Deep Chit-Chat: Deep Learning for Chatbots

Wei Wu
Microsoft Corporation
Beijing 100080, China
wuwei@microsoft.com

Rui Yan
Institute of Computer Science and Technology
Peking University
Beijing 100871, China
ruiyan@pku.edu.cn

ABSTRACT

The tutorial is based on our long-term research on open domain conversation, rich hands-on experience on development of Microsoft XiaoIce, and our previous tutorials on EMNLP 2018 and the Web Conference 2019. It starts from a summary of recent achievement made by both academia and industry on chatbots, and then performs a thorough and systematic introduction to state-of-the-art methods for open domain conversation modeling including both retrieval-based methods and generation-based methods. In addition to these, the tutorial also covers some new progress on both groups of methods, such as transition from model design to model learning, transition from knowledge agnostic conversation to knowledge aware conversation, and transition from single-modal conversation to multi-modal conversation. The tutorial is ended by some promising future directions such as how to combine non-task-oriented dialogue systems with task-oriented dialogue systems and how to enhance language learning with chatbots.

CCS CONCEPTS

• **Computing methodologies** → **Discourse, dialogue and pragmatics**.

KEYWORDS

Human-machine conversation; deep learning; chatbot.

Starting from ELIZA in 1960s, non-goal-oriented conversational agents (a.k.a., chatbots) have never been so popular as in recent years. Take Microsoft XiaoIce as an example. The chatbot was first released to Chinese users in 2014 by Microsoft, and then was launched in Japan (Rinna), U.S. (Zo), India (Ruuh), and Indonesia (Rinna). The products now have attracted more than 800 million all over the world, and the technologies behind have powered a great number of business applications such as various third party official accounts and IoT devices. Different from task-oriented dialogue systems, which are designed for helping people complete specific

tasks such as ticket booking and restaurant reservation etc., chatbots aim to engage users in human-machine conversation in open domain for entertainment and/or emotional companionship. As a result, it is easier for chatbots to go viral among end users without specific purposes than task-oriented dialogue systems. For example, until June of 2018, users from the five countries have finished more than 30 billion conversations with XiaoIce; and on average, each conversation lasts up to 23 turns. The promising user data indicate impressive popularity of the chatbot product.

Goal of the tutorial. In contrast to the prosperity of chatbots among end users, there are few systematic introductions to approaches about how to build the conversational engines behind chatbots in research communities. The reason might be that the conventional chatbots such as ELIZA are designed with hand-crafted rules which are difficult to achieve real open domain human-machine interactions. It is only until the recent two or three years that researchers (who benefit from the large scale social conversation data available on the public Internet and the rapid advancement of deep learning approaches) begin to develop principled and data-driven approaches to build open domain conversational systems. Therefore, we believe it could be useful and valuable to give a tutorial on recent progress of chatbots, particularly on applications of deep learning approaches to building open domain dialogue engines. The community would learn the insights behind chatbots in order to fulfill the gap between task-oriented dialogue systems and non-task-oriented ones. In summary, we expect that our audience can learn the following items from the tutorial: 1) new opportunities modern chatbots, such as Microsoft XiaoIce, bring to academia, industry, and the society; 2) how to build a state-of-the-art dialogue engine with deep learning methods for chatbots; and 3) what are likely to be the future solutions for making chatbots better.

Topic and relevance. Conversational AI is catching on fire: academic conferences especially add new research tracks for conversational studies and attract unexpected growth in the number of submissions to these tracks; companies from industry are making great efforts to develop conversational products. We are entering the AI era whereby large-scale big data become more easily available and learning techniques become more powerful. We may stand at the entrance of future success in more advanced conversational systems (social chatbots and/or virtual assistants). Although this research area still faces bottlenecks and obstacles, we have witnessed a rapid surge of conversational studies in the past few years, especially the chit-chat research in open domain.

The research topic "deep learning for chit-chat" is quite relevant to the information retrieval community in following ways: 1) one approach to implementing a chat system is through information retrieval technologies, such as semantic matching, short text analysis, learning to rank, etc. As a result, half of the tutorial is about

retrieval-based methods for building a chatbot; (2) generation-based methods, which represent the other approach to implementing a chatbot, can be enhanced by retrieval-based methods. In fact, combining the two approaches in one framework is emerging as a new research topic in open domain dialogue modeling; and (3) retrieval-based methods play an important role in real products such as Microsoft XiaoIce and Amazon Alexa. Although there is a tutorial entitled "Neural Approaches to Conversational AI" in SIGIR 2018, the unique value of our tutorial lies in a comprehensive introduction to open domain conversation modeling, especially retrieval-based methods that are untouched by the previous tutorial.

1 TUTORIAL OUTLINE

1.1 Introduction to chatbots

- Insights from practice and products such as XiaoIce.
- History review.
- Recent advances.
- Open domain dialogue systems.

1.2 Retrieval-based chatbots

- Single-turn response selection: message-response matching.
    Frameworks with existing models as special cases.
    Insights from the comparison on benchmark data.
- Multi-turn response selection: context-aware matching
    Frameworks with existing models as special cases.
    Introduction to benchmark datasets.
    Insights from the comparison on the benchmarks.
- Knowledge enhanced response selection
    Single-turn cases.
    Multi-turn cases.
- New trends in research of response selection.
    Matching with better representations.
    Matching with deep interaction.
    Learning a matching model from noisy data.

1.3 Coffee Break

1.4 Generation-based chatbots

- Single-turn chatbots: sequence-to-sequence generation
- Multi-turn chatbots: context-aware modeling
    Non-hierarchical context modeling.
    Various structures of hierarchical context modeling.
- Generation with additional information
    Human-like chit-chat with *persona* and *emotion*.
    Augmenting chit-chat with *topic* and *knowledge*.

1.5 Emerging trends

- Towards better evaluation for conversations
- Model optimization: reinforcement learning and adversarial learning in conversations
- Towards controllable conversations: hard constraints and soft constraints
- Towards open domain dialogue management
- Multi-modal conversation

1.6 Summary and future directions

- Combination with task-oriented dialogues
- Language learning and cognitive learning