國立中正大學

資訊工程研究所

碩士論文

A Task-oriented Chatbot Based on LSTM and Reinforcement Learning

周台亮

106級

國 立 中 正 大 學

資訊工程研究所碩士論文

基於長短期記憶遞迴模型和強化式學習的任務型聊天機器人

A Task-oriented Chatbot Based on LSTM and Reinforcement Learning

研究生：周台亮

指導教授：薛幼苓 博士

中華民國 一百零八 年 六 月

國 立 中 正 大 學

資訊工程研究所碩士論文

基於長短期記憶遞迴模型和強化式學習的任務型聊天機器人

A Task-oriented Chatbot Based on LSTM and Reinforcement Learning

研究生：周台亮

指導教授：薛幼苓 博士

中華民國 一百零八 年 六 月

請插入**學位考試同意書**到這頁。

請插入**學位論文考試審定書**到這頁。

請插入**博碩士論文授權書**到這頁。

致謝

在實驗室也過了兩年的日子，在這七百多天的日子最感謝的就是指導教授薛幼苓老師，在老師的指導下順利地完成了研究，也學到了很多技能，在老師的熱忱指導之下，獲得了寶貴的經驗 ，也能夠跟實驗室的好夥伴們一同學習、成長，覺得非常幸運能夠進到資料管理實驗室。

再來要感謝的是我的家人，在我長年的求學期間作為我的後盾，使我能夠專心於學習以及研究上，給予我很大的支持與包容，感謝家人給我的無私地付出，讓我能夠打起精神繼續下一個挑戰。

接著要感謝實驗室一起度過兩年研究的夥伴(冠佑、毓仁、敬玲)，給我扶持給我幫助，以及學長(家駿、坤緯、修宇、文仁)、學弟妹們(台亮、浩澤、紹薇、佩蓉)，一同經歷難關，與這些夥 伴們一同學習一同玩樂，減輕了學業上的壓力，度過了充實又愉快的研究所生活，謝謝你們帶給我的快樂的回憶。

最後，還有許許多多，在我研究生活中給予我幫助的人們，由於有著大家的幫忙，我才得以順利完成學業，我由衷地感謝大家給予的幫忙。此後，我也會善加使用所學，繼續向前走。

摘要

最近有個趨勢吸引大多數科技行業注意，那就是聊天機器人。基於人工智能(AI)，聊天機器人可以學習如何應對不同的情況，並與人類用自然語言交流。此外，多虧有各種雲端平台讓聊天機器人可以很容易地部署，聊天機器人應用程式可以在訊息、手機、網頁或手機等上看到。

毫無疑問，聊天機器人是人類與機器互動的一種新方式。然而，一個典型的聊天機器人也代表了一個簡單的問題回答系統，公式化的回答。傳統的對話聊天機器人通常采用基於檢索的模型。開發人員必須提供大量的對話數據，並根據不同的任務對這些數據進行分類。為了避免繁瑣的開發過程，我們提出了一種基於生成對抗性網絡生成句子的生成模型來構建聊天機器人。我們模型的體系結構包含一個生成器，該生成器依照不同的編碼生成不同的句子、一個對話管理碼和一個人物碼，還有一個判別器，用於判斷生成的句子和原始數據。在生成器中，我們結合注意模型追蹤句子狀態與使用雙向長短期記憶序列模型來提取句子信息的 對於判別器，我們計算了三種類型的回饋分數，分別為重覆句的低獎勵和不同句的高獎勵。通過大量的實驗驗證了該模型的有效性，與現有的方法相比，該模型生成的句子更加多樣化，信息量也更加豐富。

關鍵字：Deep Learning, Deep Learning, Chatbots, Service Robots, Natural Language Processing, Sentence Generation, Dialog Management.

x

Abstract

The latest trend that is catching attention of the majority of the tech industry is chatbots. Based on artificial intelligence (AI), a chatbot can learn how to react to different situations and communicate with the human in natural language. In addition, thanks for the various cloud platforms, chatbot can easy to deploy. As a result, chatbot applications can be seen tremendously on messaging applications on websites and mobile phones. Undoubtedly, a chatbot is a new way of interaction between humans and machines. However, a typical chatbot can also act as a simple question answering system that responses with formulated answers. Traditional conversational chatbots usually adopt a retrieved-based model. Developers have to provide a large amount of conversational data and classify those data to different intents. To avoid cumbersome development processes, we propose a method to build a chatbot by a sentence generation model which generates sequence sentences based on the generative adversarial network. The architecture of our model contains a generator that generates a diverse sentence with a corresponding action code, a dialogue manage code, and a persona code, and a discriminator that judges the sentences between the generated and the raw data. In the generator, we combine the attention model that responses for tracking conversational states with the sequence-to-sequence model using hierarchical long-short term memory to extract sentence information. For the discriminator, we calculate three types of rewards to assign low rewards for repeated sentences and high rewards for diverse sentences. Extensive experiments are presented to demonstrate the utility of our model which generates more diverse and

information-rich sentences than those of the existing approaches.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The types of input of digital devices have become more natural and convenient, and mainly have evolved toward speech in recent years. Talking to personal assistants such as speech interpretation and recognition interface (SIRI) [1] to enable simple tasks accomplished is widely accepted nowadays. Since the popularization of conversational interface applications has been a major trend recently, there is a need for a new application with integrated conversational interfaces. In particular, the best conversational interface with voice or text inputs is a chatbot that adopts language processing to handle inputs. Compared to traditional applications, a chatbot can provide a more intuitive user interface to users. Based on natural language processing, a chatbot can interact with humans in natural language, and use the voice recognition to control the device. Unlike humans, chatbots are free from the restriction that serves only a limited number of customers at the same time. Obviously, chatbots can provide several services and receive multiple queries any time at once. Therefore, a chatbot decreases the personnel costs to serve a greater number of customers. Furthermore, a smart chatbot not only learns user's behaver patterns and made decisions, but also needs to answer correctly to users depending on the conventional context.

More recently, a chatbot can be easily developed through the existing frameworks such as API.ai [2], Wit.ai [3] and Luis.ai [4]. However, there still exist a lot of challenges in building complex multiple-steps conversations in the processing of developing chatbots. First, a chatbot does not cover such a wide spectrum of capabilities and does not even represent a bot that can chat with people. Most of traditional chatbots do not really have natural language understanding and capability. Actually, these chatbots can only deal with simple keyword matching to inquire the corresponding sentences in the database. Second, the existing frameworks do not provide a sentence generation system for developers. Moreover, it is difficult to define the dialog states and maintain historical conversations.

When using those applications, we have to design dialogue management and responses.In particular, we have to create multiple intent components, labeling synonym entities, and tracking states of a chatbot. Most importantly, when designing the dialog management system for chatbots, the understanding of the users' queries and proper responds need to be carefully handled, because the traditional chatbots are not able to memorize a conversational status that already generated.

Recent research in the generative adversarial [5] [6] network model has inspired several efforts on conversational systems that apply an encoder to convert an ask sentence to a discrete vector representing its meaning and then generate a response from a de-

coder. The model generates more diverse, interactive, and non-repetitive responses than the traditional sequence-to-sequence models trained without using generative adversarial networks. A good dialogue model should generate utterances indistinguishable from human dialogues.

In this paper, we build a chatbot system for generating sentences. For applying the neural network to our system, each word from the user's inputs has to be transformed into word embeddings which are numerical representations. Since we adopt Chinese corpus as our dataset, where each Chinese sentence is represented as strings of Chinese characters without explicit delimiters, the procedures of word segmentation and word embedding are required for our language model construction. Subsequently, the vector of word representation is delivered to the back-end neural network model to generate respond sentences.

Our proposed model is based on a sequence natural deep learning model and reinforcement learning model. For the sentence generation, the architecture of the model consists of an encoder and a decoder. For the encoder, we use the last state from forward and backward bi-LSTM [7] to be the history state which is used to be the initial state for the decoder. The output of each encoder neuron is used in the attention model to extract the sentence information from the dialogues. We divide the decoder into two kind of condition, the training step and the prediction step. In the training step, the input at each time step of training is the result of the target which is the corresponding answers from the raw dataset. At the step of prediction, the input is the output at the previous step. We build a generative model from the above architecture. For the discriminator, we use a simple RNN as the basis to get a reward for each word and sentence.

Moreover, we utilize the reinforcement learning technique to train our model. Reinforcement learning is a general framework for decision making and can facilitate a response generator to create a sentence that is diverse and understandable. Furthermore, it generates logical responses by maximizing the total future rewards. The higher the reward score, the more fluent the sentence is. To address the discrete dialogue data problem, we apply policy gradient reinforcement learning to back-propagate for updating the parameters used in the model. Therefore, we can achieve improvement in dialogue generation tasks. In the end, the utterances that our model generates could be indistinguishable from human-generated sentences.

The rest of the paper is organized as follows. In Section **??**, we review the important related methods and the state-of-the-art techniques of chatbots. Section **??** provides the background knowledge. We survey different types of neural network models, attention methods, and generative adversarial networks. In the end, we address the reason why we adopt those methods. In Section **??**, we describe the architecture of our chatbot model. For the purpose of generating diverse and understandable sentences, we apply three types of rewards for reinforcement learning. The experiments and evaluation are provided in Section **??**. In Section **??**, we summarize the main findings and conclude the paper. In Section 7, we outline the future work for improving our model.

Bibliography

[1] E. Sadun and S. Sande, *Talking to Siri: Learning the Language of Apple's Intelligent Assistant*. Que Publishing Company, 2nd ed., 2013.

[2] "Build natural and rich conversational experiences." `https://dialogflow.com/`.

[3] "Natural language for developers." `https://wit.ai/`.

[4] "Language understanding (luis)." `https://www.luis.ai/`.

[5] J. Li, W. Monroe, A. Ritter, M. Galley, J. Gao, and D. Jurafsky, "Deep reinforcement learning for dialogue generation," *CoRR*, vol. abs/1606.01541, 2016.

[6] L. Yu, W. Zhang, J. Wang, and Y. Yu, "Seqgan: Sequence generative adversarial nets with policy gradient," *CoRR*, vol. abs/1609.05473, 2016.

[7] M. Schuster and K. Paliwal, "Bidirectional recurrent neural networks," *Trans. Sig. Proc.*, vol. 45, pp. 2673–2681, Nov. 1997.