

Accuracy: 0.899

### **Kiểm thử trên giao diện người dùng**

Thực hiện thử nghiệm tương tác với chatbot qua một số câu hỏi gần với kịch bản đã đào tạo cho chatbot



**Hình 6. Ví dụ về đặt lịch hẹn**

Người dùng cũng có thể thực hiện “trao đổi” những thông tin khác với chatbot.

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL
2022-06-09 05:07:21 INFO root - Rasa server is up and running.
Bot loaded. Type a message and press enter (use '/stop' to exit):
Your input -> xin chào
Chào bạn
Your input -> hắt đi
Mừng tám tháng ba, em ra thăm vườn, cắm một bông hoa xinh tươi tặng cô giáo
Your input -> giúp tôi với
Tôi có thể giúp gì bạn
Your input -> muốn đặt lịch hẹn
Anh/Chị muốn hẹn gặp ai?
Your input -> gặp tiến sĩ Tiến
Anh/Chị muốn gặp TS. Tiến vào thời gian nào?
Your input -> 10/9/2023
? Đặt lịch hẹn TS. Tiến vào 0 giờ 0 phút, ngày 10 tháng 9 năm 2023.\nThông tin trên đã đúng chưa ạ? 1
: Đúng (/dong_y)
Oke cảm ơn anh/chị đã sử dụng dịch vụ
Bạn đã đặt lịch thành công với TS. Tiến vào 0 giờ 0 phút, ngày 10 tháng 9 năm 2023
Your input -> tạm biệt
Hẹn gặp lại bạn
Your input -> hắt đi
Mừng tám tháng ba, em ra thăm vườn, cắm một bông hoa xinh tươi tặng cô giáo
Your input -> kể chuyện cười được không
Hôm nay, một người đàn ông gõ cửa nhà tôi và yêu cầu đóng góp cho hồ bơi của xóm. Tôi đưa anh ta một cốc nước
Your input -> 

```

**Hình 7 Một số giao tiếp cơ bản khác**

### E. Đánh giá

Từ kết quả thực nghiệm rút ra một số đánh giá như sau:

- Xác định đúng được ý định (intent) có ý nghĩa quan trọng nhất đối với chatbot. Đối với bài toán trong miền đóng cần xác định rõ ràng các intent, xây dựng

tập dữ liệu đủ lớn, gán nhãn và tiến hành training.

- Xây dựng dữ liệu đào tạo, training cho chatbot với các kịch bản là rất cần thiết để cho độ chính xác cao của chatbot.

- Chatbot ứng dụng AI có khả năng đáp ứng tốt với các kịch bản dựng sẵn, và được đào tạo. Đối với các kịch bản nằm ngoài kịch bản dựng sẵn, có thể tăng cường khả năng cho chatbot bằng cách điều hướng người dùng về các câu mặc định hoặc các dạng giao diện menu lựa chọn.

- Việc xác định và phản hồi đa ý định có thể thực hiện bằng việc kết hợp các ý định.

- Qua bài toán thực nghiệm có thể thấy rằng áp dụng bài toán Chatbot cho việc hỗ trợ trả lời thông tin khách sạn là khả thi, có tính thực tiễn cao, và hoàn toàn áp dụng được ngay trong thực tiễn.

#### 4. Ứng dụng trên robot thông minh



*Hình 7. Ứng dụng chatbot trên robot thông minh*

Kết quả của đề tài có thể được sử dụng, tiếp tục phát triển để có thể triển khai trên robot. Để làm cho robot có thể giao tiếp, tương tác với con người qua văn bản hoặc ngôn ngữ giọng nói.

#### 5. Kết luận

Kết quả của việc huấn luyện và test dữ liệu cho kết quả khá cao trên tập dữ liệu huấn luyện và test. Nếu có được nguồn dữ liệu lớn hơn để đào tạo thì độ chính



xác sẽ cao hơn và hệ thống sẽ thông minh hơn. Kết quả thử nghiệm cho thấy tính khả thi của giải pháp này khi thực hiện ứng dụng trên robot dạng người thông minh robot có thể trò chuyện, tương tác với con người qua ngôn ngữ giọng nói. Với định hướng nghiên cứu tiếp theo chúng tôi sẽ tiếp tục xử lý những vấn đề còn hạn chế như xử lý các lỗi chính tả, viết tắt dựa trên các thuật toán máy học, nhằm tăng hiệu suất của giải pháp hơn.

## TÀI LIỆU THAM KHẢO

- [1] Nguyen Thi Mai Trang, Maxim Shchebakov, “Enhancing Rasa NLU model for Vietnamese chatbot”, *International Journal of Open Information Technologies ISSN: 2307 – 8162 vol.9, no.1, 2021*
- [2] Wei-Lun Chao, “Machine Learning Tutorial”. *National Taiwan DISP Lab, National Taiwan University. 2011: - ngày truy cập 20/5/2022.*
- [3] Siwar Chibani, François-Xavier Coudert, “Machine learning approaches for the prediction of materials properties”, *HAL Id: hal-02911837,2020. pp 4-5: -ngày truy cập 20/5/2022.*
- [4] [https://vi.wikipedia.org/wiki/Xử\\_lý\\_ngôn\\_ngữ\\_tự\\_nhiên](https://vi.wikipedia.org/wiki/Xử_lý_ngôn_ngữ_tự_nhiên) -ngày truy cập 30/5/2022.
- [5] Pham Nam, “Tổng quan về Chatbot,” Viblo, 5-Nov-2021.[Online].Available: <https://viblo.asia/p/tong-quan-ve-chatbot-yMnKMByaZ7P>. [Accessed:3-Jun-2022].
- [6] P. H. Quang, “Rasa chatbot: Tăng khả năng chatbot với custom component và custom tokenization(tiếng Việt tiếng Nhật),” Viblo, 16-Mar-2020. [Online]. Available:[https://viblo.asia/p/rasa-chatbottang-kha-nang-chatbot-voi-custom-component-va\\_customtokenizationtiengviet-tieng-nhat-Qbq5QN4mKD8](https://viblo.asia/p/rasa-chatbottang-kha-nang-chatbot-voi-custom-component-va_customtokenizationtiengviet-tieng-nhat-Qbq5QN4mKD8). [Accessed:3-Jun-2022]
- [7] T. Nguyen and M. Shcherbakov, “A Neural Network based Vietnamese Chatbot,” in 2018 International Conference on System Modeling & Advancement in Research Trends (SMART), 2018
- [8] Chu Le Long, “Nghiên cứu, xây dựng chatbot hỏi đáp thông tin khách sạn sử dụng rasa framework”, *ngày truy cập 20/5/2022.*
- [9] Nguyễn Thị Duyên, Ngô Mạnh Tiến, Hà Thị Kim Duyên, Bùi Quang Tuấn, Trần Bá Hiến, Nguyễn Minh Đông, Đỗ Quang Hiệp, "Xây dựng hệ điều hướng trên bản đồ, định vị SLAM cho Robot tự hành trong nhà kính nông nghiệp dựa trên hệ điều hành ROS," *Hội nghị - Triển lãm quốc tế lần thứ 4 về Điều khiển và Tự động hoá (VCCA), 2021*