Chatbots:

- In commercial:
  - Benefits:
    - Automatic customer service
    - Reducing human labor and costs
- In academia:
  - Challenges:
    - Understanding semantics
    - Generating coherent responses
    - Deficient datasets
      - Domain specific datasets: small and oriented for a certain task. Example: Movie dialogue dataset
      - Open domain dataset: having too many turns making the data disperse too disperse to capture main idea. Example: Reddit forum posts
      - Social network datasets: too noisy, unreal and short. Example: Twitter dataset

DailyDialog:

Construction:

Data was crawled from websites for English learners practicing daily conversations.

Basic features:

More like real conversations.

Specific topics

Reasonable speaker turns (8) makes it better for training a chatbot

2 parties involved in dialogues

No grammar or misspelling

13,118 dialogues

Dialogue Purpose:

- Information exchange:
  - Key: communication intentions (dialogue act)
  - Dialogue act classes: {Inform, Questions, Directives, Commissive}
  - Interaction of acts pairs:
    - Previous utterance: Inform, response: Inform/Questions
    - Previous utterance: Questions, response: Inform
    - Previous utterance: Directives, response: commisive
- Enhancing social bonds:
  - Key: Emotions
  - Emotion classes: based on “Big six theory”: {anger, happiness, disgust, sadness, fear, surprise} and a “other emotions” class.
  - Labelling is done by experts with agreement of 78.9%
  - Emotion distribution:
    - Other>Joy>surprise>sadness>anger>disgust>fear

Characteristics:

Covers daily topics:

- Not domain specific
- Top topics: daily life, relationships, work

Bi turn:

- Natural dialogue flow
- not like QAs (like SubTle dataset) or post-reply (like reddit comment) which have ambiguous dialogue flows
- dialogue flows:

Questions-inform: respond + propose own idea

Directives-commissives

Human style Communication patterns:

Not like task oriented datasets

Pattern1:

info providers changes to info seeker

18.3%

Pattern2:

when people have different views and they exchange and persuade

Directives-directives-commissives

9.2%

Emotions:

rich and high quality labeled

shows empathy and understanding

28% happy dialogues

0.8 % dialogues changing from negative mood to positive mood

Approaches:

Embedding-based similarity:

Rank based on distance between embeddings

Feature-based similarity:

Rank with TF-IDF, fuzzy features

Feature-based similarity and reranking by intentions

Feature-based similarity and reranking by intentions, emotions

[1] Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, Shuzi Niu. “DailyDialog: A Manually Labelled Multi-turn Dialogue Dataset”