Summary of Relevant Papers

Table of Contents

[A Persona-Based Neural Conversation Model 5](#_Toc60698699)

[Description 5](#_Toc60698700)

[Datasets 5](#_Toc60698701)

[Evaluation 5](#_Toc60698702)

[An Ensemble Model with Ranking for Social Dialogue 5](#_Toc60698703)

[Description 5](#_Toc60698704)

[Datasets 5](#_Toc60698705)

[Evaluation 5](#_Toc60698706)

[Dynamic Knowledge Routing Network for Target-Guided Open-Domain Conversation 6](#_Toc60698707)

[Description 6](#_Toc60698708)

[Datasets 6](#_Toc60698709)

[Evaluation 6](#_Toc60698710)

[IEC: Towards Interest-Eliciting Neural Conversational Agents 6](#_Toc60698711)

[Description 6](#_Toc60698712)

[Datasets 6](#_Toc60698713)

[Evaluation 6](#_Toc60698714)

[Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings 6](#_Toc60698715)

[Description 6](#_Toc60698716)

[Datasets 7](#_Toc60698717)

[Evaluation 7](#_Toc60698718)

[MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling 7](#_Toc60698719)

[Description 7](#_Toc60698720)

[Datasets 7](#_Toc60698721)

[RUBER: An Unsupervised Method for Automatic Evaluation of Open-Domain Dialog Systems 7](#_Toc60698722)

[Description 7](#_Toc60698723)

[Evaluation 7](#_Toc60698724)

[Target-Guided Open-Domain Conversation 7](#_Toc60698725)

[Description 7](#_Toc60698726)

[Datasets 8](#_Toc60698727)

[Evaluation 8](#_Toc60698728)

[The Dialogue Dodecathlon: Open-Domain Knowledge and Image Grounded Conversational Agents 8](#_Toc60698729)

[Description 8](#_Toc60698730)

[Datasets 8](#_Toc60698731)

[Evaluation 8](#_Toc60698732)

[Towards a Human-like Open-Domain Chatbot 8](#_Toc60698733)

[Description 8](#_Toc60698734)

[Datasets 8](#_Toc60698735)

[Evaluation 8](#_Toc60698736)

[Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset 9](#_Toc60698737)

[Description 9](#_Toc60698738)

[Datasets 9](#_Toc60698739)

[Evaluation 9](#_Toc60698740)

[Unsupervised Context Rewriting for Open Domain Conversation 9](#_Toc60698741)

[Description 9](#_Toc60698742)

[Evaluation 9](#_Toc60698743)

[Facebook Patent (Providing Personal Assistant Service Via Messaging) 9](#_Toc60698744)

[Description 9](#_Toc60698745)

[Project MAIA: Multilingual AI Agent Assistant 9](#_Toc60698746)

[Description 9](#_Toc60698747)

[DIA: A Human AI Hybrid Conversational Assistant for Developing Contexts 10](#_Toc60698748)

[Description 10](#_Toc60698749)

[Datasets 10](#_Toc60698750)

[Evaluation 10](#_Toc60698751)

[Connecting People Through Virtual Assistant on Google Assistant 10](#_Toc60698752)

[Description 10](#_Toc60698753)

[Patent: AI Assistant for Interacting with Customers Across Multiple Communication Modes 10](#_Toc60698754)

[Description 10](#_Toc60698755)

[Datasets 10](#_Toc60698756)

[Evaluation 10](#_Toc60698757)

[Affect-LM: A Neural Language Model for Customizable Affective Text Generation 11](#_Toc60698758)

[Description 11](#_Toc60698759)

[Datasets 11](#_Toc60698760)

[Evaluation 11](#_Toc60698761)

[Affective Natural Language Generation by Phrase Insertion 11](#_Toc60698762)

[Description 11](#_Toc60698763)

[Datasets 11](#_Toc60698764)

[Evaluation 11](#_Toc60698765)

[Designing Conversation-context Recommendation Display to Support Opportunistic Search in Meetings 11](#_Toc60698766)

[Description 11](#_Toc60698767)

[Datasets 12](#_Toc60698768)

[Evaluation 12](#_Toc60698769)

[Dynamic Online Conversation Recommendation 12](#_Toc60698770)

[Description 12](#_Toc60698771)

[Datasets 12](#_Toc60698772)

[Evaluation 12](#_Toc60698773)

[Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space 12](#_Toc60698774)

[Description 12](#_Toc60698775)

[Datasets 12](#_Toc60698776)

[Evaluation 12](#_Toc60698777)

[Neural Conversation Recommendation with Online Interaction Modeling 12](#_Toc60698778)

[Description 12](#_Toc60698779)

[Datasets 12](#_Toc60698780)

[Evaluation 13](#_Toc60698781)

[Offline and Online Satisfaction Prediction in Open-Domain Conversational Systems 13](#_Toc60698782)

[Description 13](#_Toc60698783)

[Datasets 13](#_Toc60698784)

[Evaluation 13](#_Toc60698785)

[Predicting User Intents and Satisfaction with Dialogue-based Conversational Recommendations 13](#_Toc60698786)

[Description 13](#_Toc60698787)

[Datasets 13](#_Toc60698788)

[Evaluation 13](#_Toc60698789)

[The Impact of Humanoid Affect Expression on Human Behavior in a Game-Theoretic Setting 13](#_Toc60698790)

[Description 13](#_Toc60698791)

[Datasets 13](#_Toc60698792)

[Evaluation 13](#_Toc60698793)

[Topic Recommendation Method related to a Present Topic for Continuing a Conversation 14](#_Toc60698794)

[Description 14](#_Toc60698795)

[Datasets 14](#_Toc60698796)

[Evaluation 14](#_Toc60698797)

[Towards a Conversational Corpus for Human-Robot Conversations 14](#_Toc60698798)

[Description 14](#_Toc60698799)

[Datasets 14](#_Toc60698800)

# 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

# DIA: A Human AI Hybrid Conversational Assistant for Developing Contexts

## Description

* Create context learning chatbot where resources are low for scaling expert knowledge
* Chatbot learns topic-specific knowledge and local language from user interactions
* Helps in building dataset of topic and language specific dialogues
* Create a chatbot that can help sub-Saharan Africans to teach underprivileged students with limited resources
* Overcome local language written in English as those datasets are very limited
* Learn topic specific knowledge using user interactions such as WhatsApp messages

## Datasets

* TaRL (Specific teaching manuals) and use cosine similarity
* DialogFlow for basic intents
* Actual human experts for WoZ

## Evaluation

* Human evaluation

# Connecting People Through Virtual Assistant on Google Assistant

## Description

* Connect two random strangers using Google Assistant
* Uses same features of that of Google Assistant

# Patent: AI Assistant for Interacting with Customers Across Multiple Communication Modes

## Description

* Insurance assistive chatbot that helps agents to guide the users
* Several modes of assistance like text, email, call

## Datasets

* Company’s data

## Evaluation

* Human evaluation

Papers on Chat Satisfaction

# Affect-LM: A Neural Language Model for Customizable Affective Text Generation

## Description

* LSTM based generative language model

## Datasets

* Fischer
* DAIC
* SEMAINE
* CMU-MOSI

## Evaluation

* Perplexity

# Affective Natural Language Generation by Phrase Insertion

## Description

* Created Bi-Directional GRU for sentiment analysis on Cornell Movie Dialogue Corpus
* Emotions selected were sadness, anger, doubt, happiness, affirmation and love, with two intensities: moderate and intense.

## Datasets

* Cornell Movie Dialogue Corpus
* The EuroParl
* British National Corpus
* Manually Annotated Sub-Corpus of The Open American National Corpus
* Santa Barbara Corpus of Spoken American English
* The Ubuntu Dialogue Corpus

## Evaluation

* Using other datasets apart from Cornell
* BLEU
* ROUGE
* Perplexity
* Human Evaluation

# Designing Conversation-context Recommendation Display to Support Opportunistic Search in Meetings

## Description

* Provide search results based on brainstorming meeting session
* Find accidental information encountered during a meeting that makes people search that
* Gave table-top display to each user and Wizard-of-Oz approach to search
* Gave only one keyboard to a group to search
* Paper written to market the product

## Datasets

* Display the information as the user speaks
* Uses Wizard-of-Oz approach for speech recognition

## Evaluation

* People like to use keyboard more than the product

# Dynamic Online Conversation Recommendation

## Description

* I

## Datasets

* C

## Evaluation

* H

# Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space

## Description

* Virtual chat environment where people are on a video call, suggest topics to talk about when people run out of things to talk
* Analyze conversations and when there are long pauses, the bot will ask yes/no questions to each party and suggest new topic to talk about and then wait again

## Datasets

* Created knowledge base of topics that can be talked about by surveys
* Created questions that will help users to find a topic that they can talk about

## Evaluation

* Perplexity

# Neural Conversation Recommendation with Online Interaction Modeling

## Description

* I

## Datasets

* C

## Evaluation

* H

# Offline and Online Satisfaction Prediction in Open-Domain Conversational Systems

## Description

* I

## Datasets

* C

## Evaluation

* H

# Predicting User Intents and Satisfaction with Dialogue-based Conversational Recommendations

## Description

* I

## Datasets

* C

## Evaluation

* H

# The Impact of Humanoid Affect Expression on Human Behavior in a Game-Theoretic Setting

## Description

* Create model that is used to play Stackelberg’s Security Game (Strategy game with defensive / mixed strategy).
* Humans are asked to play against a bot to evaluate the game
* Goal is to find how humans can be influenced by robot’s expressions
* First human plays with a computer for baseline model and then custom-built bot
* The bot will either be encouraging or discouraging the strategy
* Created N-gram model

## Datasets

* NLTK: brown, guten- berg, inaugural, state of the union, and genesis (English text).

## Evaluation

* N-gram model works well
* Humans understood the feelings of the robot: encouraging / discouraging

# Topic Recommendation Method related to a Present Topic for Continuing a Conversation

## Description

* I

## Datasets

* C

## Evaluation

* H

# Towards a Conversational Corpus for Human-Robot Conversations

## Description

* Created a dataset for human-robot conversations
* Humans do not talk with robots as they do with other humans
* Wizard-of-Oz dataset that was created with speech recognizer and text to speech
* Could be useful to point out if humans interact with robots the same way as that of humans in an open-ended conversational framework

## Datasets

* Presented their own dataset