

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ



NGUYỄN TẤT TIẾN

**NGHIÊN CỨU VÀ XÂY DỰNG CHATBOT HỖ TRỢ
NGƯỜI DÙNG TRONG NGÂN HÀNG**

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

HÀ NỘI – 2019

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

NGUYỄN TẤT TIẾN

**NGHIÊN CỨU VÀ XÂY DỰNG CHATBOT HỖ TRỢ
NGƯỜI DÙNG TRONG NGÂN HÀNG**

Ngành: Kỹ thuật phần mềm

Chuyên ngành: Kỹ thuật phần mềm

Mã số: 8480103.01

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

HÀ NỘI – 2019

LỜI CAM ĐOAN

Tôi là Nguyễn Tất Tiến, học viên khóa K22, ngành Kỹ thuật phần mềm, chuyên ngành Kỹ Thuật Phần Mềm. Tôi xin cam đoan luận văn “Nghiên cứu và xây dựng chatbot hỗ trợ người dùng trong ngân hàng” là do tôi nghiên cứu, tìm hiểu và phát triển dưới sự dẫn dắt của TS. Nguyễn Văn Vinh. Luận văn không phải sự sao chép từ các tài liệu, công trình nghiên cứu của người khác mà không ghi rõ trong tài liệu tham khảo. Tôi xin chịu trách nhiệm về lời cam đoan này.

Hà nội, ngày tháng năm 2019

LỜI CẢM ƠN

Đầu tiên tôi xin gửi lời cảm ơn tới các thầy cô trường đại học Công Nghệ, Đại học Quốc Gia Hà Nội đã tận tình giảng dạy và truyền đạt kiến thức trong suốt khóa cao học vừa qua. Tôi cũng xin được gửi lời cảm ơn đến các thầy cô trong bộ môn Kỹ Thuật Phần Mềm cũng như khoa Công Nghệ Thông Tin đã mang lại cho tôi những kiến thức vô cùng quý giá và bổ ích trong quá trình học tập tại trường.

Đặc biệt xin chân thành cảm ơn thầy giáo, TS. Nguyễn Văn Vinh, người đã định hướng, giúp đỡ, trực tiếp hướng dẫn và tận tình chỉ bảo tôi trong suốt quá trình nghiên cứu, xây dựng và hoàn thiện luận văn này.

Tôi cũng xin được cảm ơn tới gia đình, những người thân, các đồng nghiệp và bạn bè thường xuyên quan tâm, động viên, chia sẻ kinh nghiệm, cung cấp các tài liệu hữu ích trong thời gian học tập, nghiên cứu cũng như trong suốt quá trình thực hiện luận văn tốt nghiệp.

Hà Nội, ngày tháng năm 2019

MỤC LỤC

LỜI CẢM ƠN	4
MỤC LỤC	5
DANH MỤC KÝ HIỆU VÀ CÁC CHỮ VIẾT TẮT.....	7
DANH MỤC HÌNH VẼ VÀ ĐỒ THỊ	8
MỞ ĐẦU.....	1
1. Động lực nghiên cứu	2
2. Mục tiêu luận văn	3
3. Cấu trúc luận văn.....	4
CHƯƠNG 1 : TỔNG QUAN HỆ THỐNG CHATBOT	5
1.1 Giới thiệu	5
1.2 Cấu trúc các thành phần hệ thống chatbot	6
1.3 Hiểu ngôn ngữ tự nhiên (NLU).....	7
1.4 Quản lý hội thoại (DM).....	11
1.4.1 Mô hình máy trạng thái hữu hạn FSA.....	13
1.4.2 Mô hình Frame-based	14
1.5 Thành phần sinh ngôn ngữ (NLG).....	14
1.5.1 Template-based NLG	15
1.5.2 Plan-based NLG	15
1.5.3 Class-based NLG	16
CHƯƠNG 2 : MỘT SỐ KỸ THUẬT SỬ DỤNG TRONG CHATBOT	17
2.1 Kiến trúc mạng nơ ron nhân tạo.....	17
2.2 Mạng nơ ron hồi quy RNN	20
2.3 Mạng Long short Term Memory (LSTM)	23
2.3.1 Vấn đề phụ thuộc quá dài.....	23
2.3.2 Kiến trúc mạng LSTM	24
2.3.3 Phân tích mô hình LSTM	25
2.4 Word embeddings	28
2.4.1 Word2vec	28
2.4.2 Glove	29
2.5 Ứng dụng RNN vào quản lý hội thoại	31
2.5.1 Mô hình word-based DST	31

2.5.2	Mô hình Global-Locally Self-Attentive DST (GLAD).....	32
CHƯƠNG 3 : XÂY DỰNG CHATBOT HỖ TRỢ NGƯỜI DÙNG LĨNH VỰC		
NGÂN HÀNG		34
3.1	Bài toán	34
3.2	Xây dựng chatbot hỗ trợ người dùng lĩnh vực ngân hàng	34
3.3	Ứng dụng RASA xây dựng chatbot	36
3.4	Xây dựng dữ liệu chatbot.....	37
3.4.1	Xây dựng ý định.....	38
3.4.2	Xây dựng entity.....	39
3.4.3	Xây dựng câu trả lời cho bot.....	39
3.4.4	Xây dựng khung kịch bản (history)	41
3.5	Thực nghiệm	43
3.6	Đánh giá	51
KẾT LUẬN		52
TÀI LIỆU THAM KHẢO.....		53
PHỤ LỤC		55

DANH MỤC KÝ HIỆU VÀ CÁC CHỮ VIẾT TẮT

Từ viết tắt	Từ chuẩn	Diễn giải
AI	Artificial Intelligence	Trí tuệ nhân tạo
ANN	Artificial Neural Network	Mạng nơ ron nhân tạo
CBOW	Continuous Bag of Words	
CNN	Convolution Neural Network	Mạng nơ ron tích chập
CRF	Conditional Random Fields	Mô hình xác suất trường điều kiện ngẫu nhiên
DM	Dialogue Management	Quản lý hội thoại
DNN	Deep Neural Networks	Mô hình học máy
DTS	Dialogue State Tracking	Theo dõi trạng thái hội thoại
FSA	Finite State Automata	Mô hình dựa trên máy trạng thái hữu hạn
FSM	Finite State Machine	Máy trạng thái hữu hạn
GLAD	Global-Locally SelfAttentive Dialogue State Tracker	
HMM	Hidden Markov Models	Mô hình Markov ẩn
LSTM	Long short-term memory	Mạng cải tiến để giải quyết vấn đề phụ thuộc quá dài
NLG	Natural Language Generation	Thành phần sinh ngôn ngữ
NLP	Natural Language Processing	Xử lý ngôn ngữ tự nhiên
NLU	Natural Language Understanding	Hiểu ngôn ngữ tự nhiên
ML	Machine Learning	Học máy, máy có khả năng học tập
POS	Part Of Speech	Gán nhãn từ loại
RNN	Recurrent Neural Network	Mạng nơ ron hồi quy
SVM	Vector Support Machine	Máy vector hỗ trợ

DANH MỤC HÌNH VẼ VÀ ĐỒ THỊ

Hình 1.1: Cấu trúc các thành phần cơ bản hệ thống chatbot [12]	6
Hình 1.2: Mô hình các thành phần xử lý trong chatbot [1]	7
Hình 1.3: Các bước xử lý chính trong pipeline của NLU [1].....	8
Hình 1.4: Các bước xử lý trong NLU [2]	8
Hình 1.5: Mô hình các bước xác định ý định	10
Hình 1.6: Gán nhãn từ theo mô hình B-I-O trong trích xuất thông tin	Error!
Bookmark not defined.	
Hình 1.7: Mô hình quản lý trạng thái và quyết định action trong hội thoại [2]	12
Hình 1.8: Quản lý hội thoại theo mô hình máy trạng thái hữu hạn FSA	13
Hình 1.9: Frame cho chatbot hỏi thông tin khách hàng	14
Hình 1.10: Phương pháp sinh ngôn ngữ dựa trên tập mẫu câu trả lời [1]	15
Hình 1.11: Phương pháp sinh ngôn ngữ Plan-based [1].....	15
Hình 1.12: Phương pháp sinh ngôn ngữ class-based [1].....	16
Hình 2.1: Kiến trúc mạng nơ ron nhân tạo [15]	17
Hình 2.2: Quá trình xử lý thông tin của một mạng nơ-ron nhân tạo [15]	18
Hình 2.3: Mạng RNN [15].....	21
Hình 2.4: Mạng RNN 2 chiều [15].....	22
Hình 2.5: Mạng RNN nhiều tầng [15].....	22
Hình 2.6: RNN phụ thuộc short-term [17]	23
Hình 2.7: RNN phụ thuộc long-term [17]	23
Hình 2.8: Các mô-đun lặp của mạng RNN chứa một layer [17].....	24
Hình 2.9 : Các mô-đun lặp của mạng LSTM chứa bốn layer [17].....	24
Hình 2.10 : Tế bào trạng thái LSTM giống như một băng truyền [17].....	25
Hình 2.11 : Cổng trạng thái LSTM [17].....	26
Hình 2.12 : LSTM focus f [17].....	26
Hình 2.13 : LSTM focus I [17].....	27
Hình 2.14 : LSTM focus c [17]	27
Hình 2.15 : LSTM focus o [17]	27
Hình 2.16 : Mô hình từ nhúng [16]	28
Hình 2.17 : Mô hình CBOW và Skip-Ngram [16].....	29
Hình 2.18 : Xác xuất từ k trên ngữ cảnh của từ i và j [16].....	29

Hình 2.20 : Công thức tính hàm chi phí tối thiểu [16]	30
Hình 2.21 : Hàm trọng số (weighting function) [16]	30
Hình 2.22: Mô hình word-based DST với mạng RNN [20].....	31
Hình 2.23: Mô hình Global-Locally Self-Attentive DST (GLAD) [21]	32
Hình 2.24 : Global-locally self-attentive encoder modul [21]	33
Hình 3.1: Cấu trúc chatbot cho hệ thống ebanking	35
Hình 3.2 : Cấu hình pipeline xử lý ngôn ngữ tự nhiên	37
Hình 3.3 : Các bước xây dựng chatbot	38
Hình 3.4 : Xây dựng ý định người dùng.....	39
Hình 3.6 : Mẫu câu trả lời của bot cho ý định hỏi số dư tài khoản	39
Hình 3.7: Mẫu câu trả lời mặc định của bot khi không nhận ra ý định người dùng	40
Hình 3.8: Mẫu câu trả lời cho ý định người dùng hỏi lãi suất.....	41
Hình 3.9: Custom action xử lý slot money	41
Hình 3.10: Khung kịch bản chuyển tiền bằng thẻ	42
Hình 3.11: Học tương tác với bot.....	42
Hình 3.12 : Bảng mô tả số lần thử nghiệm với người dùng	43
Hình 3.13: Hình ma trận ước lượng nhằm lần xây dựng dữ liệu intent	43
Hình 3.14: Hình ước lượng độ chính xác tập dữ liệu training intent	44
Hình 3.15 : Hình ước lượng độ chính xác trích chọn thông tin.....	44
Hình 3.16: Hình ước lượng độ nhằm lần việc xây dựng câu trả lời của bot (action)....	45
Hình 3.17 : Bảng mô tả đoạn hội thoại test với bot.....	50

MỞ ĐẦU

“Chatbot will fundamentally revolutionize how computing is experienced by everybody”

CEO Microsoft - Satya Nadella

Chatbot nói riêng hay trợ lý ảo nói chung đang ngày càng thông minh hơn, hoàn thiện hơn. Nó giúp chúng ta có những tương tác và trải nghiệm tốt hơn đối với các phần mềm.

Chúng ta sẽ đi qua một số thành tựu mà chatbot đã đạt được (thống kê trên infographic : the future of chatbots statistics & trends) [23]

- ✓ **100000** : là số lượng chatbot trên Facebook Messenger tại thời điểm tháng 11/2018
- ✓ **80%** các doanh nghiệp muốn sử dụng chatbot vào năm 2020
- ✓ **63%** người dùng sẽ xem xét một tin nhắn online từ chatbot để kết nối với doanh nghiệp hoặc nhãn hiệu
- ✓ **59%** số người đã từng sử dụng chatbot trên một app
- ✓ **37.11%** là số chatbot tăng trưởng trong suốt giai đoạn 2017-2021
- ✓ **56%** các công ty đa truyền thông và công nghệ sẽ chuyển qua sử dụng các công nghệ chăm sóc khách hàng tự động trong tương lai gần, **33%** sẽ có kế hoạch chuyển sang dùng các robot sử dụng AI trước năm 2019
- ✓ **75%** là tỉ lệ thành công của việc bot tương tác với người dùng trong lĩnh vực y tế và ngân hàng. 90% là chỉ số mong đợi sau này.
- ✓ **30%** là số chi phí giảm đi khi áp dụng chatbot vào lĩnh vực chăm sóc khách hàng
- ✓ **8 tỷ USD** là chi phí sẽ giảm khi áp dụng chatbot vào chăm sóc khách hàng đến năm 2022 so với 20 triệu USD ở năm 2017

Vậy đâu là cơ sở để chatbot phát triển nhanh như vậy. Chúng ta sẽ đi qua một số các tác nhân chính:

Đầu tiên là do thời đại bùng nổ của tin nhắn và nó là cơ sở cho chatbot phát triển mạnh mẽ:

- ✓ *28,2 tỷ tin nhắn di động đã được gửi trong năm 2017, gấp đôi so với năm 2012.* [25]
- ✓ *98% tin nhắn sẽ được đọc, với email tỷ lệ là 22%.*
- ✓ *Tỷ lệ gỡ bỏ các ứng dụng nhắn tin chỉ bằng một nửa so với các ứng dụng khác*

- ✓ 6 trong top các ứng dụng được cài đặt là ứng dụng nhắn tin: WhatsApp, Facebook messenger, Wechat, Viber, Line, KakaoTalk, Telegram [24]

Bên cạnh đó cách mạng công nghệ 4.0 đã kéo theo nhiều thay đổi đáng kể đối với nhiều lĩnh vực trong cuộc sống, đặc biệt nó thúc đẩy sự tiến bộ lớn trong lĩnh vực hiểu và xử lý ngôn ngữ tự nhiên (NLU và NLP). Hàng loạt các thuật toán ra đời bắt đầu từ năm 2013 đã giúp chatbot ngày càng thông minh và chính xác hơn.

Và cuối cùng là sự đầu tư nghiên cứu mạnh mẽ vào chatbot của các ông lớn trong lĩnh vực công nghệ đã tạo ra hệ sinh thái, platform trên nền tảng các chatbot giúp người dùng có thể tiếp cận một cách dễ dàng. Đây chính là tiền đề thúc đẩy cho sự phát triển mạnh mẽ của chatbot trong những năm gần đây.

Kỷ nguyên của Chatbot mới chỉ đang bắt đầu nhưng chúng ta đã được hưởng không ít lợi ích từ việc ứng dụng chúng. Với công nghệ ngày một phát triển, đặc biệt là với những tiến bộ của trí tuệ nhân tạo trong vài năm trở lại đây, chúng ta hoàn toàn có thể kỳ vọng ở một tương lai nơi chatbot không chỉ giúp con người giải đáp mọi khía cạnh trong cuộc sống mà còn có thể thay con người đưa ra quyết định.

1. Động lực nghiên cứu

Ở nước ta, việc giải đáp thắc mắc của bộ phận chăm sóc khách hàng qua tin nhắn trực tuyến đang được ưa chuộng. Tuy nhiên, việc này còn thực hiện một cách thủ công và gặp nhiều khó khăn như: tốn rất nhiều thời gian và chi phí chi trả cho nhân viên chỉ để trả lời những câu hỏi đơn giản và giống nhau. Chính vì vậy, nhu cầu cấp thiết là cần một hệ thống điều khiển thông minh, tự động để mang lại hiệu quả cao hơn và Chatbot là một sự lựa chọn hoàn hảo.

Hiện nay, các ứng dụng trò chuyện trực tuyến được mọi người sử dụng đang bắt đầu trở thành một phương tiện ưa thích để giao tiếp với các doanh nghiệp và giải quyết thắc mắc của khách hàng. Ứng dụng nhắn tin nhanh đã trở thành điểm đến hàng đầu cho mọi thương hiệu nhằm tiếp cận người tiêu dùng, bởi vậy không có gì đáng ngạc nhiên khi Chatbot ngày càng trở nên phổ biến.

Với một khối lượng lớn câu hỏi mà chúng ta phải giải quyết mỗi ngày như: khách hàng hỏi về sản phẩm, dịch vụ, nhân viên hỏi về các quy chế công ty, con cái hỏi về những sự việc chúng đang tò mò...ngoài ra chatbot còn được áp dụng trong rất nhiều lĩnh vực:

Giải trí: Các ChatBot giải trí trực tuyến tốt nhất dựa trên AI có thể là Mitsuku, Rose, Insomno Bot...người dùng có thể nói chuyện tương tác với chúng hàng giờ, nó

trả lời câu hỏi của bạn theo cách nhân văn nhất và hiểu được tâm trạng của bạn với ngôn ngữ bạn đang sử dụng.

Thời tiết: Poncho là ChatBot được thiết kế để trở thành một chuyên gia thời tiết, ngoài dự báo thời tiết chúng còn gửi cảnh báo khi thời tiết xấu với sự chấp thuận của người dùng.

Từ thiện: Để nâng cao nhận thức của con người về cuộc khủng hoảng nước ở Ethiopia (dưới 50% dân số được sử dụng nước sạch), tổ chức từ thiện: Nước hợp tác với Lokai để tạo ra Yeshi. Yeshi là một ChatBot đại diện các cô gái trẻ ở Ethiopia, người phải đi bộ 2,5 giờ mỗi ngày để tìm nước sạch. Khi ai đó bắt đầu trò chuyện với bot, Yeshi sẽ gửi hình ảnh, video, clip âm thanh và bản đồ để tạo ra trải nghiệm cảm xúc sâu sắc giúp người dùng khám phá ra thực tế khắc nghiệt của người Ethiopia như Yeshi.

Nhà hàng và các ngành bán lẻ: Khách hàng được Chatbot chào đón và được cung cấp các tùy chọn menu như: chọn vị trí chỗ ngồi, thanh toán và được thông báo khi nào họ có thể bắt đầu lấy thức ăn của họ.

Khách sạn và Du lịch: Chatbot có thể giúp khách sạn trong một số lĩnh vực, bao gồm quản lý thời gian, dịch vụ khách hàng và giảm chi phí nhân lực. Chúng có thể giúp khách hàng với các câu hỏi cơ bản và yêu cầu. Chúng có thể được lập trình để trò chuyện với khách bằng các ngôn ngữ khác nhau, làm cho các khách hàng nói chuyện bằng ngôn ngữ địa phương của mình dễ dàng hơn.

Y tế: ChatBot này sẽ hỏi về các triệu chứng, các thông số cơ thể và lịch sử y tế, sau đó biên soạn một danh sách các nguyên nhân gây ra hầu hết các triệu chứng và xếp hạng chúng theo thứ tự nghiêm trọng. ChatBot có thể hướng dẫn bệnh nhân điều trị các bệnh có thể được chữa khỏi mà không cần đến bác sĩ.

Hàng không: bạn có thể nhận tài liệu chuyển bay của mình qua Messenger, bao gồm xác nhận đặt vé, thông báo đăng ký, thẻ lên máy bay, và cập nhật trạng thái chuyến bay.

2. Mục tiêu luận văn

Nghiên cứu các thành phần cấu tạo chatbot. Tìm hiểu các kỹ thuật xử lý ngôn ngữ trong NLU, NLP như biểu diễn ngôn ngữ, phân loại ý định (intent classification hay intent detection), trích xuất thông tin (information extraction), quản lý hội thoại... trong việc xây dựng chatbot

Luận văn tập trung tìm cách giải quyết các bài toán mà chatbot **ứng dụng trong miền đóng (closed domain)** và trả lời theo **mô hình truy xuất thông tin (retrieval-based)**. 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. 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.

Với đề tài này thì luận văn sẽ tập trung xây dựng hệ thống chatbot hỗ trợ người dùng trong lĩnh vực ngân hàng dựa vào framework **Rasa** và áp dụng những kiến thức tìm hiểu về chatbot để có thể tùy chỉnh trên mã nguồn mở này. Đối tượng hỗ trợ người dùng chatbot cụ thể ở đây là khách hàng cá nhân sử dụng trong hệ thống ebanking.

3. Cấu trúc luận văn

MỞ ĐẦU: Giới thiệu và đưa ra hướng nghiên cứu bài toán chatbot

CHƯƠNG I: Tổng quan hệ thống chatbot: Giới thiệu tổng quan hệ thống chatbot. Cấu tạo và nhiệm vụ các thành phần hệ thống chatbot.

CHƯƠNG II: Một số kỹ thuật sử dụng trong chatbot: Hiểu và nắm được một số kỹ thuật hay thuật toán cơ bản sử dụng trong chatbot để từ đó có thể điều chỉnh phù hợp với ngôn ngữ tiếng việt, giúp chatbot xử lý thông minh hơn.

CHƯƠNG III: Xây dựng chatbot lĩnh vực ngân hàng : Xây dựng ứng dụng chatbot cho ngân hàng trên nền tảng mã nguồn mở **Rasa**. Đưa ra những kết luận và đánh giá, cải tiến hệ thống.

KẾT LUẬN : Đưa ra những kết luận, đánh giá và định hướng nghiên cứu tiếp theo

PHỤ LỤC: Danh sách các đoạn hội thoại với bot được đính kèm ở phần thử nghiệm

CHƯƠNG 1 : TỔNG QUAN HỆ THỐNG CHATBOT

Chương này sẽ giới thiệu **tổng quan** về hệ thống chatbot, **cấu trúc các thành phần** và **những vấn đề** khi xây dựng hệ thống chatbot.

1.1 Giới thiệu

Hệ thống đối thoại người máy hay còn gọi với thuật ngữ là chatbot ChatBot là một chương trình máy tính tiến hành cuộc trò chuyện thông qua nhắn tin nhanh, nó có thể tự động trả lời những câu hỏi hoặc xử lý tình huống. Phạm vi và sự phức tạp của ChatBot được xác định bởi thuật toán của người tạo nên chúng. ChatBot thường được ứng dụng trong nhiều lĩnh vực như thương mại điện tử, dịch vụ khách hàng, y tế, tài chính ngân hàng, các dịch vụ giải trí...

Chatbot có thể được chia thành 2 loại:

- Hệ thống hướng mục tiêu trên một miền ứng dụng (Task-Oriented)
- Hệ thống không có định hướng mục tiêu (chit-chat)

Miền mở (Open Domain): Mô hình trả lời tự động trên miền mở cho phép người dùng có thể tham gia trò chuyện với một chủ đề bất kỳ, không nhất thiết phải có một mục tiêu rõ ràng hay một ý định cụ thể nào. Các cuộc trò chuyện trên mạng xã hội như Facebook, Twitter thường là miền mở, chúng có thể đi vào tất cả các chủ đề. Số lượng các chủ đề thảo luận được đề cập đến là không giới hạn, do đó, tri thức yêu cầu được tạo ra để trả lời các câu đối thoại thuộc miền mở trở nên khó hơn. Tuy nhiên, việc thu thập trích rút dữ liệu từ miền này khá phong phú và đơn giản.

Miền đóng (Close Domain): Mô hình trả lời tự động thuộc miền đóng thường tập trung vào trả lời các câu hỏi đối thoại liên quan đến một miền cụ thể, ví dụ như: Y tế, Giáo dục, Du lịch, Mua sắm, ..

Trong một miền đóng cụ thể, không gian các mẫu hỏi input và output là có giới hạn, bởi vì các hệ thống này đang cố gắng để đạt được một mục tiêu rất cụ thể. Hệ thống hỗ trợ kỹ thuật (Technical Customer Support) hay tư vấn và hỗ trợ mua hàng (Shopping Assistants) là các ứng dụng thuộc miền đóng. Các hệ thống này không thể đối thoại về “Chính trị” hay “Pháp luật”, chúng chỉ cần thực hiện các nhiệm vụ cụ thể một cách hiệu quả nhất có thể. Chắc chắn, người dùng vẫn có thể hỏi đáp bất cứ gì, nhưng hệ thống không yêu cầu phải xử lý những trường hợp ngoại lệ này.

Mỗi cách tiếp cận bài toán đều có hướng giải quyết khác nhau dẫn tới các kỹ thuật sử dụng khác nhau. **Trong luận văn này, tôi sẽ tập trung vào giải quyết trường hợp thứ nhất là bài toán hướng mục tiêu trên một miền ứng dụng đóng.**

1.2 Cấu trúc các thành phần hệ thống chatbot

Hình 1.1: Cấu trúc các thành phần cơ bản hệ thống chatbot [12]

Chatbot có 3 thành phần chính là hiểu ngôn ngữ tự nhiên (**NLU**), quản lý hội thoại (**DM**), thành phần sinh ngôn ngữ (**NLG**). Các thành phần nhận dạng giọng nói Speech Recognition (text to speech hay speech to text) là các thành phần tăng cường.

Mỗi thành phần trong chatbot đều có vai trò riêng:

- **NLU**: bao gồm việc xử lý ngôn ngữ tự nhiên (NLP) có nhiệm vụ xác định được ý định câu hỏi(intent classification) và trích chọn thông tin (slots filter)
- **DM**: Quản lý hội thoại có nhiệm vụ xác định được hành động (action) tiếp theo dựa vào trạng thái hành động trước đó hay ngữ cảnh hội thoại. Các ngữ cảnh này phải được đối chiếu trong các kịch bản dự sẵn (history) đã đào tạo cho bot. Thành phần này cũng đảm nhiệm việc lấy dữ liệu từ hệ thống khác qua các API gọi trong action
- **NLG**: là thành phần sinh ngôn ngữ dựa vào chính sách (policy) và hành động được xác định trong DM thông qua các tập hội thoại. NGL có thể được sinh ra

câu trả lời dựa vào tập mẫu câu trả lời (pre-defined template) đã đào tạo cho bot.

Để rõ chi tiết luồng xử lý tin nhắn từ các thành phần chatbot ta xem mô hình 1.2:

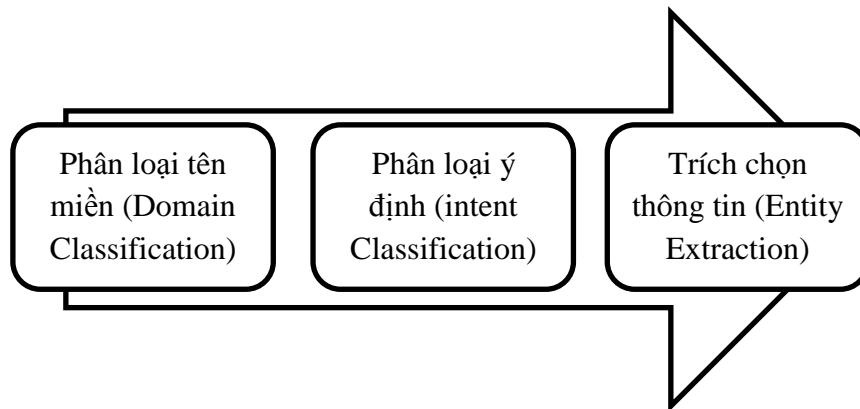
Hình 1.2: Mô hình các thành phần xử lý trong chatbot [1]

1.3 Hiểu ngôn ngữ tự nhiên (NLU)

Đây có thể nói là thành phần quan trọng nhất của chatbot. Chatbot có thông minh hay không thì đây là thành phần quyết định. Mục tiêu của thành phần này là trích xuất ra 3 thành phần thông tin từ câu nói của người dùng:

- Đầu tiên là phân loại lĩnh vực (domain classification), nó có thể là lĩnh vực ngân hàng, y tế hay bảo hiểm. Nếu trong trường hợp chỉ có một lĩnh vực thì không cần thiết cho bước phân loại này.
- Tiếp đến là phân loại ý định (intent classification), ví dụ như xác định được ý định tra cứu thông tin tài khoản hoặc ý định kiểm tra số dư.
- Cuối cùng là bước trích xuất thông tin (slot filler hay entity extraction) trong câu hỏi người dùng. Ví dụ ta phải trích chọn được thông tin số tháng trong câu hỏi người dùng: “lãi suất kỳ hạn 3 tháng là bao nhiêu”. Từ việc trích xuất được thông tin 3 tháng thì chatbot mới có cơ sở trả lời cho người dùng.

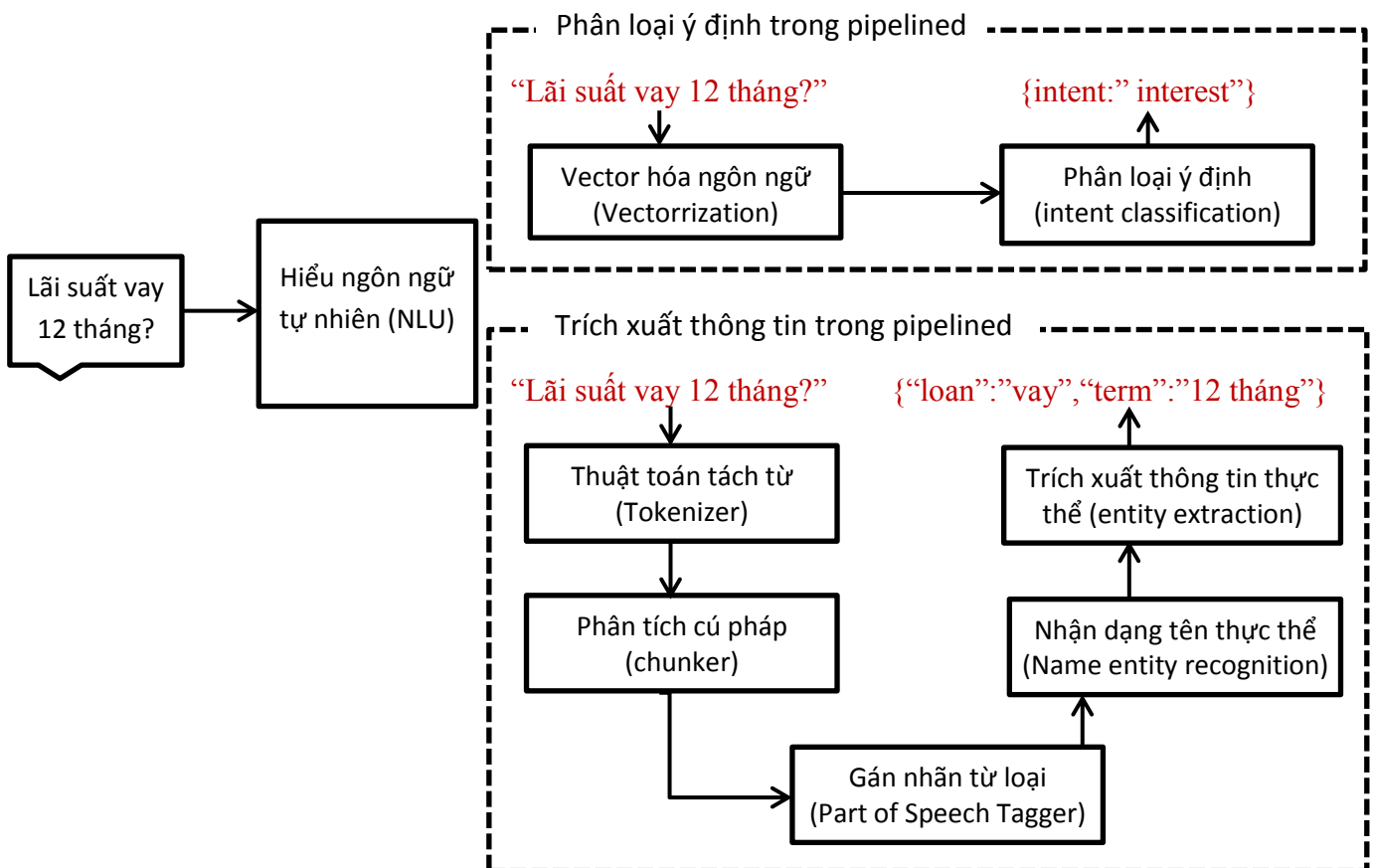
NLU xử lý tin nhắn người dùng bằng một đường ống (pipeline) nơi mà cấu hình các bước xử lý liên tiếp theo tuần tự :



Hình 1.3: Các bước xử lý chính trong pipeline của NLU [1]

Trong đường ống này thì bạn có thể tùy chỉnh các thành phần từ bước tiền xử lý dữ liệu, mô hình hóa ngôn ngữ, các thuật toán dùng để tách từ và trích xuất thông tin thực thể...

Để chi tiết các bước xử lý ta xem trong mô hình 1.4: Trong đó bước entity extraction chính là bước slot filling ở hình 1.3



Hình 1.4: Các bước xử lý trong NLU [2]

Để phân loại được ý định câu người dùng thì ta cần mô hình hóa ngôn ngữ tức là việc biểu diễn ngôn ngữ dưới dạng vector số học cho máy có thể hiểu được (vectorization). Phương pháp phổ biến nhất hiện tại là word embedding (nhúng từ). Tập nhúng từ là tên chung cho một tập hợp các mô hình ngôn ngữ và các phương pháp học đặc trưng trong xử lý ngôn ngữ tự nhiên (NLP), nơi các từ hoặc cụm từ từ vựng được ánh xạ tới vectơ số thực. Về mặt khái niệm, nó liên quan đến việc nhúng toán học từ một không gian với một chiều cho mỗi từ vào một không gian vectơ liên tục với kích thước thấp hơn nhiều. Một số phương pháp biểu diễn phổ biến như Word2Vec, GloVe hay mới hơn là FastText sẽ được giới thiệu trong phần sau.

Sau khi mô hình hóa ngôn ngữ bao gồm dữ liệu đầu vào training cho bot thì việc xác định ý định người dùng từ câu hỏi người dùng dựa trên tập đã training là bước phân loại ý định (intent classification) hay phân loại văn bản. Ở bước này ta có thể dùng một số kỹ thuật như: Naive Bayes, Decision Tree (Random Forest), Vector Support Machine (SVM), Convolution Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM, Bi-LSTM). Hầu hết các chatbot hiện tại đều ứng dụng mô hình deep learning như RNN và LSTM để phân loại ý định người dùng. Bài toán thách thức lớn nhất cho các chatbot ở bước này là xác định nhiều ý định (multiple intents) trong một câu nói người dùng. Ví dụ nếu bạn nói “xin chào, kiểm tra cho tôi số dư tài khoản” thì bot phải xác định được 2 ý định “chào hỏi” và “kiểm tra số dư” trong câu nói người dùng. Nếu bot có thể hiểu và trả lời được câu hỏi loại này sẽ giúp việc tương tác với bot trở nên tự nhiên hơn.

Tiếp đến là việc trích xuất thông tin trong câu hội thoại người dùng. Các thông tin cần trích xuất thường dưới dạng số, chuỗi hoặc thời gian và chúng phải được khai báo và huấn luyện trước.

Phân tách các từ (Tokenization hay word segmentation): Tách từ là một quá trình xử lý nhằm mục đích xác định ranh giới của các từ trong câu văn, cũng có thể hiểu đơn giản rằng tách từ là quá trình xác định các từ đơn, từ ghép... có trong câu. Đối với xử lý ngôn ngữ, để có thể xác định cấu trúc ngữ pháp của câu, xác định từ loại của một từ trong câu, yêu cầu nhất thiết đặt ra là phải xác định được đâu là từ trong câu. Vấn đề này tưởng chừng đơn giản với con người nhưng đối với máy tính, đây là bài toán rất khó giải quyết. Thông thường thì các ngôn ngữ phân tách các từ bởi khoảng trắng nhưng đối với ngôn ngữ tiếng việt thì có rất nhiều từ ghép và cụm từ. Ví dụ từ ghép “tài khoản” được tạo bởi 2 từ đơn “tài” và “khoản”. Có một số thuật toán hỗ trợ giải quyết bài toán này như mô hình so khớp từ dài nhất (longest matching), so khớp cực

đại (Maximum Matching), Markov ẩn (Hidden Markov Models- HMM) hay mô hình CRF (conditional random field)...

1.3.1 Xác định ý định người dùng

Hình 1.5: Mô hình các bước xác định ý định

Hệ thống phân lớp ý định người dùng có một số bước cơ bản:

- Tiền xử lý dữ liệu
- Trích xuất đặc trưng
- Huấn luyện mô hình
- Phân lớp

Bước tiền xử lý dữ liệu chính là thao tác “làm sạch” dữ liệu như: loại bỏ các thông tin dư thừa, chuẩn hoá dữ liệu và chuyển các từ viết sai chính tả thành đúng chính tả, chuẩn hoá các từ viết tắt... Bước tiền xử lý dữ liệu có vai trò quan trọng trong hệ thống chatbot. Nếu dữ liệu đầu vào có xử lý ở bước này thì sẽ làm tăng khả năng năng độ chính xác cũng như sự thông minh cho bot.

Tiếp đến là bước trích xuất đặc trưng (feature extraction hay feature engineering) từ những dữ liệu đã được làm sạch. Trong mô hình học máy truyền thống (trước khi mô hình học sâu được áp dụng rộng rãi), bước trích xuất đặc trưng ảnh hưởng lớn đến độ chính xác của mô hình phân lớp. Để trích xuất được những đặc trưng tốt, chúng ta cần phân tích dữ liệu khá tỉ mỉ và cần cả những tri thức chuyên gia trong từng miền ứng dụng cụ thể.

Bước huấn luyện mô hình nhận đầu vào là các đặc trưng đã được trích xuất và áp dụng các thuật toán học máy để học ra một mô hình phân lớp. Các mô hình phân lớp có thể là các luật phân lớp (nếu sử dụng decision tree) hoặc là các vector trọng số

tương ứng với các đặc trưng được trích xuất (như trong các mô hình logistic regression, SVM, hay mạng Neural).

Sau khi có một mô hình phân lớp intent, chúng ta có thể sử dụng nó để phân lớp một câu hội thoại mới. Câu hội thoại này cũng đi qua các bước tiền xử lý và trích xuất đặc trưng, sau đó mô hình phân lớp sẽ xác định “điểm số” cho từng intent trong tập các intent và đưa ra intent có điểm cao nhất

Để đưa ra hỗ trợ được chính xác, chatbot cần xác định được ý định (intent) đó của người dùng. Việc xác định ý định của người dùng sẽ quyết định hội thoại tiếp theo giữa người và chatbot sẽ diễn ra như thế nào. Vì thế, nếu xác định sai ý định người dùng, chatbot sẽ đưa ra những phản hồi không đúng, không hợp ngữ cảnh. Khi đó, người dùng có thể thấy chán ghét và không quay lại sử dụng hệ thống. Bài toán xác định ý định người dùng vì thế đóng vai trò rất quan trọng trong hệ thống chatbot.

Đối với miền ứng dụng đóng, chúng ta có thể giới hạn số lượng ý định của người dùng nằm trong một tập hữu hạn những ý định đã được định nghĩa sẵn, có liên quan đến những nghiệp vụ mà chatbot có thể hỗ trợ. Với giới hạn này, bài toán xác định ý định người dùng có thể quy về bài toán phân lớp văn bản. Với đầu vào là một câu giao tiếp của người dùng, hệ thống phân lớp sẽ xác định ý định tương ứng với câu đó trong tập các intent đã được định nghĩa trước.

Để xây dựng một mô hình phân lớp intent, chúng ta cần một tập dữ liệu huấn luyện bao gồm các cách diễn đạt khác nhau cho mỗi intent. Ví dụ, cùng một mục đích hỏi về số dư tài khoản người dùng có thể dùng những cách diễn đạt sau:

- *Thông tin tài khoản?*
- *Tra cứu tài khoản?*
- *Số dư tài khoản?*
- *Số tiền trong tài khoản?*

Có thể nói, bước tạo dữ liệu huấn luyện cho bài toán phân lớp intent là một trong những công việc quan trọng nhất khi phát triển hệ thống chatbot và ảnh hưởng lớn tới chất lượng sản phẩm của hệ thống chatbot về sau. Công việc này cũng đòi hỏi thời gian, công sức khá lớn của nhà phát triển chatbot.

1.4 Quản lý hội thoại (DM)

Trong các phiên trao đổi dài (long conversation) giữa người và chatbot, chatbot sẽ cần ghi nhớ những thông tin về ngữ cảnh (context) hay quản lý các trạng thái hội thoại