Summary of Relevant Papers

Table of Contents

[A Persona-Based Neural Conversation Model 3](#_Toc49965475)

[Description 3](#_Toc49965476)

[Datasets 3](#_Toc49965477)

[Evaluation 3](#_Toc49965478)

[An Ensemble Model with Ranking for Social Dialogue 3](#_Toc49965479)

[Description 3](#_Toc49965480)

[Datasets 3](#_Toc49965481)

[Evaluation 3](#_Toc49965482)

[Dynamic Knowledge Routing Network for Target-Guided Open-Domain Conversation 4](#_Toc49965483)

[Description 4](#_Toc49965484)

[Datasets 4](#_Toc49965485)

[Evaluation 4](#_Toc49965486)

[IEC: Towards Interest-Eliciting Neural Conversational Agents 4](#_Toc49965487)

[Description 4](#_Toc49965488)

[Datasets 4](#_Toc49965489)

[Evaluation 4](#_Toc49965490)

[Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings 4](#_Toc49965491)

[Description 4](#_Toc49965492)

[Datasets 5](#_Toc49965493)

[Evaluation 5](#_Toc49965494)

[MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling 5](#_Toc49965495)

[Description 5](#_Toc49965496)

[Datasets 5](#_Toc49965497)

[RUBER: An Unsupervised Method for Automatic Evaluation of Open-Domain Dialog Systems 5](#_Toc49965498)

[Description 5](#_Toc49965499)

[Evaluation 5](#_Toc49965500)

[Target-Guided Open-Domain Conversation 5](#_Toc49965501)

[Description 5](#_Toc49965502)

[Datasets 6](#_Toc49965503)

[Evaluation 6](#_Toc49965504)

[The Dialogue Dodecathlon: Open-Domain Knowledge and Image Grounded Conversational Agents 6](#_Toc49965505)

[Description 6](#_Toc49965506)

[Datasets 6](#_Toc49965507)

[Evaluation 6](#_Toc49965508)

[Towards a Human-like Open-Domain Chatbot 6](#_Toc49965509)

[Description 6](#_Toc49965510)

[Datasets 6](#_Toc49965511)

[Evaluation 6](#_Toc49965512)

[Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset 7](#_Toc49965513)

[Description 7](#_Toc49965514)

[Datasets 7](#_Toc49965515)

[Evaluation 7](#_Toc49965516)

[Unsupervised Context Rewriting for Open Domain Conversation 7](#_Toc49965517)

[Description 7](#_Toc49965518)

[Datasets 7](#_Toc49965519)

[Evaluation 7](#_Toc49965520)

[Facebook Patent (Providing Personal Assistant Service Via Messaging) 7](#_Toc49965521)

[Description 7](#_Toc49965522)

# A Persona-Based Neural Conversation Model

## Description

* Incorporate persona for empathetic conversations.
* One person talks differently with two people like Ross – Monica and Ross – Rachel.
* Two models: Speaker model, Speaker – Addressee model
* Speaker addressee encodes interaction patterns
* Response adaptive model where model adapts to the way person communicates.
* Lexical Entrainment
* Seq2Seq LSTM Model

## Datasets

* Twitter Persona Dataset (Collected)
* Twitter Sordoni Dataset (Old collected by Sordoni)
* IMSDb Scripts: Friends, The Big Bang Theory

## Evaluation

* BLEU

# An Ensemble Model with Ranking for Social Dialogue

## Description

* Alexa Price Challenge. Conversate coherently and engagingly with humans for 20 minutes.
* Create ensemble of existing chatbots to provide results
* Evaluation of the ensemble is the challenge as to which chatbot should be used for which sentence.
* Created score function that retains context and incorporates change of context.
* Bot Priority list – context priority list – ranking function
* When no results, random fun fact
* Tried creating Seq2Seq using Twitter and Reddit for small talk conversations

## Datasets

* Nothing but different chatbots.
* Persona: Rule based AIML / alicebot. Guarantees appropriate responses to inappropriate responses.
* Rosie: Eliza (1966) style chatbot
* NewsBot: Provides news context using newsapi.org
* Factbot, Quiz Game, WeatherBot

## Evaluation

* Custom score functions: Coherence (Semantic similarity), Flow (Prevent repetition), Questions (To continue conversation), Named Entities (Promote candidates relating to same topic), Noun Phrase, Dullness, Topic Divergence (LDA), Sentiment Polarity.

# Dynamic Knowledge Routing Network for Target-Guided Open-Domain Conversation

## Description

* Guide human conversations during an open-ended conversation.
* Introduces new dataset Sina Weibo (Social Platform)
* Problems: Define goal in open-domain conversation, define strategy to achieve that goal, achieve the goal
* Decouples the system into modules by training keyword predictor and response retrieval.
* Use semantic knowledge for smooth keyword transition
* Start off with generic topic and guide the conversation. Only agent knows the goal.
* Knowledge routing method to predict keyword for next response with current context.

## Datasets

* Sina Weibo Open domain conversations

## Evaluation

* Evaluation for keyword prediction and response retrieval.
* Custom recall-based evaluation.

# IEC: Towards Interest-Eliciting Neural Conversational Agents

## Description

* Identify user’s interests without asking. Scenario where user does not want to show interest at all.
* Detect user’s interests and changes that steer conversations E.g. comics to music.
* Leave dull and dry conversations
* 3 Models: User Behavior Simulation, Topic Detector, Topic-Aware Response Generator
* User Interest Prediction Model

## Datasets

* Reddit 10 subreddit divisions that cover 10 topics.

## Evaluation

* Average user interest prediction accuracy over 500 conversations
* BLEU, Dist 1, Dist 2, ENT 4

# Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings

## Description

* Find mutual friend based on open-domain conversations.
* Symmetric collaborative dialogue setting: task-oriented but open-domain.
* Bot-Bot conversations

## Datasets

* 11K human-human dialogues based on Amazon Mechanical Trunk

## Evaluation

* Cross-entropy
* Human paired with bot or human and scored results
* Human evaluation

# MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling

## Description

* Provides a 10,000-dialogue dataset which is fully labelled.
* Fully labelled, human-human written conversations, multiple domain topics

## Datasets

* This is a dataset

# RUBER: An Unsupervised Method for Automatic Evaluation of Open-Domain Dialog Systems

## Description

* Evaluation metrics for open ended conversational chatbots
* Human – Machine evaluation
* Scoring based on reply and previous query by the user
* Flexible and extensible to different datasets and languages

## Evaluation

* This is an evaluation metric
* BLEU
* METEOR
* ROUGE

# Target-Guided Open-Domain Conversation

## Description

* Guide open conversations to a particular arbitrary goal that defines the shape of the conversation
* Find coarse-grained keywords that control the conversation and then smoothly perform the transition to turn the conversation.
* Open-domain bot problems: what is the goal of the conversation? Define some goal and achieve it.
* Predict keywords and then generate responses
* Pairwise PMI-based Transition for changes in the keyword by creating a pairwise matrix to find association between the keywords
* Combine neural network next keyword prediction with pairwise closeness to drive the conversation

## Datasets

* Dataset derived from PersonaChat Corpus
* Broad ranged topics and topics change frequently in the conversations

## Evaluation

* Keyword Performance Evaluation using Precision, Recall and Cor. (Correlation scoring)
* Self-play simulation: chatbot must achieve its target in 8 steps
* Human evaluation

# The Dialogue Dodecathlon: Open-Domain Knowledge and Image Grounded Conversational Agents

## Description

* 12 tasks that measure if chatbot can communicate engagingly with personality, empathy, ask questions and answer questions.
* Use large datasets to help in other skills
* Uses image and text input to generate response for every dialogue

## Datasets

* ConvAI2, DailyDialog, Wiz.ofWikipedia, Empathetic Dialogues, Cornell Movie, LIGHT, ELI5, Ubuntu, Twitter, pushshift.io Reddit, Image Chat, IGC

## Evaluation

* Perplexity
* BLEU
* ROUGE-1, -2 & -L
* F1

# Towards a Human-like Open-Domain Chatbot

## Description

* Multi-turn open domain chatbot designed to chat on large topic of conversations.
* Also called as Google Meena
* Comparable to GPT-3 as they claim that it is better than GPT-2
* Heavy comparison about SSA and other metrics

## Datasets

* Public domain social media conversations

## Evaluation

* Sensibleness and Specificity Average (SSA) custom designed metric

# Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset

## Description

* Generate empathetic responses by a chatbot
* Consider and identify feelings to understand the context
* Compare with BERT

## Datasets

* Custom dataset with 25k dialogues with emotions

## Evaluation

* BLEU
* Gold Label
* Human evaluation

# Unsupervised Context Rewriting for Open Domain Conversation

## Description

* Context rewriting method that rewrites the last utterance by considering context history
* Challenges that it tackles: 1) topic identification, 2) Coreferences like he, she, it, they 3) long term dependencies
* CRN that integrates key information and last utterance to build a rewritten one
* Use RL to provide rewritten last utterance for fine tuning

## Evaluation

* BLEU-4
* Embedding Average
* Embedding Extrema
* Embedding Greedy
* Distinct 1 & 2 to evaluate response diversity

# Facebook Patent (Providing Personal Assistant Service Via Messaging)

## Description

* Provide messenger services that determine intent of request and help in conversation

# Project MAIA: Multilingual AI Agent Assistant

## Description

* Build AI assistive agents that eliminate language barriers
* Travel and Tourism industry assistive chat that helps humans with assistive suggestions
* Real-time machine translation
* Suggestive feed for human agent to reduce effort and increase customer satisfaction
* Sentiment analysis with empathy