

Paddle 기반의 중국어 Multi-domain Task-oriented 대화 시스템

등우진, 조인휘*
한양대학교, *한양대학교

dengyuchen@hanyang.ac.kr, iwjoe@hanyang.ac.kr

Chinese Multi-domain Task-oriented Dialogue System based on Paddle

Yuchen Deng, Inwhee Joe*
Hanyang Univ., *Hanyang Univ.

Abstract

With the rise of the AI wave, task-oriented dialogue systems have become one of the popular research directions in academia and industry. Currently, task-oriented dialogue systems mainly adopt pipelined form, which mainly includes natural language understanding, dialogue state decision making, dialogue state tracking and natural language generation. However, pipelining is prone to error propagation, so many task-oriented dialogue systems in the market are only for single-round dialogues. Usually single-domain dialogues have relatively accurate semantic understanding, while they tend to perform poorly on multi-domain, multi-round dialogue datasets. To solve these issues, we developed a paddle-based multi-domain task-oriented Chinese dialogue system. It is based on NEZHA-base pre-training model and CrossWOZ dataset, and uses intention recognition module, dichotomous slot recognition module and NER recognition module to do DST and generate replies based on rules. Experiments show that the dialogue system not only makes good use of the context, but also effectively addresses long-term dependencies. In our approach, the DST of dialogue tracking state is improved, and our DST can identify multiple slotted key-value pairs involved in the discourse, which eliminates the need for manual tagging and thus greatly saves manpower.

I. Introduction

Currently, task-based dialogue systems basically use the following pipeline forms. However, pipelines are prone to error propagation, so many task-oriented dialogue systems on the market only have relatively accurate semantic understanding for single-round, single-domain dialogues, and often perform poorly on multi-field, multi-round dialogue datasets. Second, most of the current mainstream DST models take the dialogue information of all previous rounds as input, which will cause the input sequence of the model to be particularly long and consume large training resources. Moreover, the long text sequences are not conducive to the model's extraction of important information and the training of the model. Furthermore, the DST model will make predictions based on each SLOT of each round during the update process, for example, if there are 100 state information, then 100 model inferences are required after each round of dialogue update, which will greatly consume time and arithmetic cost. In order to solve the above problems, we developed a paddle-based Chinese multi-domain task-oriented dialogue system. Combine the natural language understanding module and the dialogue state tracking module into one module to reduce errors. Combine the natural language understanding module and the dialogue state tracking module into one module to reduce errors. Then use NER for span slot extraction, which can not only reduce the original inference time $O(n)$ to $O(1)$, but also handle the case of out of value. Coupled with classification slot predictions, all dialog states can be predicted.

Paddle is China's first self-developed deep learning platform. It has been officially open sourced to professional communities since 2016. It is an industrial

platform with advanced technology and rich functions, covering deep learning core framework, basic model library, end-to-end-terminal development kit, tools and components, and service platforms.

II. Experiment

A. Dataset selection

CrossWOZ is the first large-scale Chinese cross-domain "human-computer interaction" task-oriented dataset. CrossWOZ contains 6K dialogues, 102K sentences, covering 5 domains (attractions, hotels, restaurants, subways, rentals). In addition, the corpus contains rich annotations of dialog states, as well as dialog behaviors on the user and system sides. About 60% of conversations have cross-domain user goals that favor inter-domain dependencies and facilitate natural transitions across domains in conversations. The large scale and rich annotations of CrossWOZ make it suitable for studying various tasks in cross-domain dialogue modeling, such as dialogue state tracking, policy learning, user simulation, etc. [2]

B. Attribute selection

Table I. Data used in CrossWOZ and Attribute Details

| | |
|--------------------|---|
| attraction_db.json | Attraction name, domain, address, subway, telephone, tickets, time required to play, score, nearby attractions, nearby restaurants. |
| hotel_db.json | Hotel name, domain, hotel type, address, subway, telephone, hotel facilities, |

| | |
|--------------------|--|
| | price, rating, nearby attractions. |
| metro_db.json | Attraction name, domain, subway station name. |
| restaurant_db.json | Domain, restaurant name, address, subway, telephone, opening time, recommended dishes, consumption per capita, rating, nearby attractions, nearby restaurants. |
| taxi_db.json | Domain, departure, destination |

C. Modules and Frameworks

The entire task-oriented dialogue system is composed of four modules: dialogue state tracking module, dialogue state decision module, database query module, and dialogue reply module. The dialogue state tracking module consists of three sub-modules: intent recognition module, two-category Slot recognition module, and Span Slot extraction module. Figures 1 and 2 show the architecture of the intent recognition and span slot extraction modules.

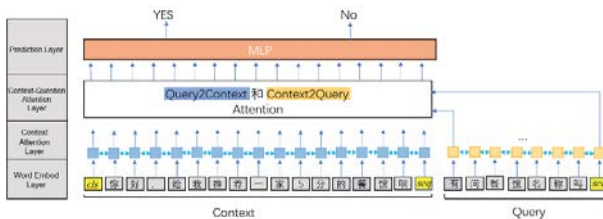


Fig.1. Intent Recognition Module

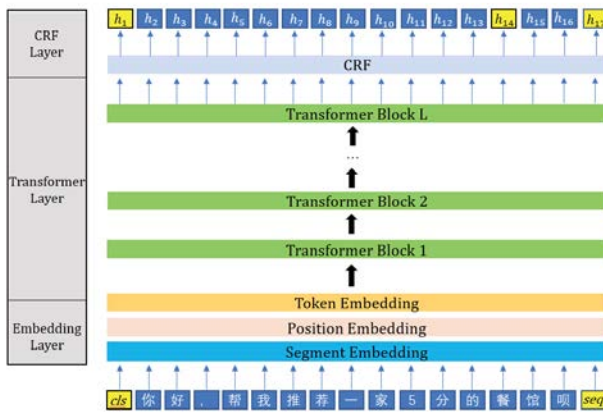


Fig.2. Span Slot Extraction Module

D. Experiment Results

Table II. Dialogue Results

| User/ System | Dialogue |
|-----------------|----------|
|-----------------|----------|

| | |
|---------|--|
| User: | 你好，我想吃美食街，帮我推荐一个人均消费在50-100元的餐馆，谢谢。 Hello, I want to eat in the food court, help me recommend a restaurant with a per capita spending of 50-100 yuan, thank you. |
| System: | 为您推荐鲜鱼口老字号美食街，人均消费75元，有您想吃的美食街哦。 We recommend the time-honored food street in Xianyukou for you. The per capita consumption is 75 yuan. There is a food street you want to eat. |
| User: | 营业时间是什么时候？ What are the opening time? |
| System: | 周一至周日 10:00-22:00。 Monday to Sunday 10:00-22:00. |
| User: | 他家周边有什么景点吗？ Are there any attractions around? |
| System: | 有故宫，前门大街，恭王府，天安门广场。 There are the Forbidden City, Qianmen Street, Gong Wang Fu, Tiananmen Square. |
| User: | 哦，我想在这些附近景点里找一个4.5分以上的，有吗？ Oh, and I'm looking for a 4.5 or higher in these nearby spots, any? |
| System: | 故宫就是哦，4.7分。 The Forbidden City is, rating is 4.7 out of 5. |
| User: | 好的，电话和地址告诉我一下。 Okay, let me know the phone number and address. |
| System: | 010-85007938；北京市东城区景山前街4号。 010-85007938; No. 4, Jingshan Qianjie, Dongcheng District, Beijing, China |
| User: | 好的，收到，谢谢你！ Ok, thank you! |
| System: | 不客气。 You're welcome. |

III. Conclusion

In this paper, we implemented a Chinese multi-domain task-oriented dialogue system based on flying paddles, combines the natural language understanding module and the dialogue state tracking module into one module to reduce the propagation error. The prediction of Slot is divided into Classification Slot and Span Slot in the Dialogue State Tracking module, which not only reduces the prediction time of most Slots from O(n) to O(1), which greatly reduces the time and arithmetic power consumption of Dialogue State Tracking (DST), but also enhances the robustness of the whole model by handling the Out of Values (OOV) case.

References

[1] Wei, Zhongyu, et al. "Task-oriented dialogue system for automatic diagnosis." Proceedings of the 56th Annual

Meeting of the Association for Computational Linguistics
(Volume 2: Short Papers) . 2018.

- [2] Wu, Chien-Sheng, et al. "Transferable multi-domain state generator for task-oriented dialogue systems." arXiv preprint arXiv:1905.08743 (2019).
- [3] CrossWOZ: A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset <https://github.com/thu-coai/CrossWOZ>