Knowledge-enhanced Response Generation in Dialogue Systems: Current Advancements and Emerging Horizons

Priyanshu Priya\textsuperscript{1}, Deeksha Varshney\textsuperscript{1}, Mauajama Firdaus\textsuperscript{2}, Asif Ekbal\textsuperscript{1} \vspace{1mm}
\textsuperscript{1}Indian Institute of Technology Patna, India \hspace{1mm} \textsuperscript{2}University of Alberta, Edmonton, Canada

\texttt{(priyanshu\_2021cs26@iitp.ac.in, 1821cs13@iitp.ac.in, mauajama.03@gmail.com, asif@iitp.ac.in)}

Abstract

This tutorial provides an in-depth exploration of Knowledge-enhanced Dialogue Systems (KEDS), diving into their foundational aspects, methodologies, advantages, and practical applications. Topics include the distinction between internal and external knowledge integration, diverse methodologies employed in grounding dialogues, and innovative approaches to leveraging knowledge graphs for enhanced conversation quality. Furthermore, the tutorial touches upon the rise of biomedical text mining, the advent of domain-specific language models, and the challenges and strategies specific to medical dialogue generation. The primary objective is to give attendees a comprehensive understanding of KEDS. By delineating the nuances of these systems, the tutorial aims to elucidate their significance, highlight advancements made using deep learning, and pinpoint the current challenges. Special emphasis is placed on showcasing how KEDS can be fine-tuned for domain-specific requirements, with a spotlight on the healthcare sector. The tutorial is crafted for both beginners and intermediate researchers in the dialogue systems domain, with a focus on those keen on advancing research in KEDS. It will also be valuable for practitioners in sectors like healthcare, seeking to integrate advanced dialogue systems.

1. Introduction

In the realm of artificial intelligence, dialogue systems have evolved as crucial interfaces facilitating human-machine interaction through natural language conversations. These systems are broadly categorized into task-oriented and open-domain dialogue systems. While task-oriented systems are designed to assist users in specific tasks like restaurant booking (Firdaus et al., 2020d, 2021c; Varshney and Singh, 2021), open-domain systems engage in a broader spectrum of conversational topics without a defined objective (Firdaus et al., 2020a; Varshney et al., 2020). The integration of deep learning, particularly neural language models, has significantly elevated the performance of these systems, yet challenges like understanding user opinions, integrating visual data, and ambiguity in open-domain interactions persist (Chen et al., 2017). In addressing the limitation of generating bland or generic responses common in traditional dialogue systems, Knowledge Enhanced Dialogue Systems (KEDS) have emerged as a prominent solution. The crux of KEDS lies in grounding the dialogues in external or internal knowledge, thereby enriching the conversation with insightful and contextually relevant responses. This tutorial provides an in-depth examination of KEDS, shedding light on its integral components, various approaches, and the benefits derived from such systems.

In this tutorial, we first introduce the foundational frameworks of Knowledge-enhanced Dialogue Systems (KEDS), establishing a solid understanding of how they augment dialogue systems. Following this, we explore the diverse methodologies employed to incorporate both internal and external knowledge sources, thereby enriching the conversational experience. We delve into internal knowledge sources embedded in the input text, such as topics, keywords, and internal graph structures, as discussed in (Ahmad et al., 2023; Mishra et al., 2022b; Firdaus et al., 2021a; Xie and Pu, 2021; Priya et al., 2023a). Concurrently, we investigate external knowledge acquisition from resources like uni-and-multi-modal knowledge bases, knowledge graphs, and grounded text such as persona information, Wikipedia information as elucidated in (Dinan et al., 2018; Zhou et al., 2018b; Firdaus et al., 2020a; Varshney et al., 2021; Ghazvininejad et al., 2018; Varshney et al., 2022a). The discourse further extends to domain-specific applications, particularly in the healthcare sector. In the healthcare domain, having a thorough understanding of a person’s medical history, mental state, symptoms, and treatment plan is crucial. Studies have indicated that the integration of extensive knowledge resources into healthcare dialogue systems presents multiple significant benefits. These include improving the system’s understanding of medical terminology and concepts, equipping the system with the ability to reason and make inferences, grasping the emotional nuances within conversations, and discerning beneficial response patterns that contribute to emotional alleviation (Varshney et al., 2023b; Liang et al., 2021). Motivated by these insights, this tutorial session aims to explore various research endeavors that incorporate external knowledge into healthcare dialogue systems, thereby facilitating personalized and effective support (Shen et al., 2022; Deng et al., 2023; Varshney et al., 2023a).
In the conclusion section, we highlight the shortcomings of conventional dialogue systems to provide a clearer pathway for newcomers to further research in KEDS systems.

2. Target Audience

We believe that the potential target audience could be the students at all levels (Doctorals, Masters, Bachelors), and anyone who is associated with healthcare, customer care, & related application areas, and researchers. We would assume an acquaintance with basic concepts about chatbots and neural networks, such as those included in most introductory Machine Learning (ML), Deep Learning (DL) and Natural Language Processing (NLP) courses. We expect an audience size of about 25-30 participants.

3. Outline

This tutorial is organized as follows:

- **Introduction (15 minutes)** We will briefly introduce dialogue systems, including the different types of dialogue systems and limitations of traditional dialogue systems (Chen et al., 2017). Afterward, we will discuss the notion of knowledge-enhanced response generation in dialogue systems and the different categories of knowledge sources, viz. internal knowledge and external knowledge. Precisely, we will delve into the concepts of (i) Internal knowledge sources embedded in the input text, including but not limited to topic, keyword, and internal graph structure (Xing et al., 2017; Xu et al., 2020; Li and Sun, 2018; Chen and Yang, 2023), and (ii) External knowledge acquisition, including but not limited to the multimodal information, persona, knowledge base, external knowledge graph, and grounded text (Firdaus et al., 2020b, 2022d; Dinan et al., 2018; Zhou et al., 2018b; Ghazvininejad et al., 2018).

- **Need and Challenges of Knowledge-enhanced Response Generation in Dialogues (15 minutes)**

An effective dialogue system should be able to generate coherent, contextually relevant, user-centric, and informative responses. To achieve this, these systems require diverse information sources, including textual and structured data from external sources, user attributes (like sentiment, emotions, politeness, personal profile information - age, gender, persona, etc.), and contextual information (Wang et al., 2023a).

Integrating the knowledge into the generated responses poses challenges concerning the retrieval or selection of pertinent knowledge and effective comprehension and utilization of the acquired knowledge to facilitate response generation (Wang et al., 2023b).

In this section, we will discuss how the varied knowledge resources enhance response generation and improve the interpretability of dialogue systems by incorporating explicit semantics. Subsequently, we will address the challenges inherent in knowledge-enhanced response generation within dialogue systems.

- **Internal Knowledge-enhanced Response Generation in Dialogue Systems (60 minutes)**

In this part of the tutorial, we aim to delineate the internal knowledge-enhanced response generation methods and applications. The information from internal knowledge sources helps enlighten and drive the generated responses to be informative and avoids generating universally relevant replies with little semantics. The internal knowledge can be obtained from topical information, keywords, and internal graph structures. We will point out the works that incorporate these knowledge sources for response generation.

(i) **Response enhanced by Topic**: A dialogue system frequently employing responses such as “I don’t know”, “Okay” “I see” may appear repetitive and uninformative. While these off-topic replies are generally harmless for addressing various inquiries, they lack engagement and are likely to prematurely conclude conversations, significantly diminishing the overall user experience (Xing et al., 2017; Ahmad et al., 2023). Consequently, there is a pressing demand for on-topic response generation. This part of the tutorial delves into the works that have incorporated topical knowledge to guide the informative response generation (Xing et al., 2017; Xu et al., 2020).

(ii) **Response enhanced by Keywords**: Recent research has incorporated personalized data into the dialogue generation process to enhance the quality of dialogue responses, particularly concerning emotional aspects, viz. emotion (Rashkin et al., 2019), sentiment (Chen and Nakamura, 2021), and politeness (Mishra et al., 2022b; Wang et al., 2020). We will discuss the works that attempt to integrate emotion (Zhou et al., 2018a; Firdaus et al., 2021a; Madasu et al., 2022; Majumder et al., 2022; Mishra et al., 2022c; Samad et al., 2022), sentiment (Firdaus et al., 2021b, 2022a), politeness (Golchha et al., 2019; Firdaus et al., 2020c; Mishra et al., 2022a; Firdaus et al., 2022a; Mishra et al., 2023a, c, b; Priya et al., 2023b), and intent (Xie and Pu, 2021) into the generated responses to make them personalized and engaging.
(iii) **Response enhanced by Internal Knowledge Graph**: Internal knowledge graphs are valuable for comprehending lengthy input sequences. They serve as intermediaries to consolidate or eliminate redundant data, resulting in a concise representation of the input document (Fan et al., 2019; Priya et al., 2023a). Furthermore, KG representations enable the creation of structured summaries and emphasize the connections between related concepts, particularly in cases where complex events associated with a single entity extend across multiple sentences (Huang et al., 2020). In this part of the tutorial, we will present works integrating an internal knowledge graph to enhance response generation capabilities (Liang et al., 2022; Firdaus et al., 2020e).

- **External Knowledge-enhanced Response Generation in Dialogue Systems (60 minutes)**

(i) **Persona Information**: Research focused on personas in dialogue systems requires that the agent adopts a specific character when engaging with users. This persona is closely linked to personality, which influences the emotional and personal aspects of users. In this section of the tutorial, we discuss studies that have employed persona-aware techniques to enhance the efficacy of response generation in dialogue systems (Firdaus et al., 2020f; Saha and Ananiadou, 2022; Firdaus et al., 2022d,b; Zhong et al., 2022). Findings from these studies suggest that persona information drives empathetic and personalized conversations more than non-empathetic ones.

(ii) **Multimodal Information**: Lately, the utilization of multimodal information has witnessed a surge in popularity in the field of dialogue systems. This approach is instrumental in comprehensively understanding users’ emotional and mental states, as it leverages textual and non-textual attributes (Firdaus et al., 2023). In this part of the tutorial, we aim to discuss several notable studies in the literature that have harnessed multimodal data to enhance response generation within dialogue systems (Tavabi et al., 2019; Firdaus et al., 2020a, 2022c).

(iii) **External Knowledge Bases**: Knowledge-grounded systems utilize external resources such as Wikipedia documents to enhance response generation. (Dinan et al., 2018) released the first Wikipedia knowledge-grounded conversation dataset. (Varshney et al., 2023a) utilized the knowledge on various topics such as politics, and movies using the Topical Chat (Gopalakrishnan et al., 2019) and CMU_DoG (Zhou et al., 2018c) dataset to propose a knowledge-emotion enabled conversational model. (Lin et al., 2020) introduced a model that combined knowledge decoders with a pointer network to effectively handle out-of-vocabulary words. Experts suggest converting unstructured knowledge into organized knowledge graphs, composed of triplets (entity, relation, entity/item). Models, such as CCM, retrieve subgraphs from these graphs, especially using knowledge bases like ConceptNet (Speer and Havasi, 2012), and employ attention mechanisms to blend this knowledge into conversations (Zhou et al., 2018b). Concept Flow expands this by including extended subgraph ranges, integrating knowledge from two sources (Zhang et al., 2019). (Varshney et al., 2022a) utilizes both knowledge graphs and Wikipedia documents with a coreference-based knowledge graph augmenting method to improve factual accuracy in dialogue systems.

- **Knowledge-grounded Dialogue Systems in Healthcare (20 minutes)** In healthcare, background knowledge is vital in understanding an individual’s medical history, mental condition, symptoms, and treatment plan. Research has shown that integrating comprehensive knowledge resources in the healthcare dialogue systems offers several key advantages, such as enhancing the system’s grasp of medical concepts and terminology, empowering the system with reasoning and inference capabilities, comprehending emotional dynamics in conversations, and identifying useful response patterns leading to emotional relief (Varshney et al., 2022b; Liang et al., 2021). Driven by these considerations, in this tutorial session, we will discuss the studies that infuse external knowledge in healthcare dialogue systems for providing personalized and effective support (Shen et al., 2022; Deng et al., 2023; Varshney et al., 2022c, 2023b,c; Liu et al., 2021).

- **Hands-on Session (50 minutes)**

1. Setting up a basic knowledge-enhanced dialogue system for healthcare domain (Varshney et al., 2023c,b).
2. Integrating a sample knowledge base (e.g., Unified Medical Language System).
3. Evaluating the performance of the dialogue using automated metrics such as BLEU, F1, and embedding-based metrics.

- **Conclusion and Future Perspectives (20 minutes)**

This tutorial explores notable studies on knowledge-enhanced dialogue generation, showcasing how leveraging diverse information sources can enhance dialogue model efficacy. Despite advancements, several challenges remain, highlighting exciting future re-
search avenues. We’ll delve into four key research directions: (i) Knowledge Acquisition from Pre-trained Language Models: Pre-trained models harbor vast implicit knowledge without external memory reliance (Lewis et al., 2020), opening avenues for efficient knowledge extraction methods like knowledge distillation, data augmentation using pre-trained models as knowledge sources (Petroni et al., 2019), and prompting of language models (Li and Liang, 2021). (ii) Knowledge Acquisition from Limited Resources: In real-world scenarios, new domains often have scarce examples, necessitating rapid adaptation of knowledge-enhanced dialogue models via efficient meta-learning algorithms that minimize task-specific fine-tuning. (iii) Continuous Knowledge Acquisition: A noteworthy exploration is done in (Mazumder et al., 2018), where authors devised a knowledge acquisition engine for chatbots, enabling continuous learning from diverse information sources during interactions. (iv) Leveraging Emotional Knowledge through External Sources: Utilizing emotional knowledge bases like SenticNet aids in discerning user emotional states and background, thus generating emotionally coherent responses, crucial in healthcare and social good applications like persuasion and negotiation.

4. Proposed Length of the tutorial

Half-day (4h long including a coffee break (30m long))

5. Diversity Considerations

This tutorial on Knowledge-enhanced Dialogue Systems (KEDS) emphasizes inclusivity and diversity in three ways: (i) Enhancing Fairness: It educates on designing less biased, more inclusive dialogue systems, promoting equity in healthcare communication tools. (ii) Addressing Unique Needs: It’s relevant to underrepresented groups like healthcare professionals and researchers from certain countries, offering tailored insights. (iii) Diverse Presenters: The presenters, originating from an underrepresented country, embody the commitment to diversity and inclusivity in computational linguistics.

6. Reading List

Extensive reading list is available at Reading List for Knowledge-enhanced Dialogue Systems.

7. Presenters

1. Priyanshu Priya, Indian Institute of Technology Patna, India (priyanshu_2021cs26@iitp.ac.in; priyanshu528priya@gmail.com; LinkedIn)

2. Deeksha Varshney, Indian Institute of Technology Patna, India (deeksha_1821cs13@iitp.ac.in; deeksha.varshney2695@gmail.com; LinkedIn)

3. Mauajama Firdaus, University of Alberta, Canada (mauzama.03@gmail.com; LinkedIn)

4. Asif EKbal, Indian Institute of Technology Patna, India. (asif@iitp.ac.in; asif.ekabi@gmail.com); Webpage: http://www.iitp.ac.in/ asif/; LinkedIn.

8. Other Information

While we are dedicated to accommodate a flexible number of participants, we anticipate an audience of 25-30 people. Our estimate is based on the previous attendance at the tutorial delivered on the topic “Empathetic Conversational Artificial Intelligence Systems: Recent Advances and New Frontiers” was presented at the 32nd International Joint Conference on Artificial Intelligence, held from 19-25 August, 2023 at Macao, S.A.R, China., as well as the outreach efforts we have undertaken to promote the tutorial.

We would appreciate access to standard audio-visual equipment, such as microphones, projectors, and screens, to guarantee the tutorial’s success. Furthermore, a high-speed internet connection is essential to ensure a seamless hands-on session during the tutorial, and an interactive whiteboard might be useful during the presentation for explanatory reasons. This configuration will assist us in facilitating interesting and informative discussions.

9. Ethics Statement

Dialogue systems are becoming ubiquitous in daily applications like healthcare and customer care, necessitating ethical considerations in development and usage. Key considerations include: (i) Knowledge-enhanced dialogue systems can collect sensitive user information, including personal and health data. To safeguard users’ privacy, the data used in the research presented here has been anonymized, and personal details have been protected; (ii) In the context of knowledge-enhanced dialogue systems, user-centric design is essential, ensuring that users have control over the conversation and information sharing. Respecting user autonomy, these systems should offer options to conclude the conversation or seek further assistance. The datasets created for various research topics covered in this tutorial have been crafted to preserve user autonomy.

10. Bibliographical References


Mauajama Firdaus, Nidhi Thakur, and Asif Ekbal. 2021c. Aspect-aware response generation for multimodal dialogue system. ACM Transactions

Mauajama Firdaus, Naveen Thangavelu, Asif Ekbal, and Pushpak Bhattacharyya. 2020f. Persona aware response generation with emotions. In 2020 International Joint Conference on Neural Networks (IJCNN), pages 1–8. IEEE.


Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal. 2022b. Predicting politeness variations in goal-oriented conversations. *IEEE Transactions on Computational Social Systems*.


Deeksha Varshney, Aizan Zafar, Niranshu Kumar Behera, and Asif Ekbal. 2023b. Knowledge graph assisted end-to-end medical dialog generation. Artificial Intelligence in Medicine, 139:102535.


