

Conclusion and Future Directions

What we have so far

Task-oriented Dialogues

- Task-oriented dialogue systems require task-/domain-specific data
 - Strong dependence on individual task characteristics, constraints, etc.
- Task-specific data require modeling the task/domain through schemas, ontologies, etc.
 - In data augmentation there is a chance to make this data driven, but not in zero-shot
- LLMs are proven good UX towards consuming and producing text
 - Including generating dialogue goals
- ... but passing task/domain constraints remains a challenge; even when leveraging LLMs, we need access to constraints such as schemas, or ontologies. They are mostly human-generated and not easily integrated in an e2e process

What we have so far

Open Domain Dialogue systems

- Data augmentation is proven effective for various types of open domain dialogue systems
- Methods have moved from Generative to Prompting based
 - Minimizes the need for human involvement
 - It is faster and more accessible
- General trend in LLM-based data augmentation:
 - Create Large-scale LLM-generated datasets; e.g., using GPT* models
 - (Parameter-efficient) Finetune another LLM (e.g., LLaMA) to generate a dialogue agent
 - E.g., for role-specified open domain dialogue systems, information seeking systems
- It still requires domain-specific knowledge (i.e., seed data, structural constraints)

Open Challenges

- There is less control over the generated data
 - Limited guards against unsafe and toxic content
 - Large-scale automatic evaluation and human evaluation is still an open problem
- LLM-generated dialogues lead to self-reinforcement of LLM-based dialogue systems
 - We already know LLM-based evaluation models prefer LLM-generated text
- Large scale data generation for complex and personalized tasks remains a challenge
 - E.g., tutoring tasks, modeling personas and preferences,