IJCAI/2023 MACAO

32nd International Joint Conference on Artificial Intelligence

Empathetic Conversational Artificial Intelligence Systems: *Recent Advances and New Frontiers*

Priyanshu Priya
priyanshu528priya@gmail.com

Dr. Mauajama Firdaus
mauzama.03@gmail.com

Kshitij Mishra
kmishra.kings@gmail.com

Dr. Asif EKbal
asif.ekbal@gmail.com

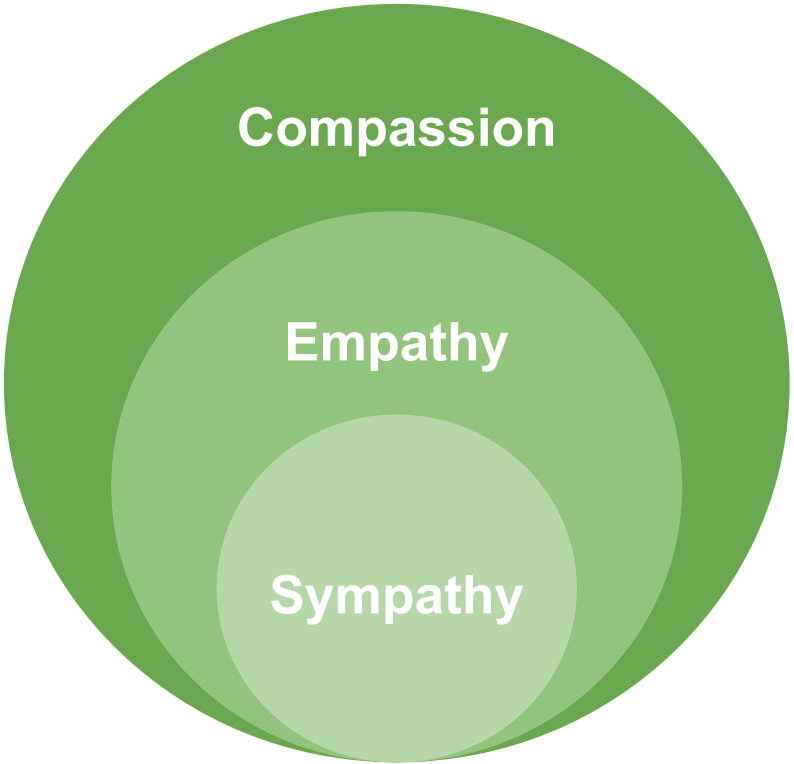
Conceptual Models of Empathy

(15 minutes)

Empathy

“An emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one’s perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with the recognition that the source of the emotion is not one’s own” - *Benjamin MP Cuff*

Understanding Empathy



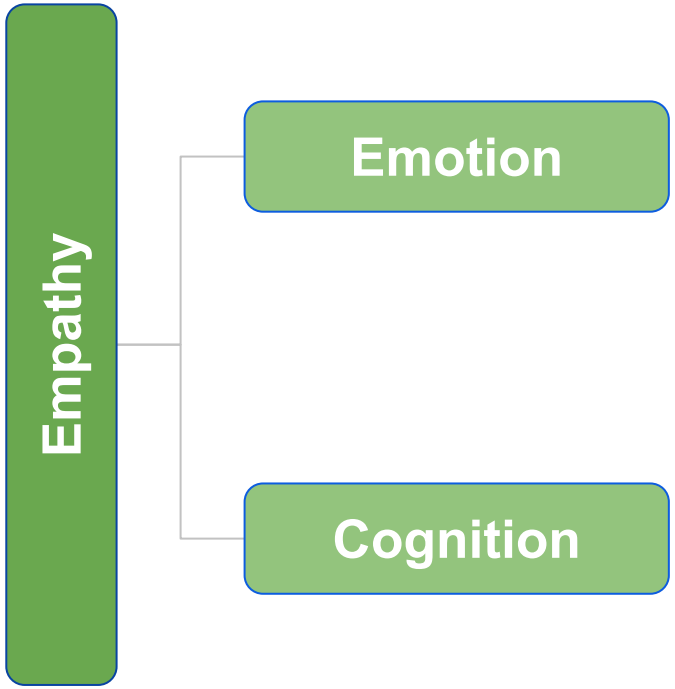
Seeing with the ***eyes of another***, listening with the ***ears of another*** and feeling with the ***heart of another***.

Sympathy: “I’m sorry that happened to you.”

Empathy: “I see your pain and I understand.”

Compassion: “How do you need me to help?”

Empathy as Multi-dimensional Construct

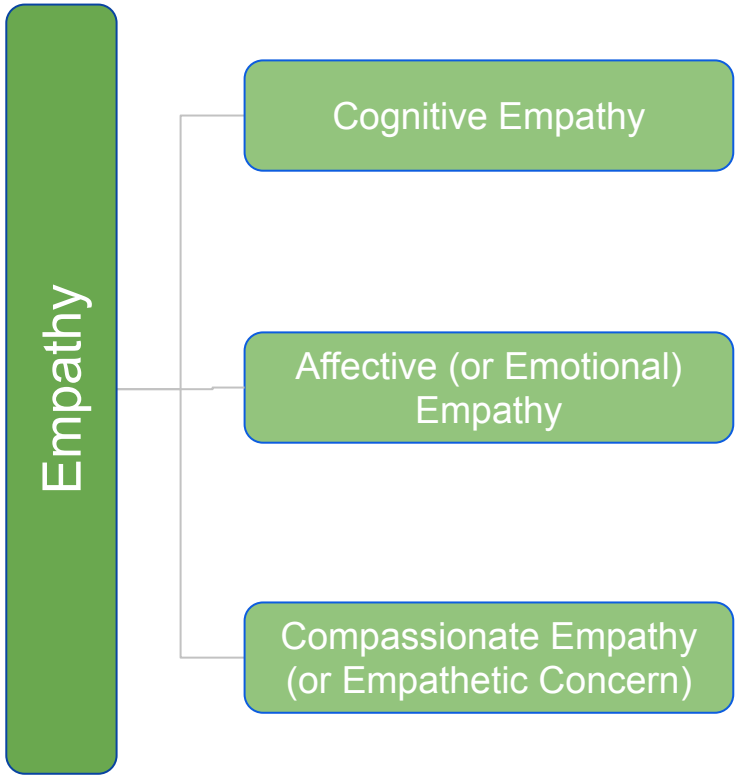


Relates to the *emotional stimulation in reaction to the experiences and feelings* expressed by a user

A more deliberate process of *understanding and interpreting the experiences and feelings* of the user and communicating that understanding to them

Mark H. Davis 1980

Empathy as Multi-dimensional Construct



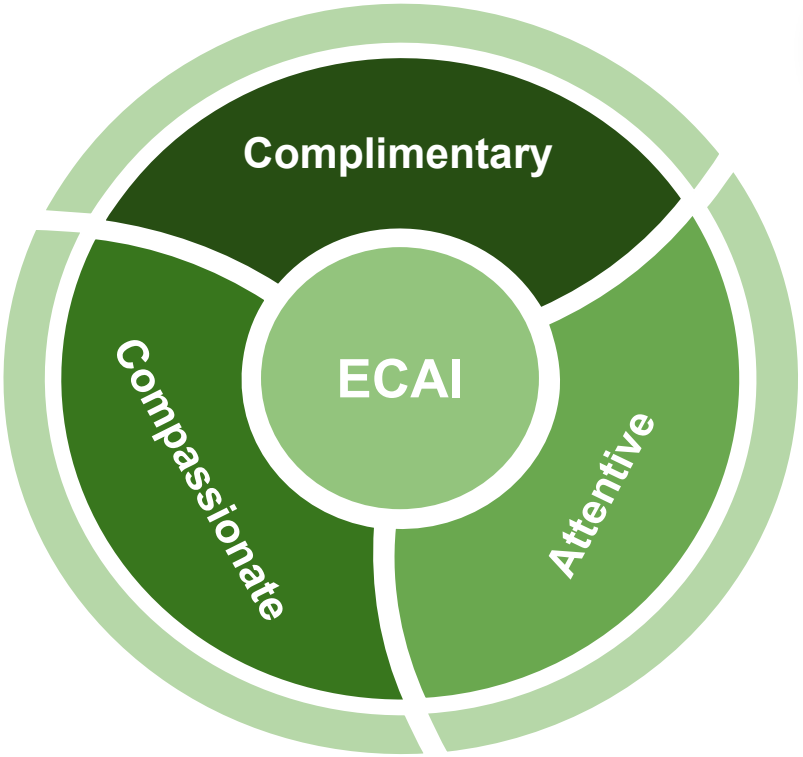
Ability to understand another person's point of view or perspective ("*putting oneself in others' shoes*")

Ability to feel what someone else feels ("*your pain in my heart*")

Ability to sense what another person needs from you and do something helpful ("*demonstrates helpful behaviors*")

D. Goleman 1995

Empathy as Multi-dimensional Construct Cont..



- Human-like systems (bridge human-machine gap)
- Better and more meaningful user engagement
- Enhances emotional bond with users
- Perceived as social actors by users
- Leads to positive user experience and effective communication

Need for Empathy in Conversational AI Systems

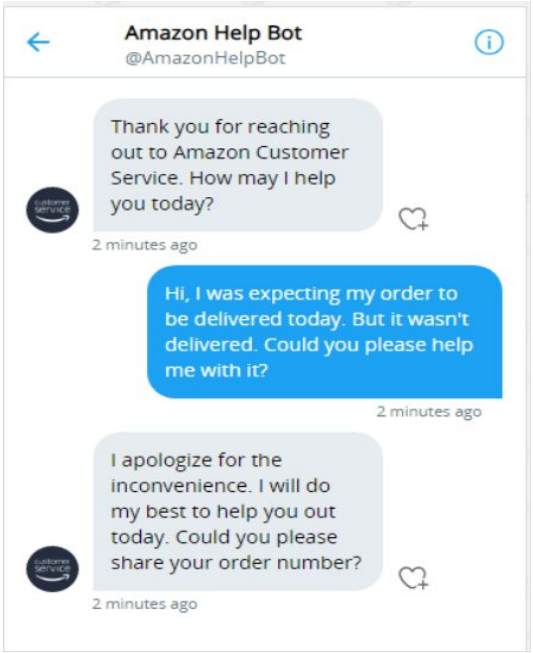
(15 minutes)

Empathy in AI-assisted Customer Care System

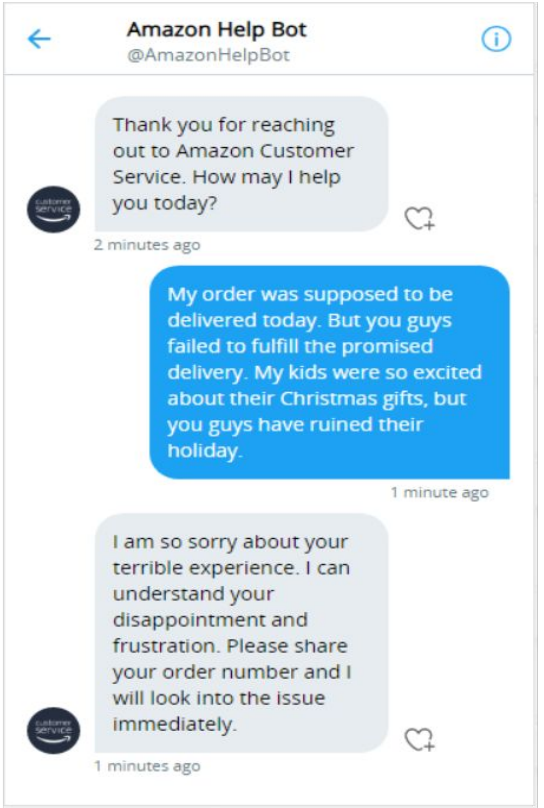
Compared to scenario 1, the customer sounds more distressed in scenario 2.

An ideal response may start with first acknowledging the understanding of the customer's frustration and displaying empathy to subdue their negatively charged emotions.

A compassionate choice of words in the response can not only alleviate some of the customer annoyance, but it may also help in better customer retention in the long run.



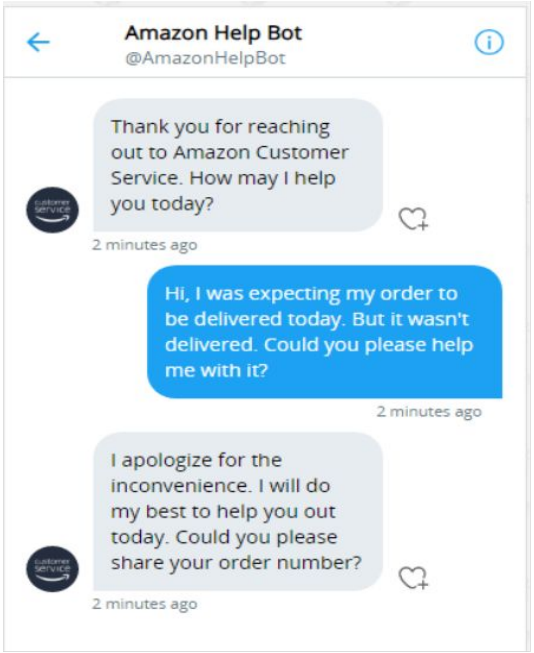
Scenario 1



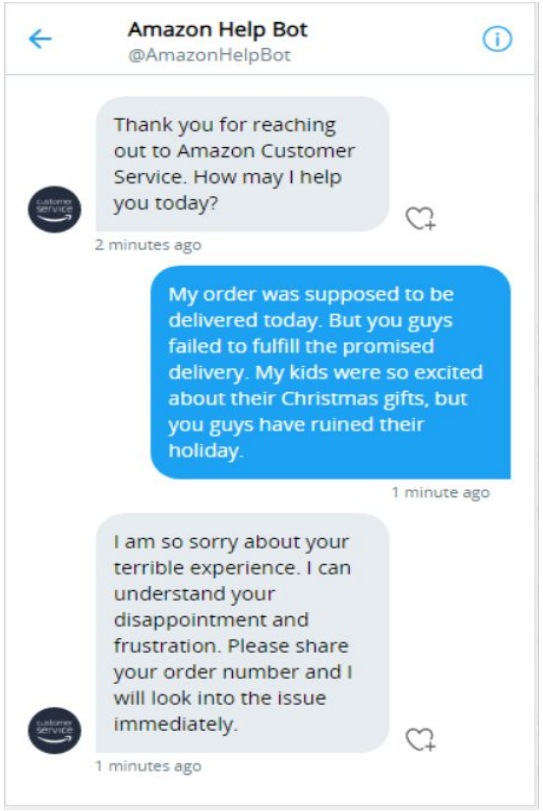
Scenario 2

PC: https://blog.reachsumit.com/posts/2020/12/generating-empathetic-responses/

Amazon Help Bot



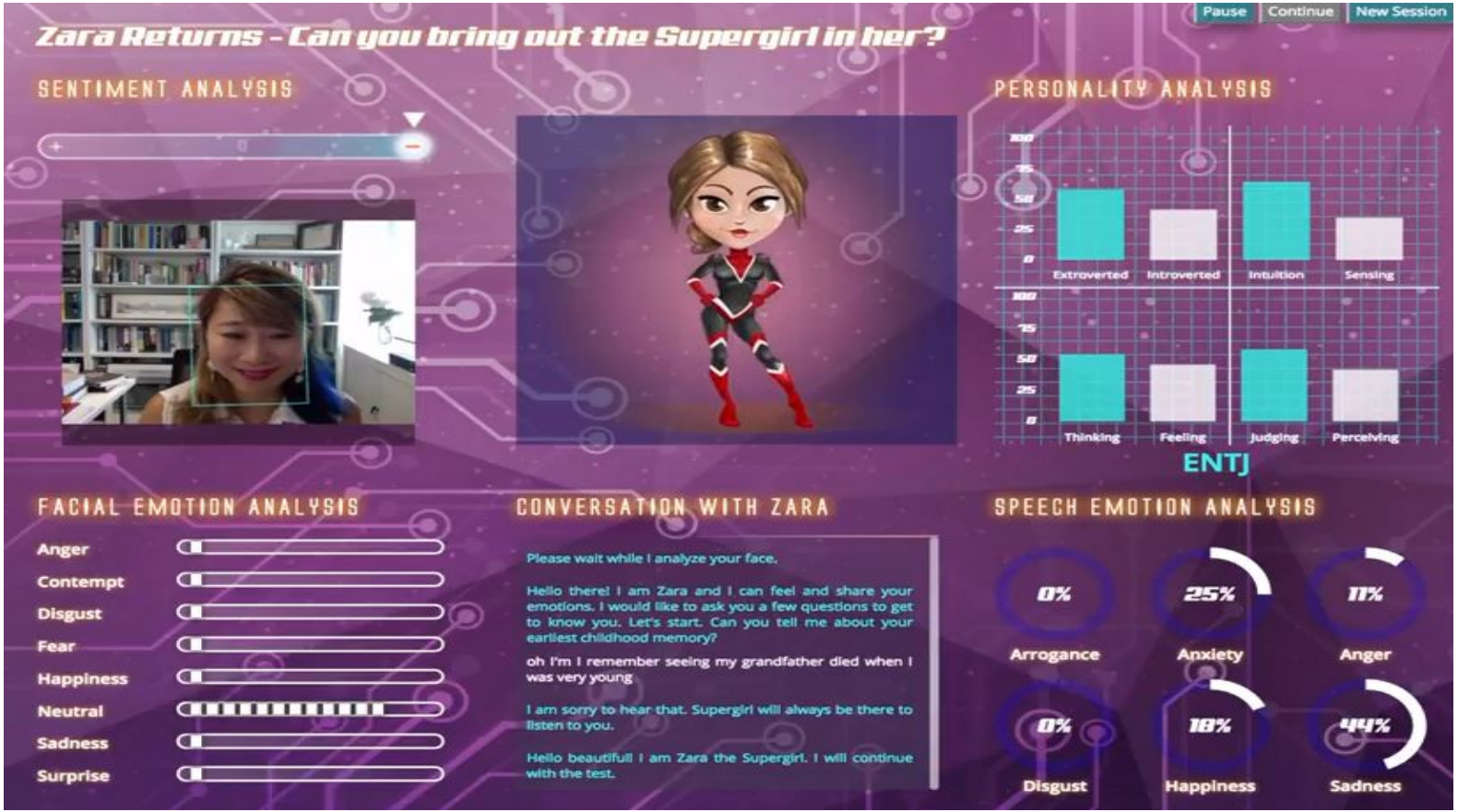
Scenario 1



Scenario 2

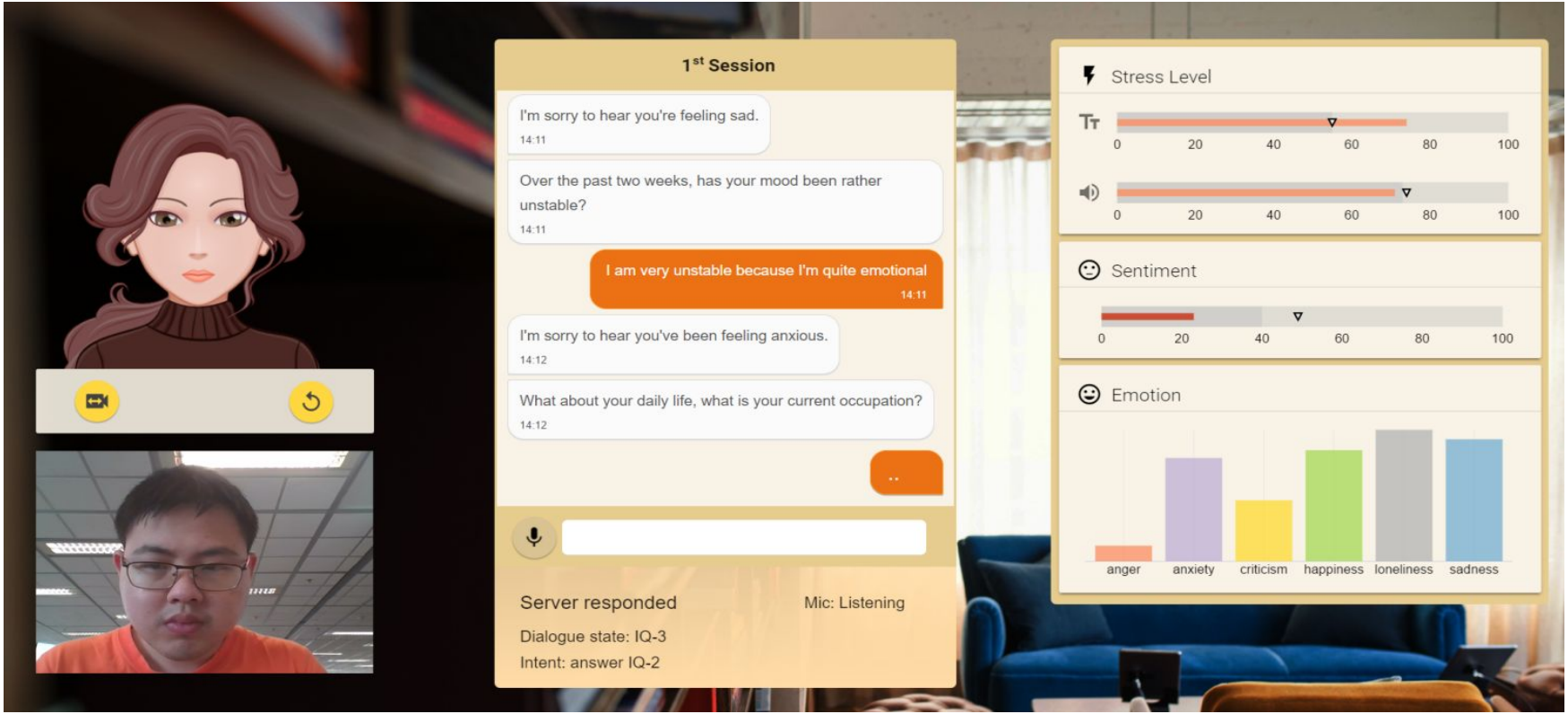
Source: https://blog.reachsumit.com/posts/2020/12/generating-empathetic-responses/

Zara The SuperGirl



Source: https://blog.reachsumit.com/posts/2020/12/generating-empathetic-responses/

Nora the Empathetic Psychologist



Source:: https://blog.reachsumit.com/posts/2020/12/generating-empathetic-responses/

Empathy-related Concepts in Conversational AI Systems

(135 minutes)

Empathy and Related Concepts

01

Emotion/Sentiment

Problem Definition

- Generate the response conditioned on different emotions along with emotion intensity
- For Example, *“Oh my God!!! How could you treat him in this manner.”* (surprise, anger)
    - Absence of one of the emotion - the entire meaning of the utterance left incomplete
    - Here *“Oh my God”* emphasizes that the anger of the user is due to unawareness of the situation leading to surprise emotion as well in the utterance

Motivation

- Existing systems lack completeness in terms of emotional content in the responses as humans inherently express multiple feelings in their day-to-day conversations
- In addition, the intensity of emotions fluctuates in an utterance providing variations for a particular emotion in conversations

Firdaus et al. More the merrier: Towards multi-emotion and intensity controllable response generation. In AAAI 2021

Multiple Emotion and Intensity aware Multi-party Dialogue (MEIMD) Dataset

- Every utterance is labeled with corresponding emotions and intensity values to facilitate multi-emotion and intensity controlled response generation
- For data collection, 8 famous TV shows belonging to the different genres were considered:
  - Drama: Breaking Bad, Castle, Game of Thrones, Grey's Anatomy, and House M.D.
  - Comedy: Friends, How I Met Your Mother and The Big Bang Theory
  - In total, there are 507 episodes, spanning 456 hours
- 7 emotion labels
  - anger, acceptance, disgust, fear, joy, sadness, surprise
  - intensity values ranging from 0-3

Firdaus et. al More the merrier: Towards multi-emotion and intensity controllable response generation. In AAAI 2021

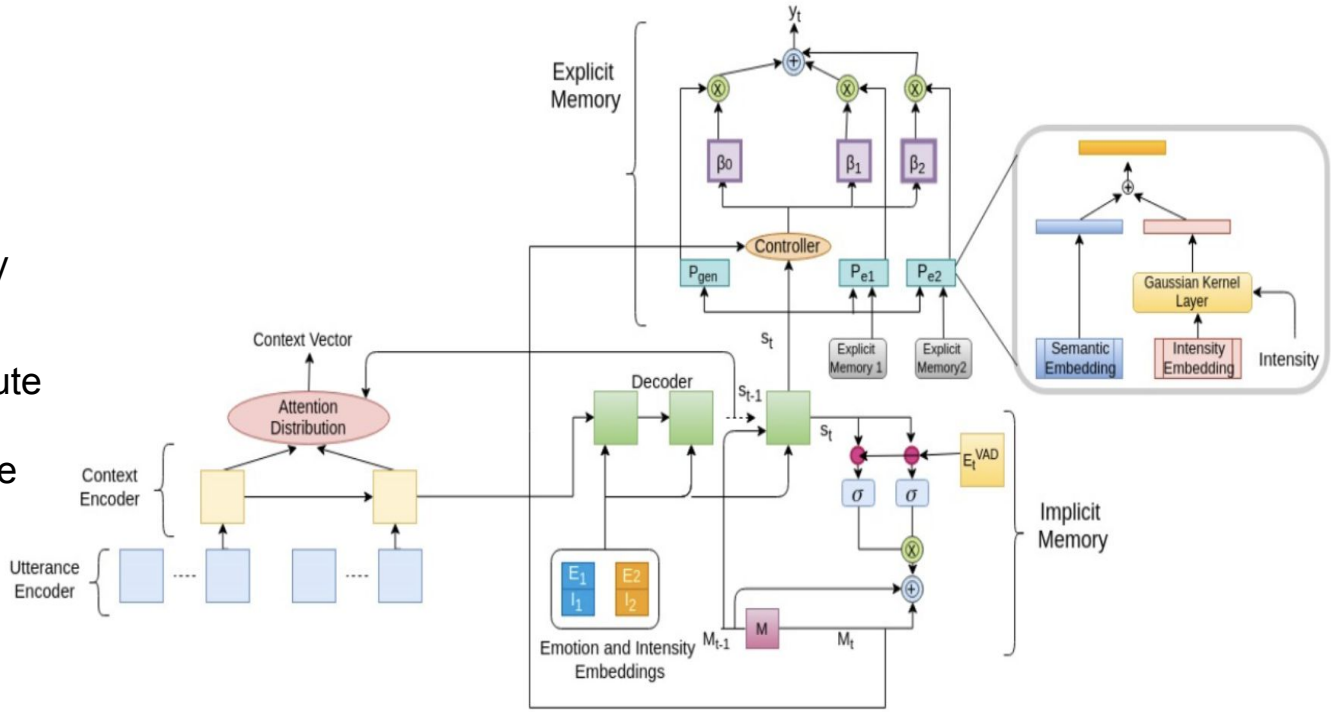
Multiple Emotion and Intensity aware Multi-party Dialogue (MEIMD) Dataset

	Conversations	Emotions
1	It's amazing, I am thrilled you got promoted I have loads of work now and am afraid to complete it. Stop sulking, I am sure you will manage it.	Surprise (0.3), Joy (0.9) Disgust(0.3), Fear(0.6) Anger(0.3), Acceptance(0.6)
2	I am sorry this could be an infection or cancer. I am afraid but I know you could help me.	Sadness(0.6), Fear(0.3) Acceptance(0.3), fear(0.6)

Firdaus et. al More the merrier: Towards multi-emotion and intensity controllable response generation. In AAAI 2021

Multiple Emotion with Intensity-based Dialogue Generation Framework

- *Explicit Memory:* Determine whether to generate an emotion or generic word, while focusing on the intensity of the desired emotions
- *Implicit Memory:* Compute the number of words remaining to express the emotion completely, thereby regulating the generation accordingly



Firdaus et. al More the merrier: Towards multi-emotion and intensity controllable response generation. In AAAI 2021

Evaluation Metrics

Automatic Evaluation

- **Perplexity**
- **Embedding Scores-based Metrics**
  - Average
  - Greedy
  - Extreme
- **Emotion Content**
  - *Macro average weighted F1*: Evaluate the generated response at content level
  - *Pearson Correlation Coefficient*: Evaluate the generated response at emotion level

Human Evaluation

- **Fluency and Relevance:** Evaluate if the generated response is linguistically fluent and relevant
- **Emotion:** Evaluate the emotion quotient of the generated response
- **Intensity:** Evaluate the degree of the emotion expressed in the generated response

Firdaus et. al More the merrier: Towards multi-emotion and intensity controllable response generation. In AAAI 2021

Problem Definition

Generate the emotional and personalized response in accordance to the conversational history, sentiment and the persona information of the speaker

Motivation

- Persona information in a reply may not sufficient to produce interactive responses. To render it more human-like, the emotional element must also be integrated into the replies.
- Every emotion is associated with sentiments, hence using the sentiment information of the utterances can assist in narrowing down the set of emotions for generating contextually correct emotional responses.

Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

Dataset

- Used ConvAI2 benchmark dataset, which is an extended version (with a new test set) of the persona-chat dataset
- Dataset Statistics
  - 10,981 dialogues with 164,356 utterances
  - 1,155 personas, each consisting of at least four personality texts.
  - 1,016 dialogues in the testing set and 200 never before seen personas.

Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

Dataset Annotation

- Annotated the dataset utilizing emotion annotated version of the dataset used in (Firdaus et al., 2020)
- Use the emotion information to annotate the PersonaChat dataset with sentiments due to the high correlation between emotions and sentiment
    - Utterances with emotions such as *excited, grateful, joyful, caring, hopeful, faithful, impressed* are labeled as *positive sentiment*
    - Utterances with emotions such as *angry, sad, annoyed, disgusted, terrified, furious, disappointed, jealous* has a negative undertone, hence are labelled as *negative sentiment*
    - For the other emotion labels such as *surprise, proud, nostalgic, guilty, confident, prepared, sentimental* that can either be positive, neutral or negative depending on the utterance and the context we resort to manual annotation

Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

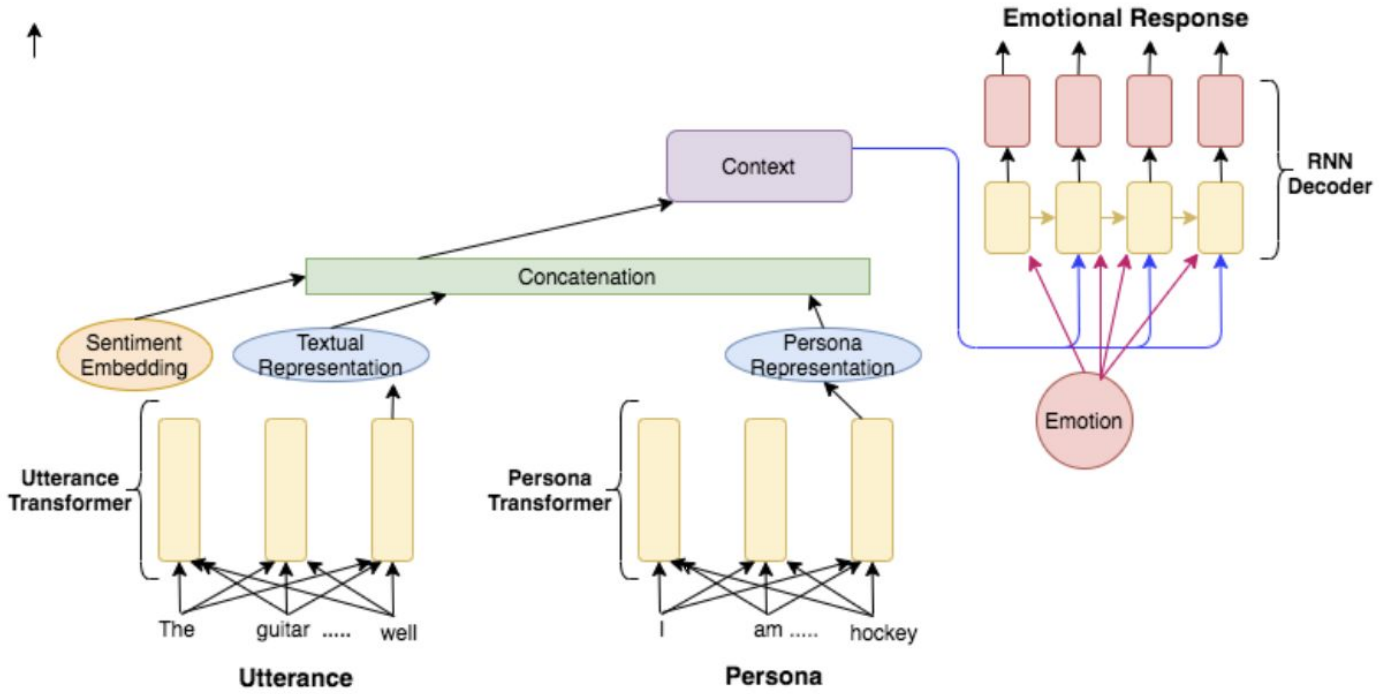
A conversation from the PersonaChat dataset with sentiments

Persona 1	Persona 2
I am primarily a meat eater.	*I've a sweet tooth.*
I am a guitar player.	*I'm a babysitter and drive a mercedes.*
Welding is my career field.	*I'm the middle child of 3 siblings.*
My parents don't know I am gay.	*I'm getting married in six months.*

[Person 1] What do you do for career? (Neutral)
[Person 2] I like to watch kids. (Positive)
[Person 1] I actually play guitar and do a lot of welding. (Positive)
[Person 2] What do you weld? houses?(Neutral)

Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

Sentiment and Persona guided Emotional Dialogue Generation Framework



Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU-4 and ROUGE-L**
- **Distinct-1 and Distinct-2:** Measure the distinct n-grams in the generated responses.
- **Emotion Accuracy:** Measure the emotional content in the generated responses

Human Evaluation

- **Fluency and Relevance:** Evaluate if the generated response is linguistically fluent and relevant on a scale of 1-5
- **Persona Consistency, Sentiment Coherence, Emotion Appropriateness:** Evaluate the persona, sentiment and emotion inclusion in the generated response

Firdaus et al. SEPRG: Sentiment aware Emotion controlled Personalized Response Generation. In INLG 2021

02

Emotion Cause

Problem Definition

- Recognize emotion cause words in dialogue utterances
- Make dialogue models better focus on these targeted words to generate more specific empathetic responses

Motivation

- Empathy is a complex cognitive ability based on the reasoning of others' affective states
- To better understand others and express stronger empathy in dialogues, underlying emotion cause words need to be identified and emphasized

User: I got a gift from my friend, last vacation!
Agent: Wow! how was the vacation?
User: You're not listening, are you?

vs.

User: I got a gift from my friend, last vacation!
Agent: Wow! what kind of gift?
User: Lego Discovery Space Shuttle! I'm so excited to build it.

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

How do we humans recognize emotions?

Do we use emotion cause labels? 🤔

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

We put ourselves in the other's shoes

Simulating what it would be like if we were in that situation

Perspective-taking

The act of perceiving a situation or understanding a concept from alternating point of view

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

The work aims to reason the emotion-cause weight

of each word in utterances, while satisfying the following three desiderata

[A] Do not require word-level supervision
We human do not need them

[B] Simulate the observed interlocutor's situation within the model
Much evidence for this behavior is found from cognitive science including empathetic perspective-taking and mirror neurons

[3] Reason other's internal emotional states in Bayesian fashion
Studies from cognitive science argue that affective reasoning can be described via Bayesian inference [1]

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

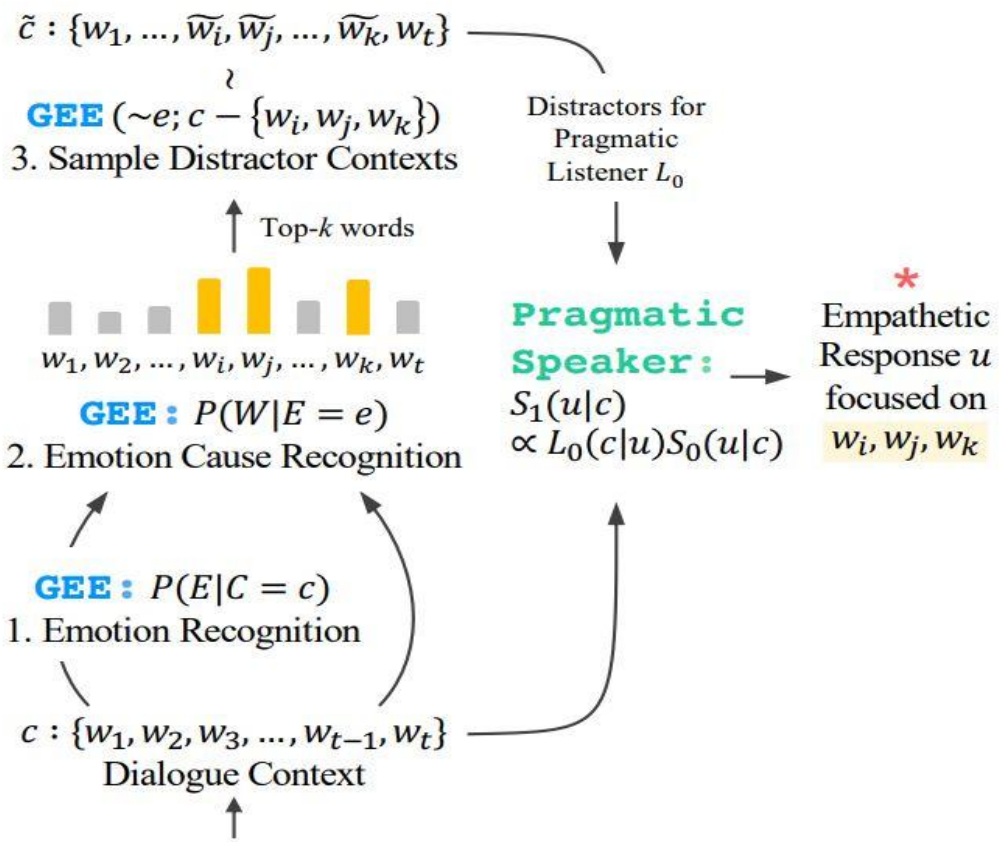
Overview of the proposed Approach

Generative Emotion Estimator (GEE)
models P(C, E) = P(E)P(C|E) with text sequence (e.g. context) C and emotion E

First, the generative estimator computes the likelihood of C by generating C given E, which can be viewed as a simulation of C.

Second, it estimates P(E|C) via Bayes' rule.

Finally, the association between the emotion estimate and each word comes for free by using the likelihood of each words; without using any word-level supervision *(weakly supervised)*.



Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

Dataset

- To train GEE, the EmpatheticDialogues dataset [2] was used
  - a multiturn English dialogue dataset
  - the speaker talks about an emotional situation and the listener expresses empathy
  - 24,850 conversations in total
  - 32 emotion labels that are evenly distributed.

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

Automatic Evaluation: *Recognize emotion cause words*

EmoCause

- Evaluation set to measure the performance of GEE
- Annotate emotion cause words on the situation in EmpatheticDialogues [2] validation and test sets

Metrics

Report the **Top-1, 3, 5 recall scores**

Emotion	Situation
Surprised	Man, I did not expect to see a bear on the road today.
Afraid	I have to take a business trip next week, I'm not looking forward to flying.
Sad	I feel sad that I am spending so much time this late on the internet.
Joyful	I'm excited I get to go to Disney in October!

Examples of annotated emotion cause words

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

Automatic Evaluation: *Generate empathetic responses based on those words*

Coverage: Average no. of emotion-cause words included in the generated responses

Exploration and Interpretation [3]
Metrics for empathy expressed in text
Requires responses to focus on interlocutor's utterance and to be specific
Measured by trained RoBERTa models

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

Human Evaluation: *Generate empathetic responses based on those words*

Evaluation on 4-point Likert Scale

- **Empathy/Sympathy:** *did the responses show understanding of the feelings of the person talking about their experience?*
- **Relevance:** *did the responses seem appropriate to the conversation? Were they on-topic?*
- **Fluency:** *could you understand the responses? Did the language seem accurate?*

Kim et al. Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes. In EMNLP 2021.

03

Intent

Problem Definition

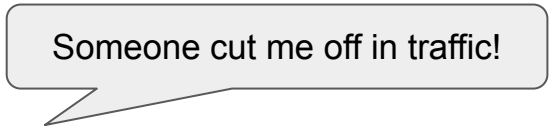
- Generate responses that are contextually coherent and emotionally appropriate
- Model a fine-grained set of empathetic response intents in an empathetic dialog model to ensure a more precise learning of the emotional interactions revealed in the dialogue

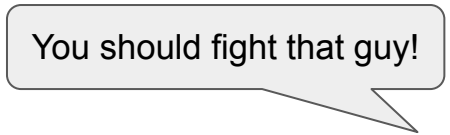
Motivation

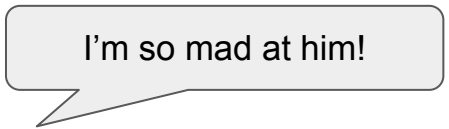
- Incorporate empathetic response intents explicitly into the design of dialog systems
    - To capture the subtle interactions in human conversations, where the listener often exhibits empathetic intents that are more neutral

Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

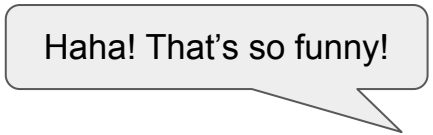
A conversation Example

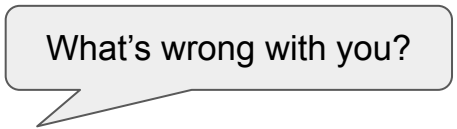




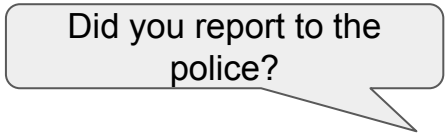


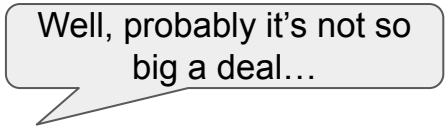
Simply following the speaker's emotion state





Reversing the speaker's emotion state





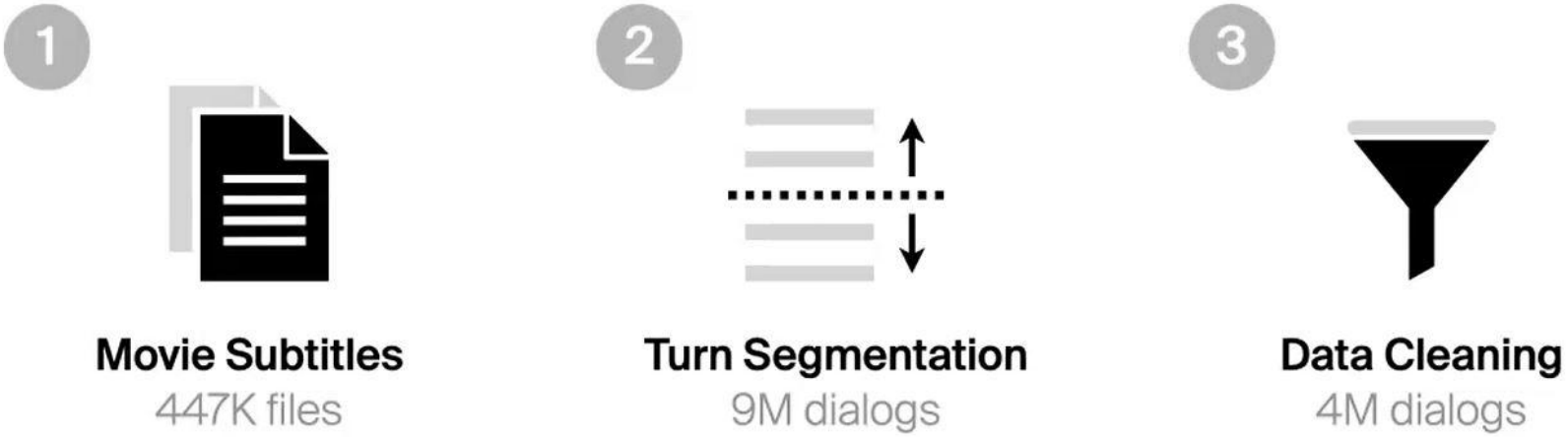
Responding with questioning

Leaves the speaker in angry state (or even escalates the situation)

Calms down the speaker and drives the conversation to a empathetic direction.

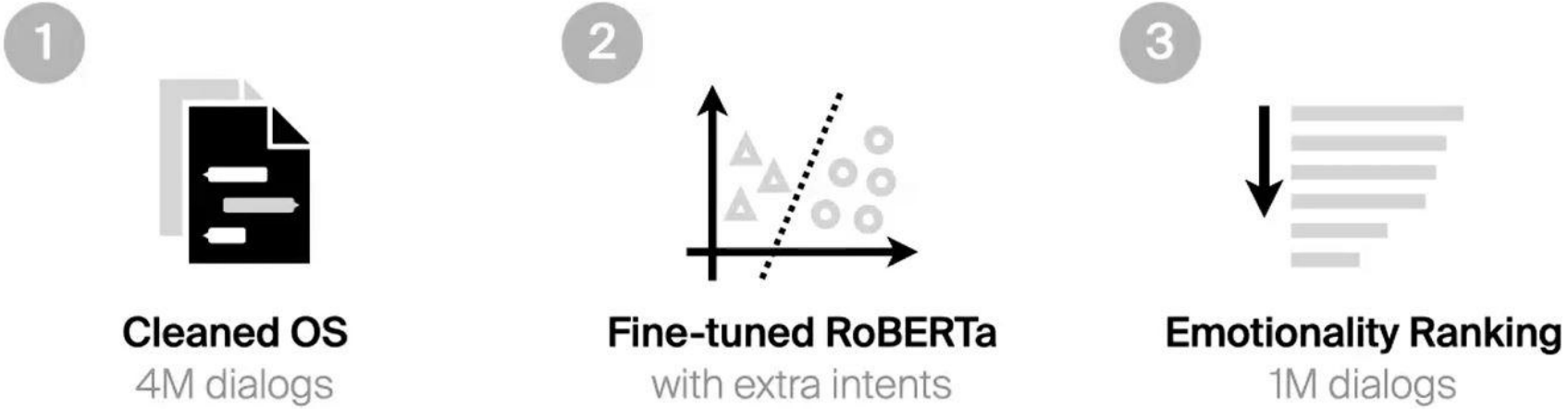
Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

Data Curation (OpenSubtitles Dialogs)



Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

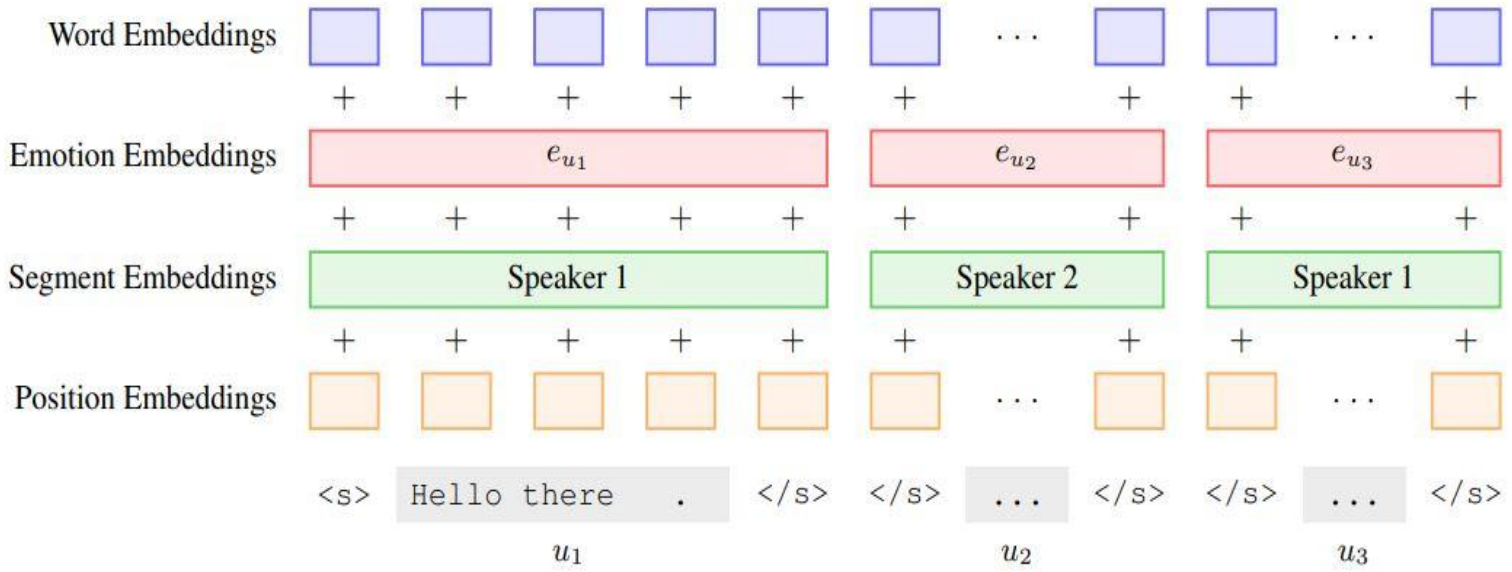
Data Curation (Emotional Dialogs in OpenSubtitles)



Empathetic intents: *questioning, agreeing, acknowledging, sympathizing, encouraging, consoling, suggesting, and wishing, neutral*

Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

Proposed Method (Input Representation)



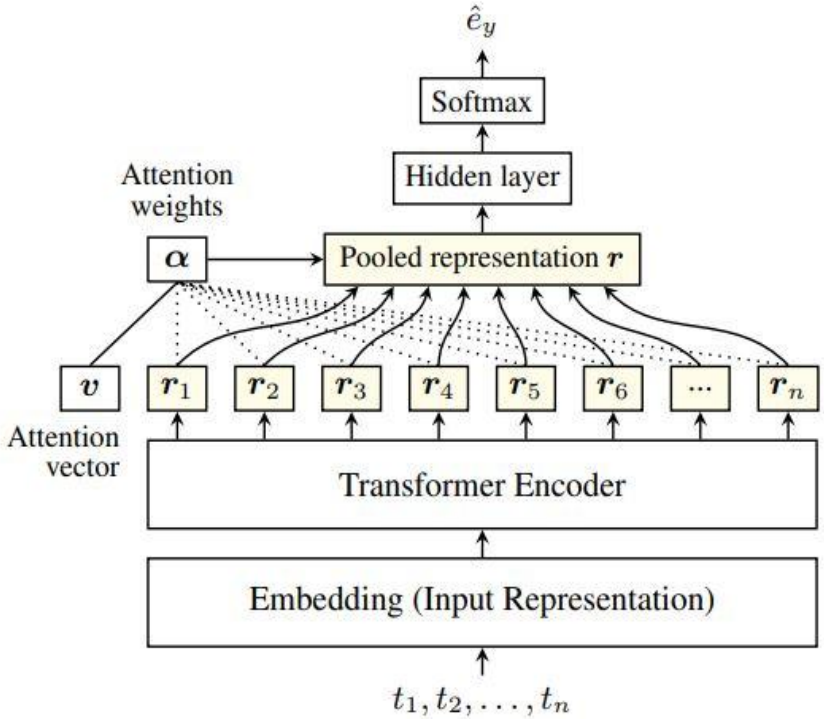
Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

Proposed Method (Response Emotion/Intent Predictor)

- Transformer encoder to get context-dependent representations
- Pool using attention (Dotted lines in Figure):

$$\boldsymbol{\alpha}_i = \frac{\exp(\boldsymbol{v}^T \boldsymbol{r}_i)}{\sum_{j=1}^{N} \exp(\boldsymbol{v}^T \boldsymbol{r}_j)}$$

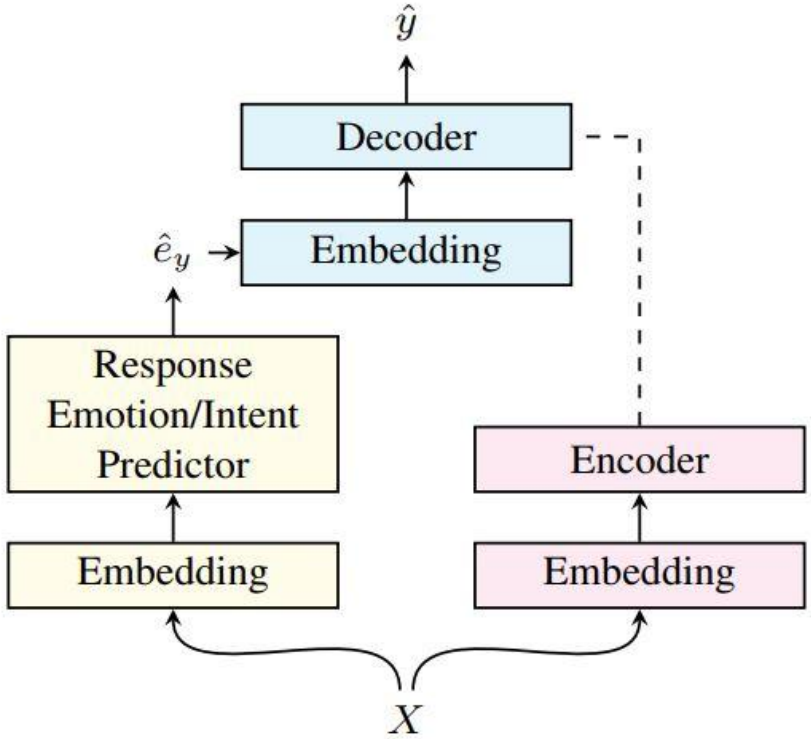
$$\boldsymbol{r} = \sum_{i=1}^{N} \boldsymbol{\alpha}_i \boldsymbol{r}_i$$



Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

Proposed Method (Training)

- Trained separately
- Response emotion/intent Predictor
  - Cross-entropy loss between predicted and gold emotion/intent category
- Encoder-Decoder
  - Cross-entropy loss between predicted and gold response



Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

Experiments

- **Datasets**
  - OpenSubtitles dialog (OS): 3M
  - Emotional dialog in OpenSubtitles (EDOS): 1M
  - EmpatheticDialogues (ED): 25K
- **Splitting**
  - 80%, 10%, 10% for training, validation and testing
  - 6,000 testing dialogs for human evaluation (2,000 from each dataset)

Evaluation

Automatic Evaluation

- **Response Emotion/Intent Predictor**
  - *Weighted Precision, Recall and F1-scores*
- **Encoder-Decoder**
  - **Perplexity:** Measure how well a probability model predicts a given sample
  - **Distinct-1 and Distinct-2:** Measure the degree of diversity by calculating the no. of distinct unigrams and bigrams, respectively in the generated responses
  - **Sentence Embedding Similarity:** Calculate the cosine similarity between the embeddings of predicted and ground-truth responses

Human Evaluation Setup

- **A new evaluation strategy**
  - Combination of Likert Scale or side-by-side comparison (A/B testing)
  - Drag and drop the candidate replies to one of three areas: *Good, Okay,* and *Bad*

Xie et al. Empathetic Dialog Generation with Fine-Grained Intents. In CoNLL 2021.

04

Persona

Problem Definition

- Persona aware emotional response generation
  - The system is able to generate specific and consistent responses in accordance to the provided personality information and the conversational history

Motivation

- Every individual has a personality and is driven by emotions
  - The ability to converse with a consistent personality helps in bringing consistency and specificity in responses
- This work intend to infuse the emotions in the responses that help in making the responses more human-like
  - Make the responses interactive and interesting

Firdaus et al. Persona aware Response Generation with Emotions. In IJCNN 2020

Example from the Dataset

Persona 1	Persona 2
As a child , I won a national spelling bee.	*I'm very athletic.*
I've been published in the new yorker magazine.	*I have brown hair.*
I am a gourmet cook.	*I love bicycling.*
I've perfect pitch.	*I hate carrots.*
[Person 1] Hi! i work as a gourmet cook.	
[Person 2] I don't like carrots. I throw them away.	
[Person 1] Really. But, I can sing pitch perfect .	
[Person 2] I also cook, and I ride my bike to work.	

- Speakers maintain the persona information while conversing with each other
  - Make the conversation interactive and also facilitate building user's trust and confidence.

- Response to Person 1 could be more empathetic like *"That's a great job, but I don't like carrots and throw them away."*

Firdaus et al. Persona aware Response Generation with Emotions. In IJCNN 2020

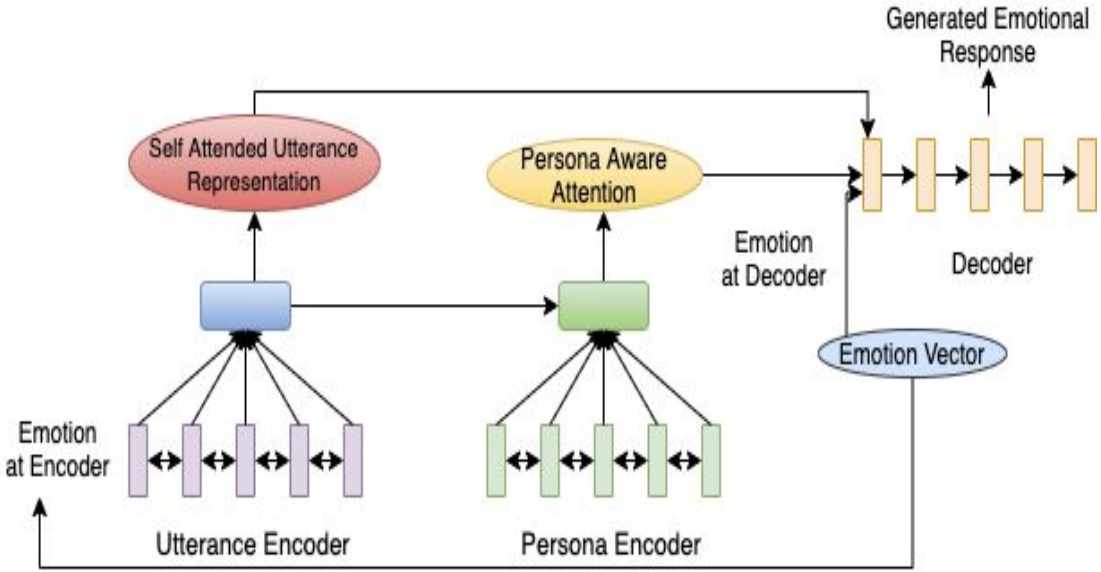
Dataset

- Used ConvAI2 benchmark dataset, which is an extended version (with a new test set) of the persona-chat dataset
- Dataset Statistics
  - 10,981 dialogues with 164,356 utterances
  - 1,155 personas, each consisting of at least four personality texts.
  - 1,016 dialogues in the testing set and 200 never before seen personas.

Firdaus et al. Persona aware Response Generation with Emotions. In IJCNN 2020

Proposed Framework

- Utterance and Persona encoders followed by a decoder for generating the desired emotional responses
- Persona-aware attention enables the model to focus on different personas mentioned in the utterance



Firdaus et al. Persona aware Response Generation with Emotions. In IJCNN 2020

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU-4 and ROUGE-L:** Measure the ability of the generated response for capturing the correct information
- **Emotion Accuracy:** Measure the emotional content in the generated responses

Human Evaluation

- **Fluency:** Measures the grammatical correctness of the generated response
- **Emotion:** Judges whether the generated response is in accordance with the desired emotions
- **Persona Consistency:** Measures the response generated is in accordance with the persona information of the speaker provided in the form of texts and is also coherent with the conversational history

Firdaus et al. Persona aware Response Generation with Emotions. In IJCNN 2020

Problem Definition

- Generate empathetic, personalized responses while considering the persona information and implicitly the emotion in the responses through the dialogue context

Motivation

- Social chatbots have gained immense popularity, and their appeal lies in their capacity to respond to diverse requests, and their ability to develop an emotional connection with users
- To develop and promote social chatbots, both the intellectual and emotional quotient needs to be introduced in conversational agents
    - Increases user interaction

Firdaus et al. I enjoy writing and playing, do you: A Personalized and Emotion Grounded Dialogue Agent using Generative Adversarial Network. In IEEE Transactions on Affective Computing 2022

Example from the Dataset

Persona 1	Persona 2
I am primarily a meat eater.	*I've a sweet tooth.*
I am a guitar player.	*I'm a babysitter and drive a mercedes.*
Welding is my career field.	*I'm the middle child of 3 siblings.*
My parents don't know I am gay.	*I'm getting married in six months.*
[Person 1] What do you do for career?	
[Person 2] I like to watch kids.	
[Person 1] I actually play guitar and do a lot of welding.	
[Person 2] What do you weld? houses?	

- Speakers maintain the persona information while conversing with each other
  - Make the conversation interactive and also facilitate building user's trust and confidence.

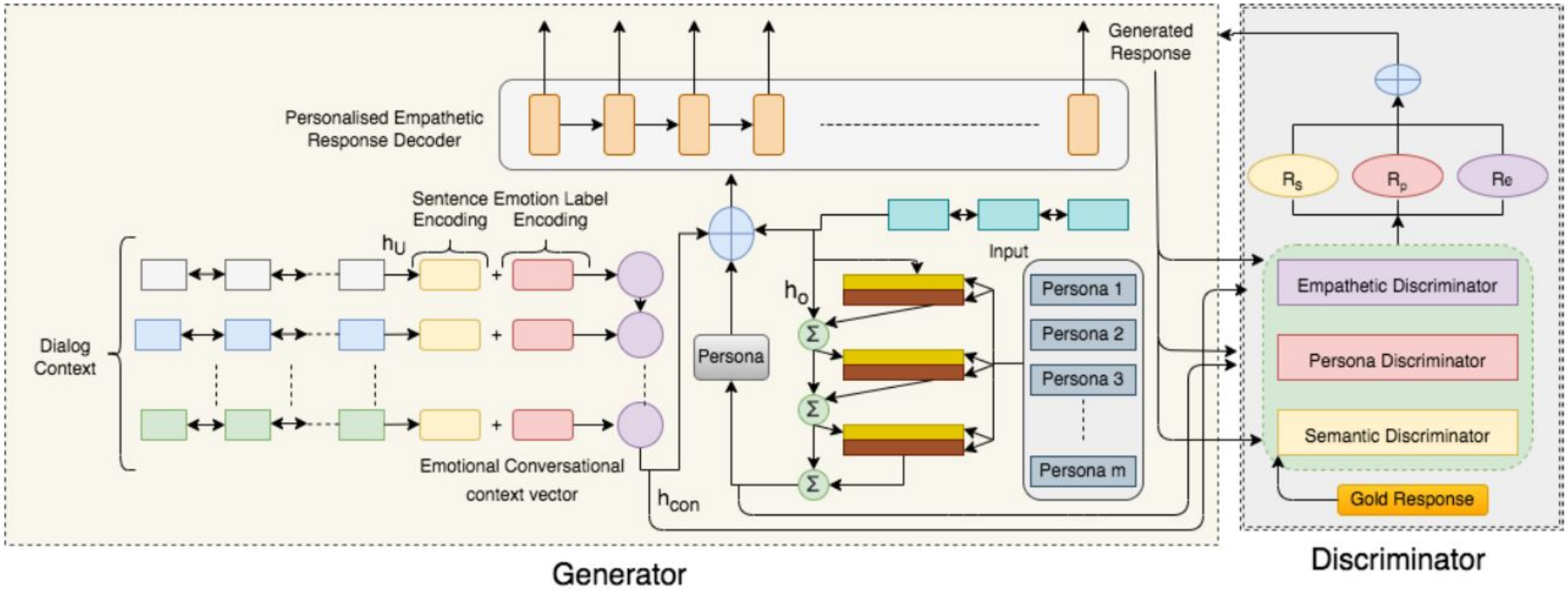
Firdaus et al. I enjoy writing and playing, do you: A Personalized and Emotion Grounded Dialogue Agent using Generative Adversarial Network. In IEEE Transactions on Affective Computing 2022

Dataset

- Used ConvAI2 benchmark dataset, which is an extended version (with a new test set) of the persona-chat dataset
- Dataset Statistics
  - 10,981 dialogues with 164,356 utterances
  - 1,155 personas, each consisting of at least four personality texts.
  - 1,016 dialogues in the testing set and 200 never before seen personas.
- Emotion Annotation
  - Semi-supervised approach
  - Used DistilBERT fine-tuned on EmpatheticDialogues dataset with 32 emotion labels

Firdaus et al. I enjoy writing and playing, do you: A Personalized and Emotion Grounded Dialogue Agent using Generative Adversarial Network. In IEEE Transactions on Affective Computing 2022

Empathy and Persona aware Generative Adversarial Network



Firdaus et al. I enjoy writing and playing, do you: A Personalized and Emotion Grounded Dialogue Agent using Generative Adversarial Network. In IEEE Transactions on Affective Computing 2022

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU-4 and ROUGE-L:** Measure the ability of the generated response for capturing the correct information
- **Distinct-1 and Distinct-2:** Measure the degree of diversity by calculating the no. of distinct unigrams and bigrams, respectively in the generated responses
- **Emotion Accuracy:** Measure the emotional content in the generated responses

Human Evaluation

- **Fluency:** Measures the grammatical correctness of the generated response
- **Relevance:** Evaluates whether the responses are on-topic with the dialogue history
- **Emotion Appropriateness:** Judges whether the generated response is in accordance with the desired emotions
- **Persona Consistency:** Measures the response generated is in accordance with the persona information of the speaker provided in the form of texts and is also coherent with the conversational history

Firdaus et al. I enjoy writing and playing, do you: A Personalized and Emotion Grounded Dialogue Agent using Generative Adversarial Network. In IEEE Transactions on Affective Computing 2022

05

Politeness

Problem Definition

To **transform** a generic chatbot response into a response which uses courteous phrases and emoticons to display appreciation, empathy, apology, assurance, in coherence with the state of conversation

Domain: Customer Care on Twitter

Motivation

To transform a generic chatbot reply into one that:

- Is emotionally aware and intelligent
- Uses courteous phrases and emoticons to display appreciation, empathy, apology, assurance
- End motive is to increase user satisfaction and to build customer relations

Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

Example I (Expressing Apology / Empathy)

somebody from @VerizonSupport please help meeeeee 😩😩😩😩 I'm having the worst luck with your customer service

@115719 How can we help?

@VerizonSupport I finally got someone that helped me, thanks!

@115719 Awesome!

somebody from @VerizonSupport please help meeeeee 😩😩😩😩 I'm having the worst luck with your customer service

@115719 **Help has arrived! We are sorry to see that you are having trouble.** How can we help?

@VerizonSupport I finally got someone that helped me, thanks!

@115719 Awesome! **If you ever need us we are just a tweet away.**

Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

Resource Creation: Data Source and Description

- Source the requisite Twitter data from the dataset made available on Kaggle

- Segment the tweet into sentences
  - Purely courteous (and non-informative) sentences must be removed
  - Purely informative sentences must be retained
  - Informative sentences with courteous expressions must be transformed (to remove only the courteous part from the sentence)

Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

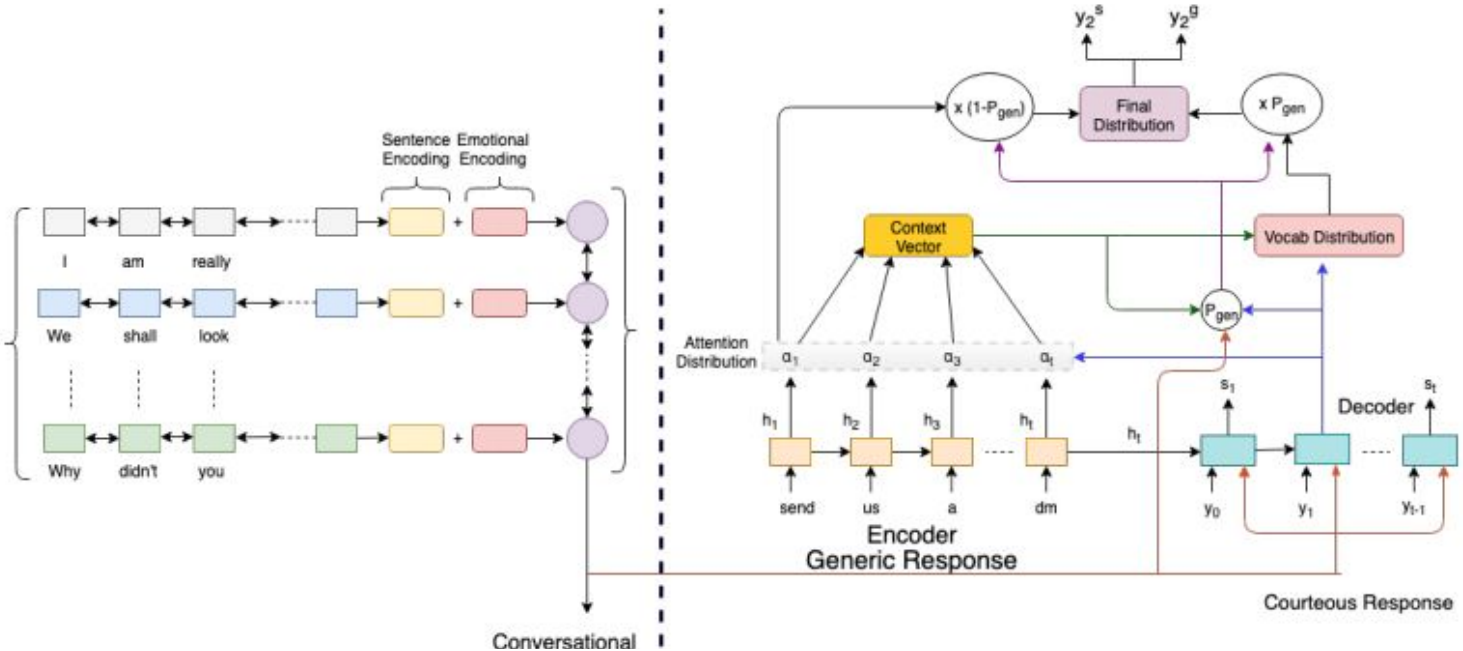
Resource Creation: Scaling up for large data creation

- **Clustering**
  - The vector-semantic representations of sentences are obtained using the sentence encoder trained on the SNLI corpus.
  - Use the K-Means clustering(k = 300) to cluster these sentences.
- **Annotations**
  - purely courteous,
  - purely informative,
  - hybrid
- **Preparing generic responses**
  - Obtain the generic response by removing the courteous sentences, retaining the informative sentences, and replacing the hybrid sentences with the prepared generic equivalents

Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

Proposed Methodology

- Based on a reinforced pointer-generator model for the sequence to sequence task
- The model is also conditioned on a hierarchically encoded and emotionally aware conversational context



Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU and ROUGE-L:** Measure the ability of the generated response for capturing the correct information
- **Content Preservation:** Measures how much of the informative content from the original generic response is reflected in the generated courteous response
- **Emotional accuracy:** Measures the consonance between the generated courteous expressions (source of emotion) and the gold

Human Evaluation

- **Fluency:** The courteous response is grammatically correct and is free of any errors
- **Content Adequacy:** The generated response contains the information present in the generic form of the response and there is no loss of information while adding the courteous part to the responses
- **Courtesy Appropriateness:** The courtesy part added to the generic responses is in accordance to the conversation history

Golchha et al. Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In NAACL-HLT 2019

Problem Definition

To **induce** courteous behaviour in generic customer care response (appreciation, empathy, apology, assurance, etc.) in a multi-lingual scenario (Hindi and English languages)

Domain: Customer Care on Twitter

Motivation

- Polite behavior of the agent give humanly essence to the conversational systems
- Develop systems that can converse with humans in their preferred language
    - using polite/courteous response,
    - leading to user satisfaction and high customer retention

Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

Example of Polite Response

Generic Response	Polite Response	Behavior
Provide your booking info via dm.	We're here to help you, please provide your booking info via dm.	*Assurance*
हम मामले पर गौर करेंगे। (We will look into the matter.)	यह सुनकर निराशा हुई, कृपया तब तक धैर्य रखें जब तक हम इस मामले पर गौर न करें। (That's disappointing to hear, please have patience until we look into the matter.)	*Empathy*

Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

Resource Creation: Data Source and Description

- Use CYCCD dataset in English [10]
- Prepared Hindi Conversational Data
  - Source the requisite Twitter data from the dataset made available on Kaggle in Hindi
  - Segment the tweet into sentences
    - Purely courteous (and non-informative) sentences must be removed
    - Purely informative sentences must be retained
    - Informative sentences with courteous expressions must be transformed (to remove only the courteous part from the sentence)

Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

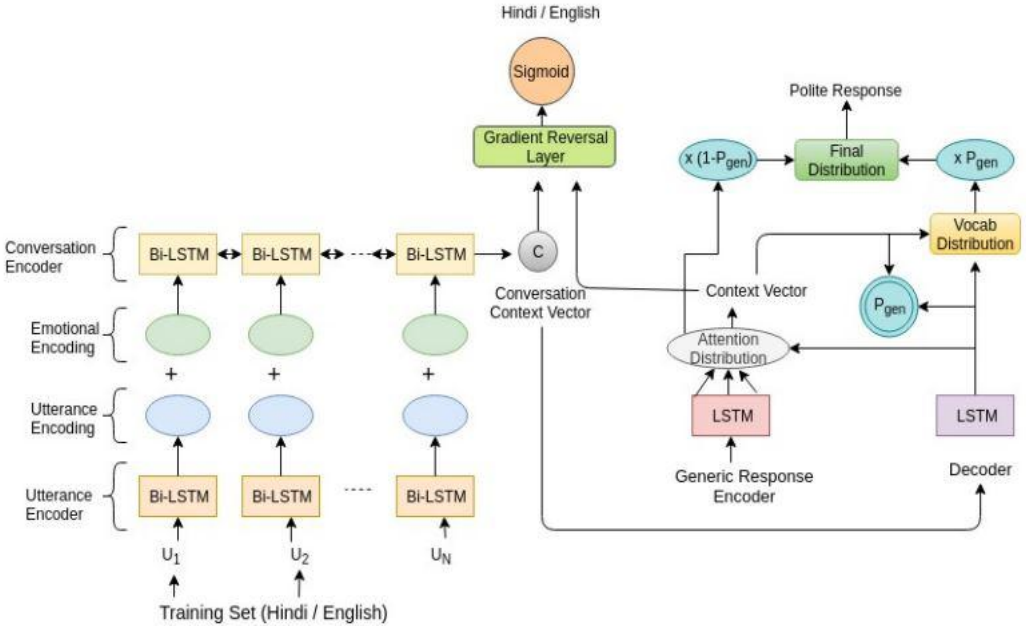
Resource Creation: Scaling up for large data creation

- **Clustering**
  - The vector-semantic representations of sentences are obtained using the sentence encoder trained on the SNLI corpus.
  - Use the K-Means clustering(k = 300) to cluster these sentences.
- **Annotations**
  - purely courteous,
  - purely informative,
  - hybrid
- **Preparing generic responses**
  - Obtain the generic response by removing the courteous sentences, retaining the informative sentences, and replacing the hybrid sentences with the prepared generic equivalents

Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

Proposed Methodology

- Based on a reinforced pointer-generator model for the sequence to sequence task
- The model is also conditioned on a hierarchically encoded and emotionally aware conversational context
- Model is jointly training of Hindi and English
- Gradient reversal layer is used to learn language invariant features



Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU and ROUGE-L:** Measure the ability of the generated response for capturing the correct information
- **Emotional Accuracy:** Ensures that the emotional states of the generated courteous behavior is consistent with the gold
- **Content Preservation:** Measures how much of the informative content from the original generic response is reflected in the generated courteous response

Human Evaluation

- **Fluency:** The courteous response is grammatically correct and is free of any errors
- **Content Adequacy:** The generated response contains the information present in the generic form of the response and there is no loss of information while adding the courteous part to the responses
- **Politeness Consistency:** Evaluates whether politeness added to the generic responses is consistent with the history of the conversation

Firdaus et al. Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In LREC 2020.

Problem Definition

- Identify the sentiments from the user utterances, and
- Using the sentiment information to transform the generic customer care responses into polite responses which are contextually appropriate to the dialog history and the user sentiments.

Motivation

- The usage of the user feedback in the form of sentiments is crucial to get contextually correct polite responses
- If the user has a negative sentiment towards the customer care system
    - Possible polite response should be towards apology, assurance, and empathy rather than greet or appreciation.

Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Examples of polite responses in accordance to the user sentiments

Dialog Context	User Sentiment	Generic Response	Polite Response	Polite Behaviour
Hey, i got food poisoning from your inflight meal on sunday	Negative	Send us a dm	That's disappointing to hear, we are sorry please send us a dm.	Apology
I need the software update urgently, the battery lasts literally half a day	Negative	How can we help?	Don't worry, we are here for you, please say how can we help?	Assurance
Dear this new update is awesome, got great new apps!	Positive	The update has many features.	Thank you very much, please checkout the exciting features in the update.	Appreciation
Order 2 zinger box meals n got free popcorn chicken, yayyyy	Positive	Enjoy your meal.	That's nice to hear, enjoy your meal.	Acknowledge
How do i go about getting a monthly ride pass ?	Neutral	We have send the link	Hello, good morning we have send the link.	Greet

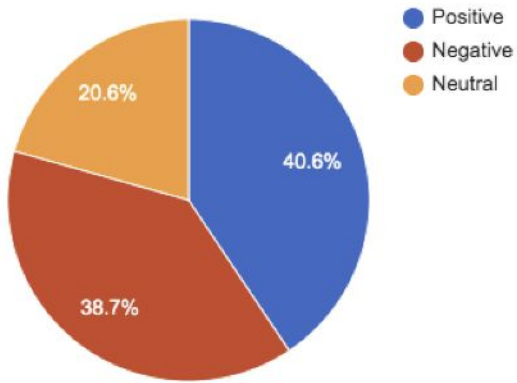
Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Dataset

- Utilized the CYCCD dataset []
- CYCCD sentiment labels
    - Positive, Negative, Neutral

	Train	Valid	Test
# Conversations	130898	19762	39665
# Utterances	168534	24724	49788

CYCCD Dataset Statistics

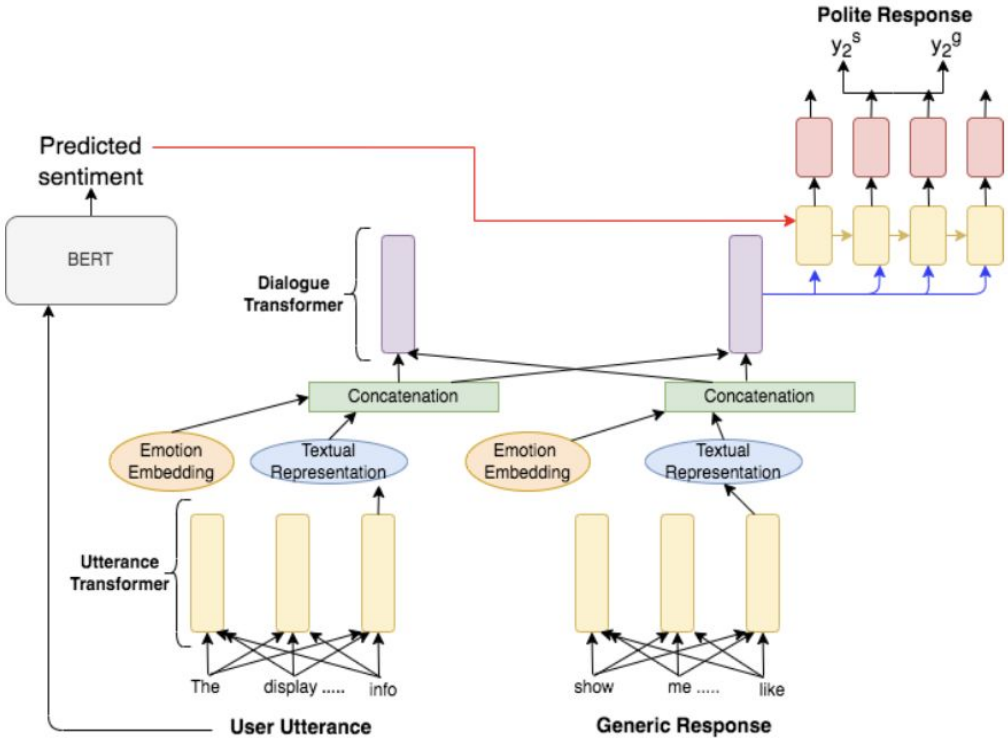


Sentiment distribution in the CYCCD dataset

Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Proposed Approach

- Transformer Encoder-Decoder (TED) architecture
- Utilized a hierarchical transformer having two encoders
    - One is to encode the sentences named as sentence encoder
    - Another Transformer is used to encode the output of the sentence encoder to capture the dialog context



Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Methodology

- The predicted sentiment information along with the contextual information is used to initialize the Transformer decoder
- We design task-specific rewards to ensure that the users' sentiments and politeness are induced appropriately in the generated responses
  - ***BLEU metric***: Ensures the content matching between the generated response and the ground-truth response
  - ***Sentiment consistency***: Measured by the cosine similarity of the sentiment prediction distributions of the user utterance and generated responses
  - ***Politeness accuracy***: The politeness accuracy of the generated response is computed using the pre-trained BERT based politeness classifier
- We jointly train the entire model by simultaneously minimizing the sentiment classification loss and generation loss

Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Evaluation Metrics

Automatic Evaluation

- **Perplexity*:*** Evaluate the generated responses at content level
- **BLEU-4 and ROUGE-L*:*** Measure the ability of the generated response for capturing the correct information
- **Politeness accuracy**: Measures the degree of politeness in the generated responses.

Human Evaluation

- **Fluency*:*** The courteous response is grammatically correct and is free of any errors.
- **Relevance*:*** Evaluate if the generated response is contextually relevant
- **Politeness Appropriateness:** Measures whether the politeness induced in the response is in accordance with the user sentiment and the dialogue history

Firdaus et al. PoliSe: Reinforcing Politeness using User Sentiment for Customer Care Response Generation. In COLING 2022

Problem Definition

- Generate polite responses by varying the degree of politeness in dialogues while considering the user's persona information

Motivation

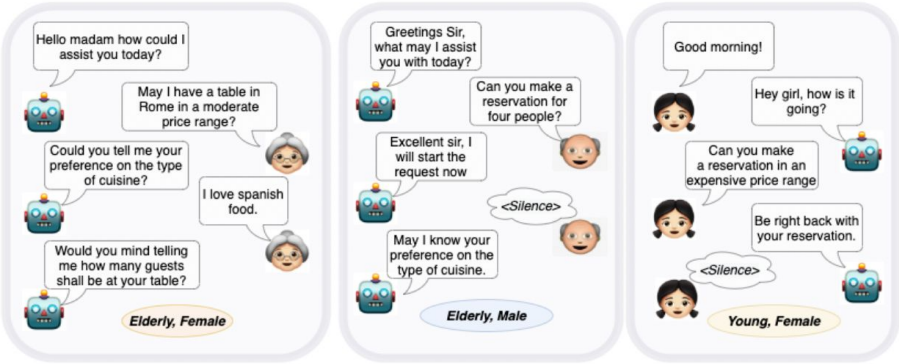
- Politeness in itself has different facets that is difficult to inculcate in a conversational agent
- Politeness varies among the different age groups. While communicating with elders, humans are found to be more polite than their conversations with younger people
- Also, the degree of communication and politeness vary when we communicate with people of different genders

Firdaus et al. Being Polite: Modeling Politeness Variation in a Personalized Dialogue Agent. In IEEE TCSS 2022.

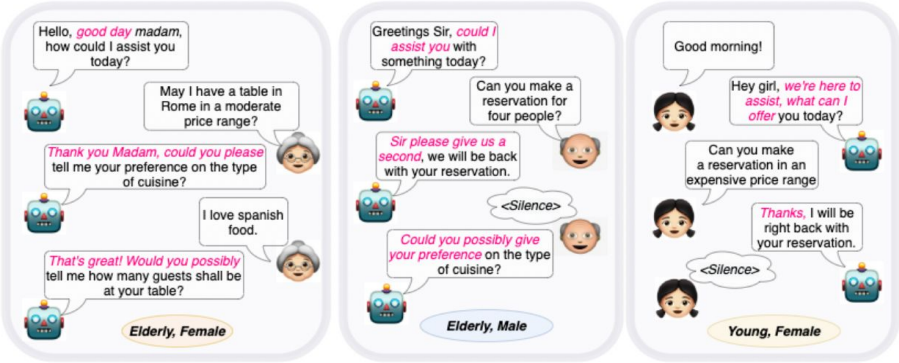
Example of Polite Personalized Conversation

The personalized dialogue conversation shows the difference in language styles with the change in users' information such as age and gender.

For example, communication with females tends to comprise more appreciation, making the conversation more polite than a male person.



Personalized Dialogue Conversations



Polite Personalized Dialogue Conversations

Firdaus et al. Being Polite: Modeling Politeness Variation in a Personalized Dialogue Agent. In IEEE TCSS 2022.

Dataset

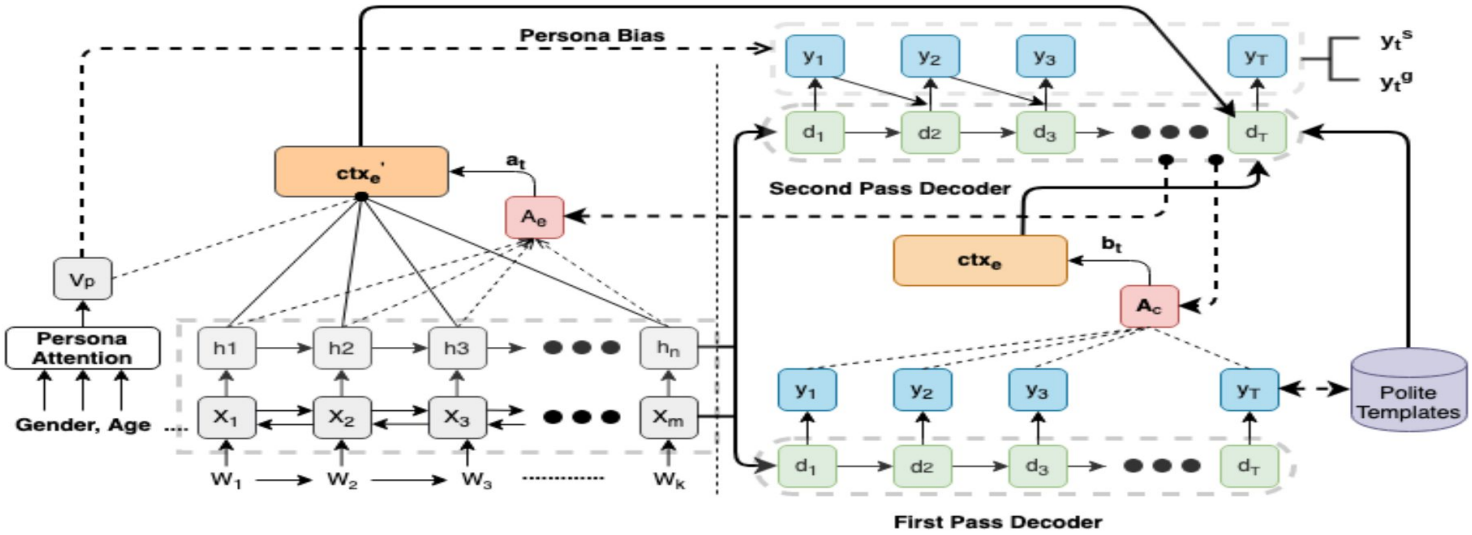
- bAbI dataset [9]
  - A multi-turn personalized dialogues
  - Use personalized features in the language style (such as Sir, Madam, etc.) to respond to the different users
- To provide politeness according to the user profiles, create polite templates in accordance with the dialogs

	Female	Male
Elderly	Thank you for your time, we are glad to help. I will finish your reservation.	Excellent sir, we enjoyed helping! I will finalize your request.
Middle-Aged	Thank you, let me complete the reservation.	Great! I will finalize the request.
Young	We are happy to help, your reservation is done.	Cool! It's done.

Firdaus et al. Being Polite: Modeling Politeness Variation in a Personalized Dialogue Agent. In IEEE TCSS 2022.

Polite Personalized Dialogue Generation Framework

- Encode the given persona information (i.e., the age, gender, etc.) and the polite templates for inducing politeness in accordance with the persona information
- Use attention over the persona information to selectively pass only those information that is important to generate polite responses for a given dialogue



Firdaus et al. Being Polite: Modeling Politeness Variation in a Personalized Dialogue Agent. In IEEE TCSS 2022.

Evaluation Metrics

Automatic Evaluation

- **Perplexity:** Evaluate the generated responses at content level
- **BLEU and ROUGE-L:** Measure the ability of the generated response for capturing the correct information
- **Politeness Accuracy:** Measure the degree of politeness in the generated responses

Human Evaluation

- **Fluency*:*** The courteous response is grammatically correct and is free of any errors.
- **Relevance*:*** Evaluate if the generated response is contextually relevant
- **Politeness Appropriateness:** Measures whether the politeness induced in the response is in accordance with the user sentiment and the dialogue history

Firdaus et al. Being Polite: Modeling Politeness Variation in a Personalized Dialogue Agent. In IEEE TCSS 2022.

Problem Definition

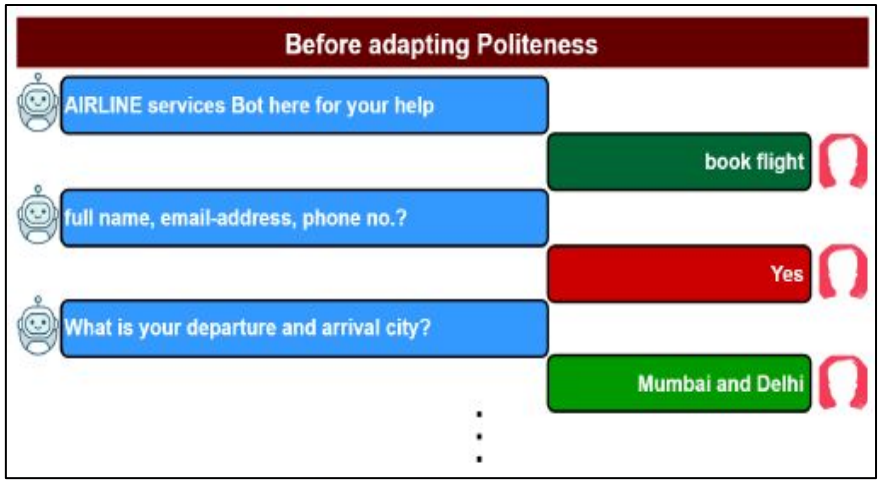
- Incorporate politeness in an end-to-end learning framework based on the agent's last action and the user's current response politeness feedbacks
- Four reinforcement learning based rewards have been designed to ensure that the conversation adapts to the different degrees of politeness

Motivation

- Lacking the ability to learn from user interactions to improve and adapt to the demands of the user
- The agent requires to pacify the aggrieved users for products in a polite manner
- Reinforcement learning ability to focus on long-term rewards have enhanced the performance of dialogue agents

Mishra et al. Please be Polite: Towards building a Politeness Adaptive Dialogue System for Goal-oriented Conversations. In Neurocomputing 2022.

An Example Conversation



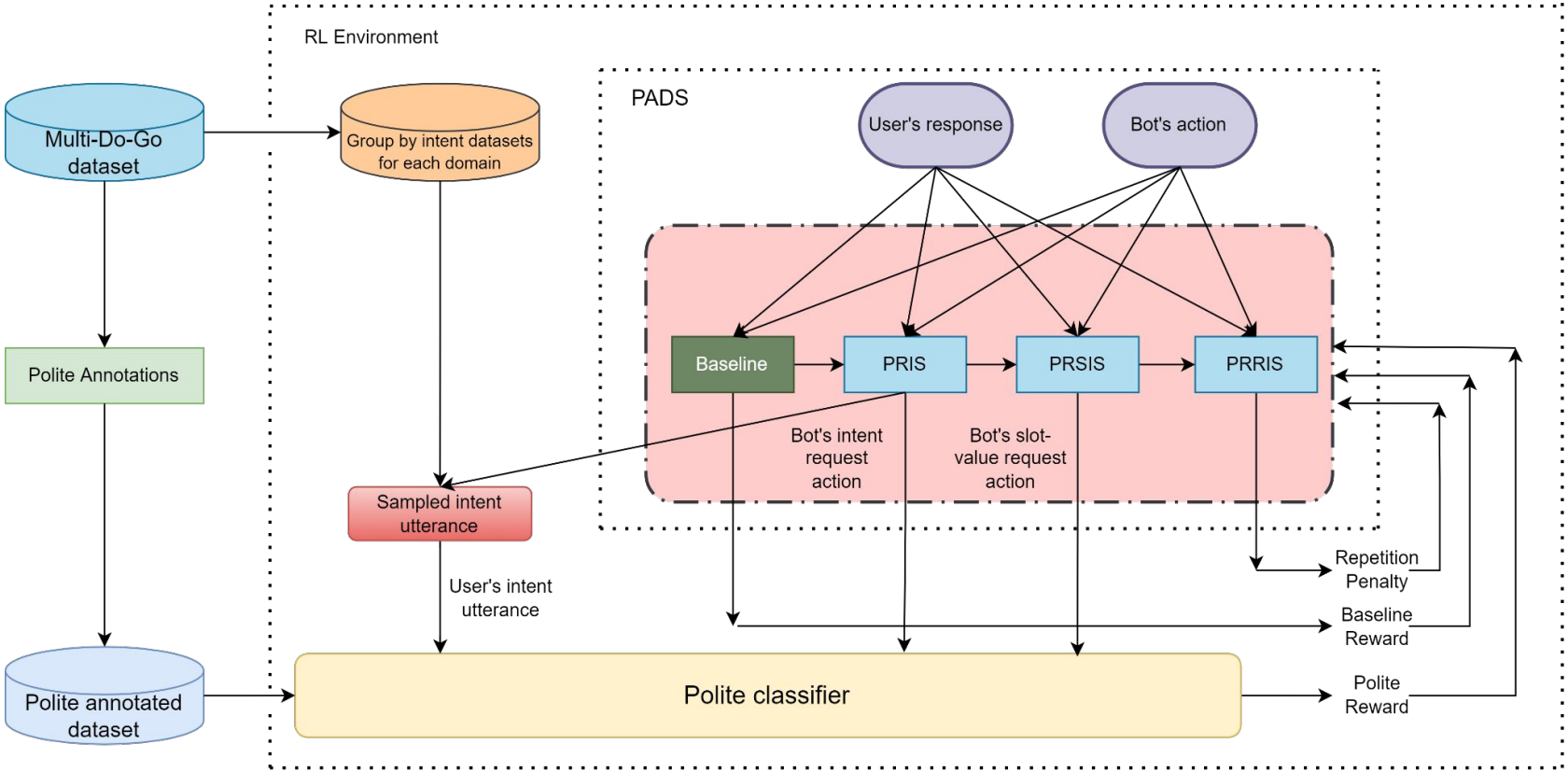
Mishra et al. Please be Polite: Towards building a Politeness Adaptive Dialogue System for Goal-oriented Conversations. In Neurocomputing 2022.

Dataset

- Annotated MultiDoGo [4] dataset with politeness labels
- **Domain**
  - *airline, fastfood, finance, insurance, media, and software*
- **Politeness Labels**
  - *polite, somewhat_polite, somewhat_impolite, impolite*
- **Politeness Annotation**
  - ***Phase 1:*** The politeness score of each utterance of the MultiDoGo dataset is obtained via Stanford Politeness Classifier [5]
  - ***Phase 2:*** Every utterance is assigned one of the four fine-grained politeness classes

Mishra et al. Please be Polite: Towards building a Politeness Adaptive Dialogue System for Goal-oriented Conversations. In Neurocomputing 2022.

Politeness-adaptive Dialogue System



Mishra et al. Please be Polite: Towards building a Politeness Adaptive Dialogue System for Goal-oriented Conversations. In Neurocomputing 2022.

Evaluation

Automatic Evaluation

- **Perplexity**
- **BLEU**
- **METEOR**
- **Rogue-2 f-1 score**
- **Polite Level score**

Human Evaluation Setup

- **Politeness Adaptability (PA)**
- **Task Completion Rate (TCR)**

Mishra et al. Please be Polite: Towards building a Politeness Adaptive Dialogue System for Goal-oriented Conversations. In Neurocomputing 2022.

Problem Definition

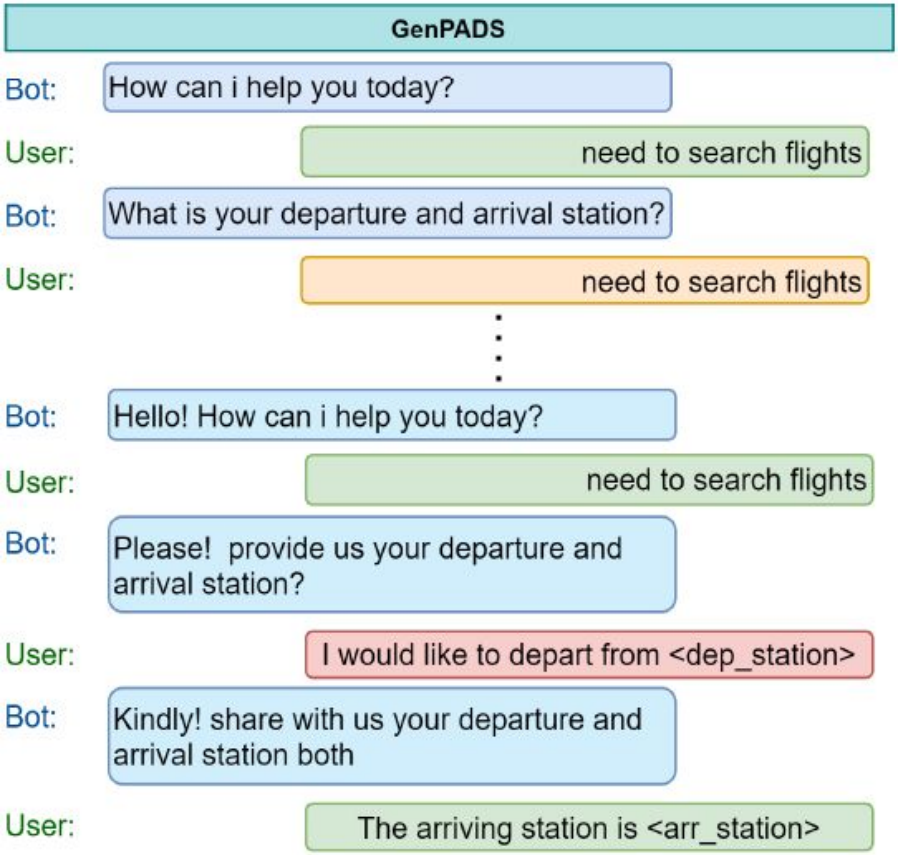
- Develop a generative politeness adaptive dialogue system
- Incorporate all three aspects, *viz.* informativeness, politeness and diversity in an end-to-end RL based learning framework

Motivation

- Address the problem of non-informative conversations or conversation drop-off in task-oriented dialogue settings
- Build a dialogue model that can learn the user's behavior online and generate polite, diverse, and interactive responses to improve the quality of the conversation
- Design the system to reinforce politeness in its dialogue generation and adapt to changes in the user's mood and demands during an ongoing conversation

An Example Conversation

- GenPADS in action for a flight domain scenario
- The user expresses dissatisfaction with the ongoing dialogue through impolite or noisy utterances (depicted by light orange boxes) or partial information (depicted by light red boxes).
- GenPADS adapts to this feedback and generates polite and diverse responses that are tailored to the user's and agent's politeness feedback



Mishra et al. GenPADS: Reinforcing Politeness in an End-to-End Dialogue System. In PLOS ONE 2023.

Dataset

- **Taskmaster**
  - Domains: *flights, food ordering, hotels, movies, music, restaurant search, and sports*
- **Dialogue Generation Dataset (DG-Dataset)**
  - Topics: *news, weather, and sports*
- **Politeness Labels**
  - *polite, somewhat_polite, somewhat_impolite, impolite*
- **Politeness Annotation**
  - Employ crowd-workers from Amazon Mechanical Turk (AMT) that labels every utterance with the provided set of polite labels

Mishra et al. GenPADS: Reinforcing Politeness in an End-to-End Dialogue System. In PLOS ONE 2023.

Generative politeness adaptive dialogue system (GenPADS)

- **Politeness Classifier (PC)**
  - The Politeness Classifier takes an input utterance and predicts its politeness level using a pre-trained model
  - The output of this component is used to determine the appropriate response generation strategy
- **Dialogue Generator (DG)**
  - The Dialogue Generator is a sequence-to-sequence model that generates responses given an input utterance.
  - It is trained on human-human dialogues from the DG-Dataset.
- **GenPADS Generation Module (G)**
  - Combines the outputs of the Politeness Classifier and Dialogue Generator to generate a polite and diverse response tailored to the user's and agent's politeness feedback
  - Uses a reinforcement learning approach to learn from user feedback during an ongoing dialogue

Mishra et al. GenPADS: Reinforcing Politeness in an End-to-End Dialogue System. In PLOS ONE 2023.

Evaluation

Automatic Evaluation

- **Politeness classifier's (PC)**
  - **F1-score**
- **Dialogue Generator (DG) and GenPADS Generation Module (G)**
  - **Perplexity, BLEU and NIST**
  - **Success Rate or Task Completion Rate**
  - **Dialogue Length:** Average number of turns needed to complete a task
  - **Average politeness score:** Judge agent's adaptation towards polite actions
  - **Average meteor score:** Measure the semantic similarity
  - **Average rogue-2 f-1 score:** Measure the diversity of the generated response.

Human Evaluation Setup

- **Fluency:** The generated response is grammatically correct and is free of any errors
- **Informativeness:** Measure how well the response addresses the user's query or request.
- **Politeness Adaptability:** Measure how well the response adapts to different levels of politeness exhibited by users
- **Diversity:** Measures how varied and creative the generated responses are.

Mishra et al. GenPADS: Reinforcing Politeness in an End-to-End Dialogue System. In PLOS ONE 2023.

06

External Knowledge

Problem Definition

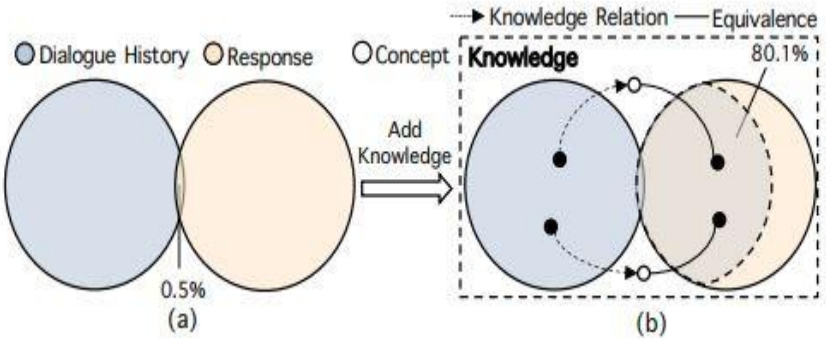
- To improve empathetic dialogue generation by
    - leveraging external knowledge, including commonsense knowledge and emotional lexical knowledge
- Explicitly understand and express emotions

Motivation

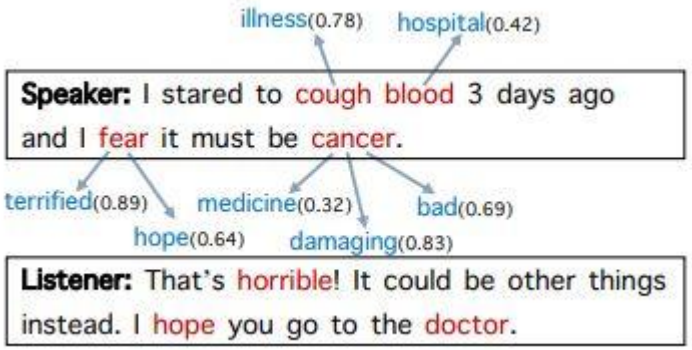
- Humans usually rely on experience and external knowledge to acknowledge and express implicit emotions
- Lack of external knowledge makes empathetic dialogue systems difficult to perceive implicit emotions and learn emotional interactions from limited dialogue history
- Leverage external knowledge, including commonsense knowledge and emotional lexical knowledge, to explicitly understand and express emotions in empathetic dialogue generation
- Enhance emotional interactions between humans and machines

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

Dialogue History, Responses, and External Knowledge in Empathetic Dialogue Generation



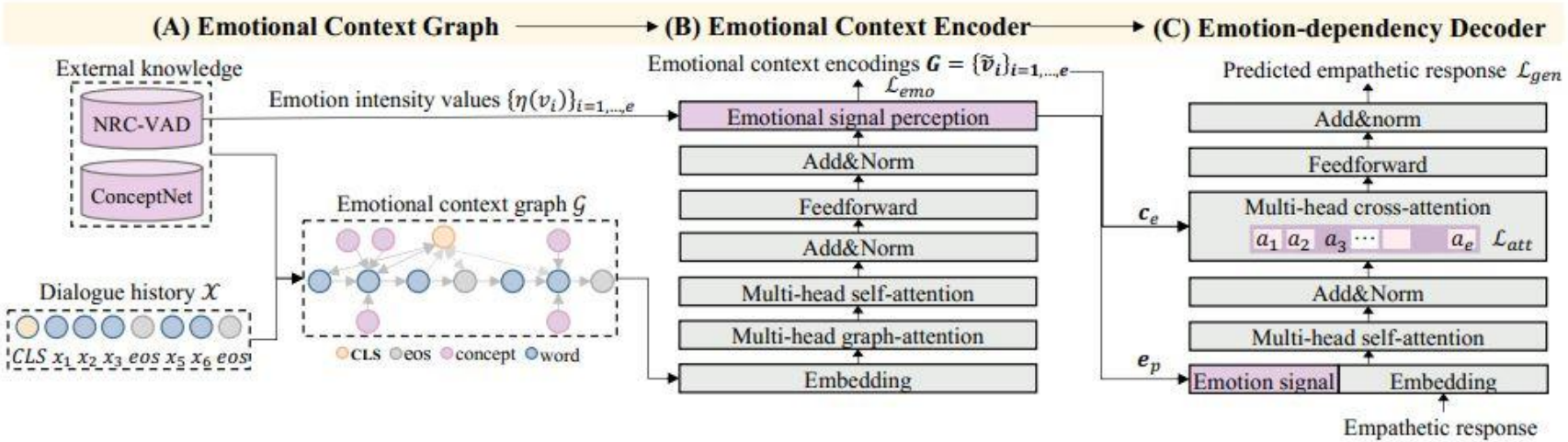
Relationships among dialogue history, responses and external knowledge.



An example of empathetic dialogues with external knowledge from EMPATHETICDIALOGUES.

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

Knowledge-aware EMPathetic dialogue generation (KEMP)



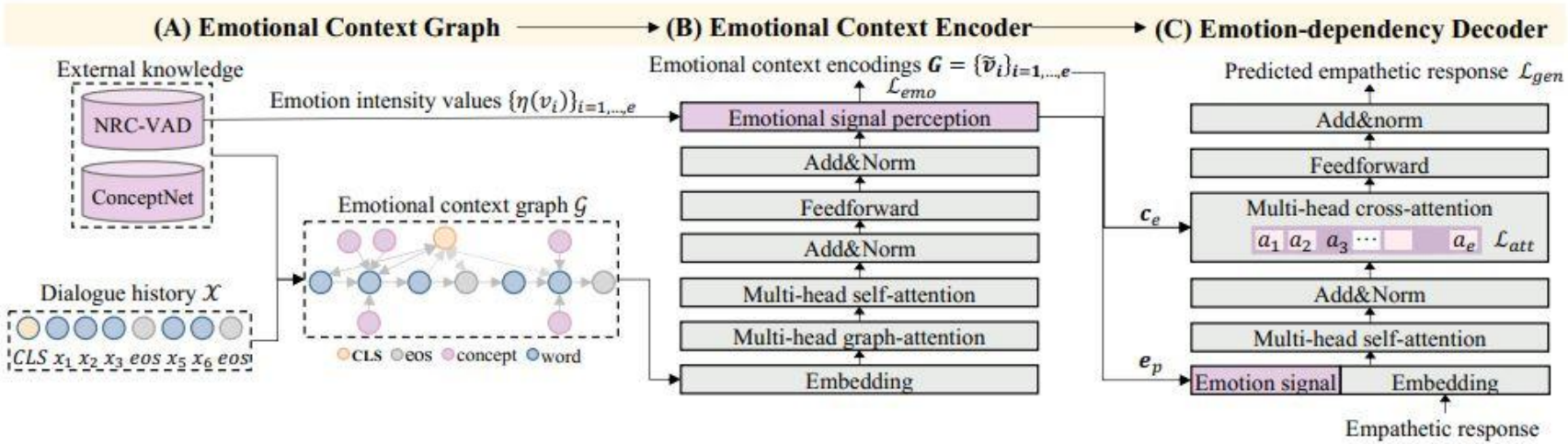
Emotional context graph: Constructed via integrating the dialogue history with external knowledge

Emotional context encoder: Employs the graph-aware transformer to learn the graph embeddings, and propose an emotional signal perception procedure to perceive context emotions that lead the response generation.

Emotion-dependency decoder: Conditioned on the knowledge-enriched context graph, it particularly models emotion dependencies to generate empathetic response

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

Knowledge-aware EMPathetic dialogue generation (KEMP)



(A) Constructed via integrating the dialogue history with external knowledge

(B) Employs the graph-aware transformer to learn the graph embeddings, and propose an emotional signal perception procedure to perceive context emotions that lead the response generation

(C) Conditioned on the knowledge-enriched context graph, it particularly models emotion dependencies to generate empathetic response

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

Dataset

- EMPATHETICDIALOGUES dataset [2]
  - 25k one-to-one open-domain empathetic conversations
  - 32 evenly distributed emotion labels
  - 17,802 dialogues in the training set, 2,628 in the validation set, and 2,494 in the testing set

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

Evaluation

Automatic Evaluation

- **Perplexity**
- **Distinct-1 and Distinct-2**
- **Emotion Accuracy:** Measures the agreement between the ground truth and predicted emotion labels

Human Evaluation Setup

- **Empathy:** Measures whether the generated responses express the appropriate emotions;
- **Relevance:** Evaluates whether the responses are on-topic with the dialogue history;
- **Fluency:** Measures the grammatical correctness and readability of the generated responses

Li et al. Knowledge Bridging for Empathetic Dialogue Generation In AAAI 2022.

07

Multimodal Information

Problem Definition

- Generate sentiment and emotion controlled textual responses conditioned on the conversational history
- The dialogue consists of text utterances along with audio and visual counterparts, and given a context of k turns the task here is to generate the next text response
- For the given task, emotion and sentiment categories will be provided to generate the response

Motivation

- Simultaneous use of sentiment and emotion information is useful to generate more human-like responses
- Leads to better user experience and retention
- Multimodal information (audio and video) provide important cues for correctly identifying sentiment and emotion

Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Sentiment and Emotion aware Multi-modal Dialogue (SEMD) Dataset

- Large-scale multi-party dataset that seamlessly employs multimodal information along with sentiment and emotion in the dialogues.
- Dataset was created utilizing the 10 famous TV shows belonging to different genres:
    - ***Comedy:*** *Friends, The Big Bang Theory, How I Met Your Mother, The Office*;
    - ***Drama:*** *House M.D., Grey's Anatomy, Castle and Game of Thrones, House of Cards, Breaking Bad*
- Total 55k dialogues
- Emotion labels:
    - Ekman's six universal emotions: J*oy, Sadness, Anger, Fear, Surprise, and Disgust*
    - *Extended Emotion annotation list: Acceptance and Neutral*
- Sentiment labels: Positive, Negative and Neutral

Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

An Example from SEMD dataset

You know, we had all this cool stuff in basement.
(Surprise, Positive)

No no, I am paddling away.
(Disgust, Positive)

See, Yeah
(Joy, Positive)

Really, you got all this rustic crap for free.
(Anger, Negative)

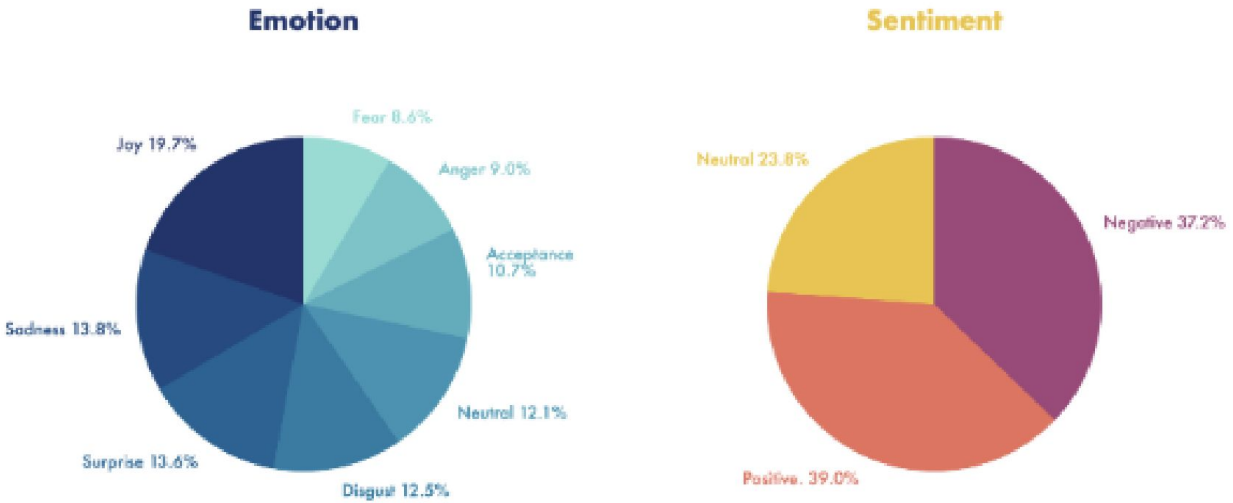
Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Data Collection

- We create our dataset from the 10 famous TV shows belonging to different genres:
  - ***Comedy:*** *Friends, The Big Bang Theory, How I Met Your Mother, The Office*;
  - ***Drama:*** *House M.D., Grey's Anatomy, Castle and Game of Thrones, House of Cards, Breaking Bad*
- The dataset contains dialogues mostly from all the episodes belonging to the different seasons of the TV series, giving us a *wide variation in conversations*.
- In total, there are 1258 episodes, spanning 746 hours.
- First, we extract all the subtitles and transcripts for every episode.

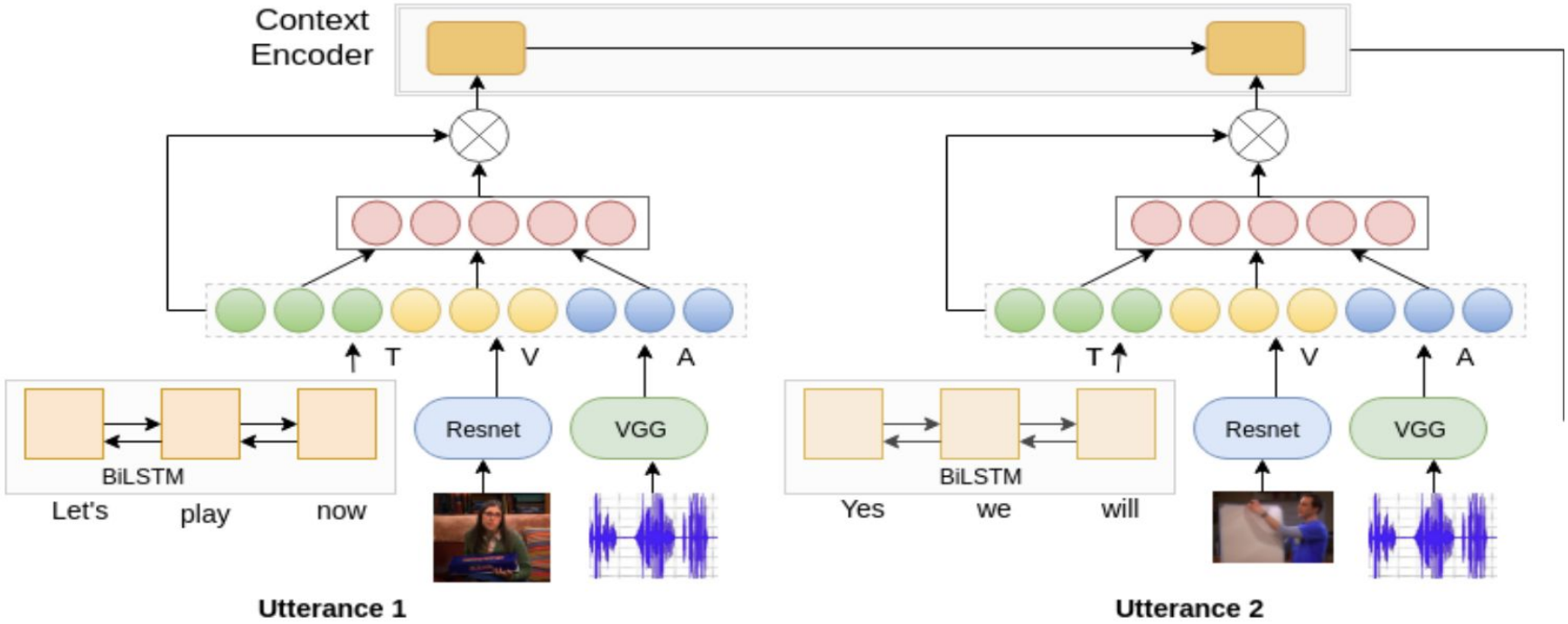
Data Annotation

- We create a balanced dataset (SEMD-annotated) by manually annotating all the 10 TV series.
- For annotating the dataset, we consider Ekman's six universal emotions, viz. *Joy, Sadness, Anger, Fear, Surprise, and Disgust* as emotion labels for all the utterances in dialogue. The annotation list has been extended to incorporate two more emotion labels, namely *acceptance and neutral*.
- We label every utterance in a dialogue with sentiment labels (*positive, negative, and neutral*).



Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Multimodal Hierarchical encoder with Attention

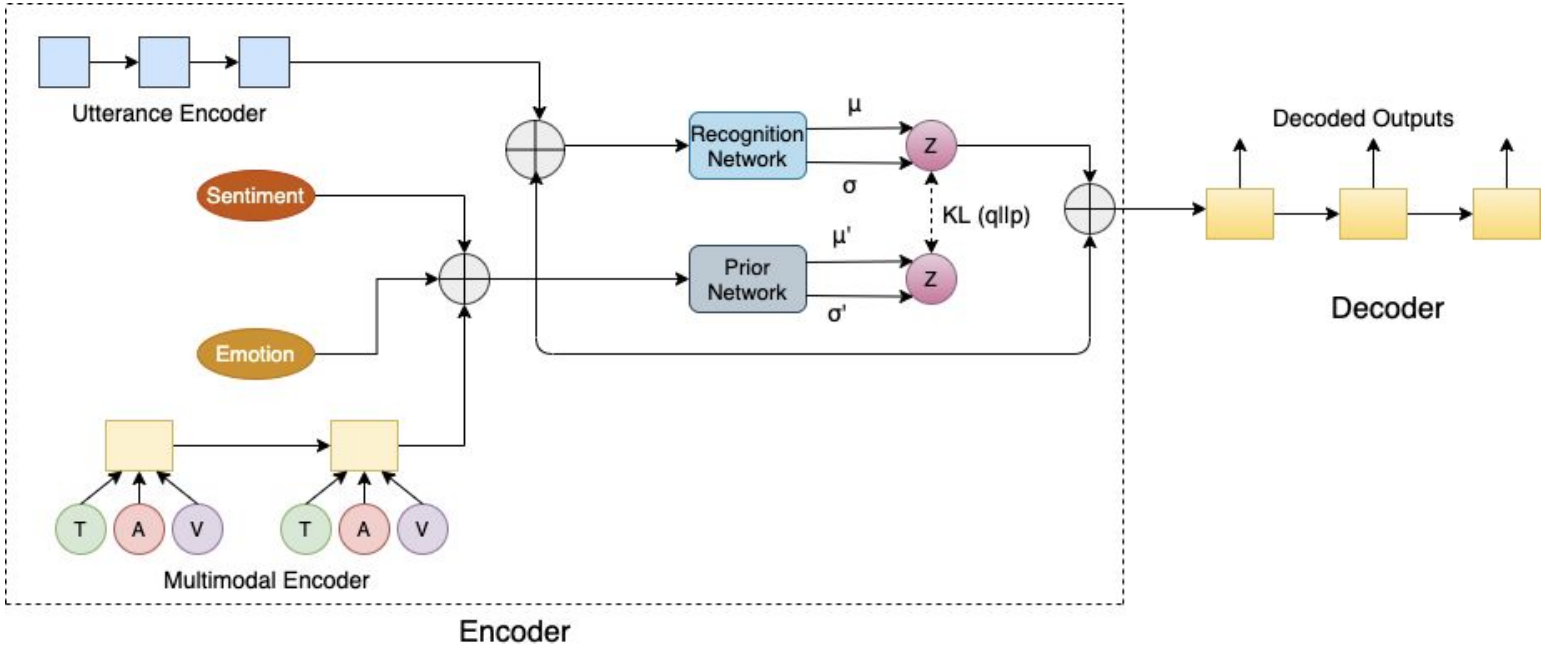


The attended utterance representation(with features from all the three modality) is passed to the context encoder

Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Multimodal Conditional Variational Autoencoder (M-CVAE)

- In M-CVAE, dialog response *y* is generated conditioned on dialog context h_c along with the desired emotion V_e and sentiment V_s embedding and latent variable *z*.



Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Evaluation Metrics

Automatic Evaluation

- **Perplexity**: Evaluate the generated responses at content level
- **Distinct-1 and Distinct-2:** Measure the degree of diversity by calculating the no. of distinct unigrams and bigrams, respectively in the generated responses
- **Sentiment Accuracy:** Evaluates the sentiment information in the generated responses.
- **Emotion Accuracy:** Evaluates the sentiment information in the generated responses.

Human Evaluation

- **Fluency:** The courteous response is grammatically correct and is free of any errors.
- **Emotion:** Evaluate whether the emotional category of the generated response is consistent with the user-specified emotion
- **Sentiment:** Evaluate whether the sentiment category of the generated response is consistent with the user-desired sentiment.

Firdaus et al. EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System. In IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2022.

Problem Definition

- Propose the task of sentiment guided aspect controlled response generation for multimodal dialogue systems

Motivation

- Growing requirements in various fields require conversational agents to communicate by incorporating information from the different modalities to build a robust system
- Users are the ultimate evaluators of dialogue systems. Therefore, research on the dialogue framework should aspire for greater user satisfaction

Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

Sentiment annotated examples from the Multi-domain Multi-modal Dialogue(MDMMD) dataset

Dialogue	Sentiment
User: I don't like the speakers shown so far, could you please show something portable.	Negative
Agent: Don't worry, we have some great portable speakers for you, please have a look. 	Neutral
User: The 2nd one looks great, could you show more in this pattern.	Positive

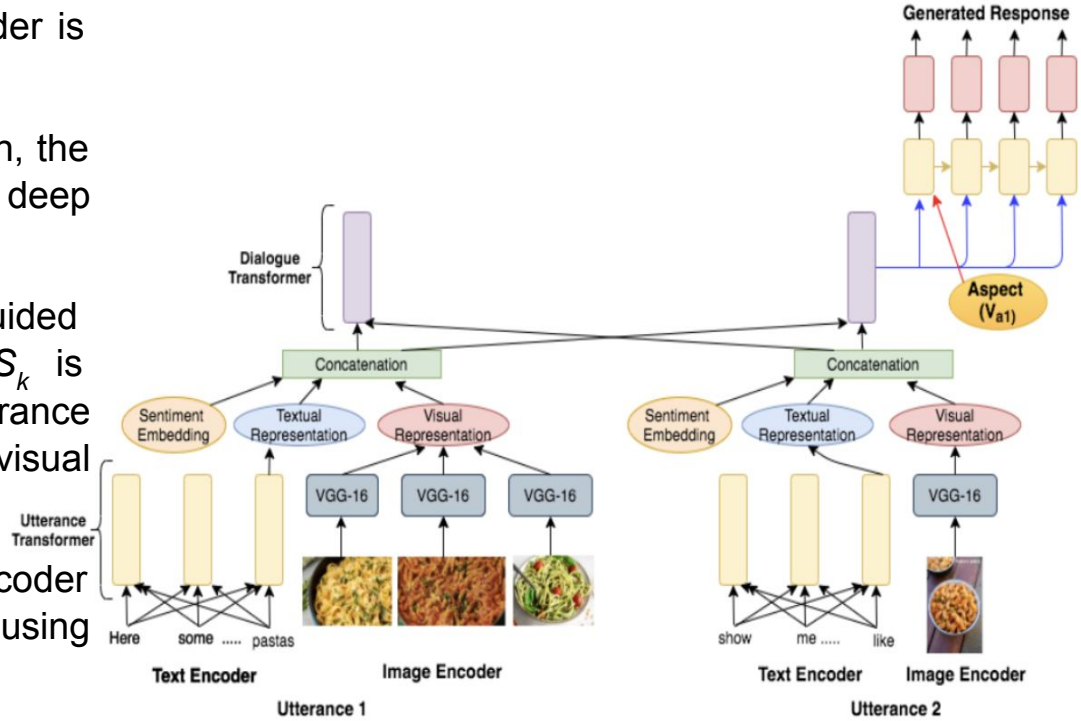
Dialogue	Sentiment
Agent: Great Choice! We have some nice colors for you, please see 	Positive
User: I love red! Please show the 3rd one from different orientations.	Positive
Agent: Nice to know, but we don't have any images of the 3rd one to show.	Neutral

- Providing extra feedback from the user in the form of sentiment
  - guide the model to adapt to user behaviour
  - assist in generating appropriate and accurate responses according to the user requirements

Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

Proposed Framework

- **Utterance Encoder:** A transformer encoder is used to encode the textual utterances
- **Visual Encoder:** For visual representation, the pre-trained VGG-16 having 16-layer deep convolutional network is used
- **Context Encoder:** For sentiment guided response generation, sentiment label S_k is concatenated with the final utterance representation having both textual and visual representation.
- **Aspect conditioned Decoder:** RNN decoder is used to construct the next textual reply using the specified aspect embedding



Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

Training and Inference

Reinforcement learning (RL) and machine learning (ML) are jointly used to train the model

- **For ML:** the maximum-likelihood objective using teacher forcing is employed
- **For RL:** The final reward function is the weighted mean of the three terms as given below:
  - ***BLEU metric:*** Ensures the content matching between the generated response and the ground-truth response to avoid loss of information
  - ***Sentiment consistency:*** Measured by the cosine similarity of the sentiment prediction distribution of the user utterance and generated responses (using pre-trained BERT classifier). It ensures that the sentiment states of the generated response is consistent with the user sentiment
  - ***Fluency:*** The above rewards do not assess if the response content expressed is linguistically fluent

Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

Dataset

This work is built upon the Multi Domain Multi Modal Dialogue (MDMMD) dataset that comprises of 130k chat sessions between the customer and sales agent

Data Annotation

- Due to the absence of sentiment labels in the MDMMD dataset, a semi-supervised approach is used for labeling it with sentiments for which we annotate a portion of the dataset
- Create a balanced dataset (MDMMD-annotated) by manually annotating 10k dialogues for all the three domains
- Label every utterance in a dialogue with three sentiment labels, *positive, negative, and neutral*

Sentiment Classifier:

- Apply a semi-supervised approach for annotating the entire MDMMD dataset with sentiment labels
- For labeling the entire MDMMD dataset, the best-performing classifier, RoBERTa is used

Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

Evaluation

Automatic Evaluation

- **Perplexity:** Evaluate the model at relevance and grammatical level
- **BLEU-4:** Measure the ability of the generated response for capturing the correct information.
- **Aspect F1:** F1 score of the requested aspects present in the generated response. Used to balance both recall and precision
-
- **Sentiment Accuracy:** Checks if the user opinion has been expressed in the generated responses

Human Evaluation

- **Fluency:** The courteous response is grammatically correct and is free of any errors.
- **Relevance:** Evaluate if the generated response is contextually relevant
- **Informativeness:** Evaluates how much information is contained in the generated responses.
- **Aspect Consistency:** Evaluate whether the response generated is in consonance to the specified aspects (e.g., cuisine, color, type, etc.)
- **Sentiment Appropriateness:** Evaluate whether the response generated is coherent to the sentiment of the user utterance and the conversational history

Firdaus et al. Sentiment guided Aspect conditioned Dialogue Generation in a Multimodal System. In ECIR 2022.

ECAI Systems for Persuasion and Therapy

(30 minutes)

Problem Definition

- Build a dialogue system which is capable of persuading the users empathetically for the task of charity donation

Motivation

- A high quality conversation is often derived by understanding and acknowledging implied feelings towards the conversing partner
- Subtle dependency between the different personalization techniques, such as empathy, sentiment, persuasion etc.
- People are more likely to engage in the conversation when they are motivated with empathetic responses

User: I am not ready to donate right now.

Bot(Without empathy): Do you reconsider for 10?

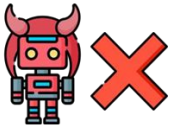
Bot (With empathy): Only a little help may save the children as a whole. Would you like to reconsider for 10?

Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Empathetic Persuasion Conversational AI

I am not ready to donate right now.

Do you reconsider for 10?



Only a little help may save the children as a whole. Would you like to reconsider for 10?

Empathetic Persuasion

Reinforcing Empathy and Persuasiveness in Dialogue Systems

Yes I think so, we are so involved in ourselves.

You are right, I know. I feel like it has become so important to me to help others and to be a part of the solution. [Agreeing, Emotional Appeal, Caring]

Consistency Persuasive Strategy Empathy

Maintaining Dialogue Consistency, Applying Persuasive strategy and generating Empathetic response

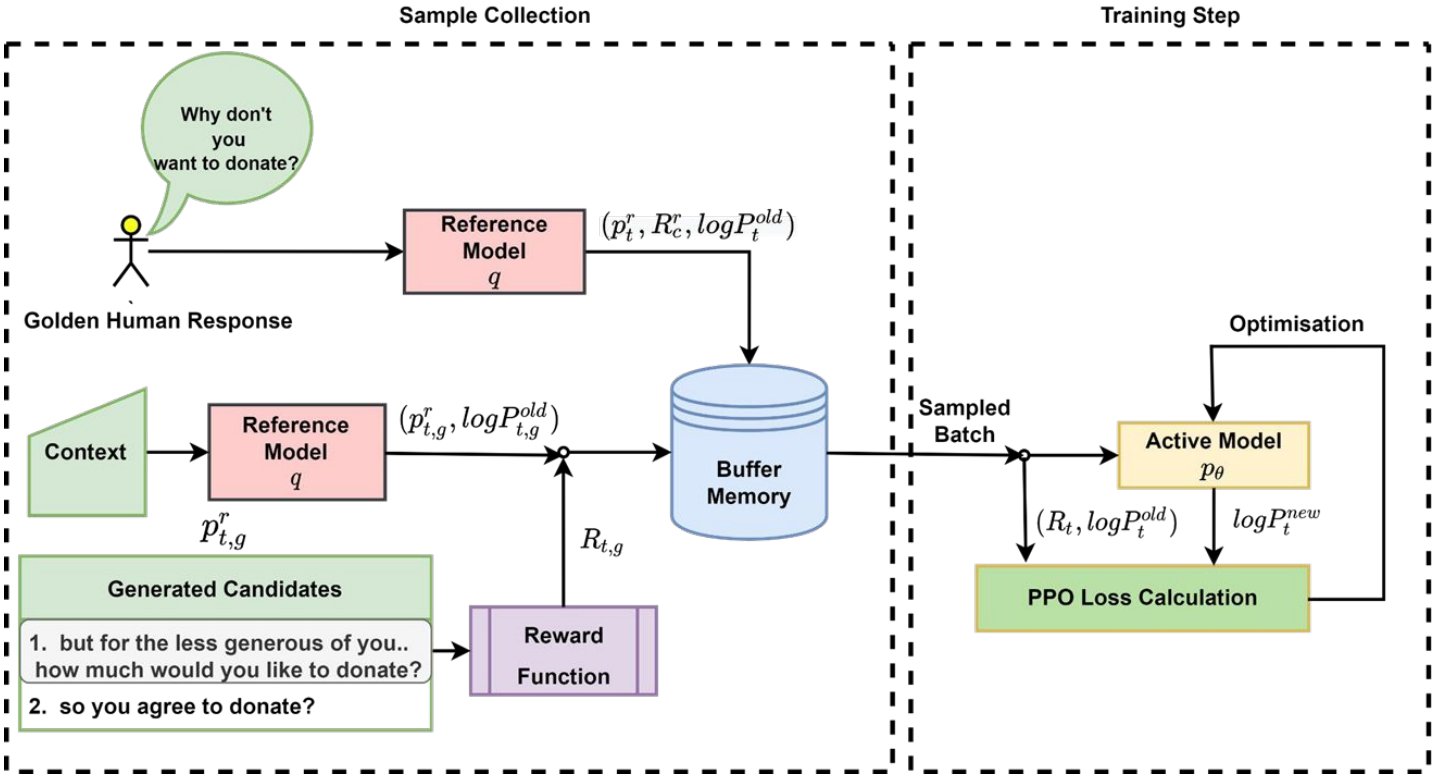
Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Dataset

- PERSUASION FOR GOOD dataset [6]
  - 1,017 persuasive conversations
    - For donation to a charity organization "*Save the Children*"
  - Between two humans
    - One acts as a *Persuader* and the other acts as a *Persuadee*
  - 11 Persuasion strategies
    - Persuader's utterances are grounded in one of the 11 persuasion strategies
  - Annotated PERSUASION FOR GOOD dataset with 23 emotion labels

Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems



Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Architecture Details

- **Two models: A Reference Model (RM) and an Active Model (AM)**
- **Reference Model (RM)**: Used for generating response candidate given a context (persuadee's utterance).
- **Reward Function:** Calculates rewards for the generated candidates.
- Gold Response and generated candidates are stored in the **buffer memory**, and sampled during the training.
- **Active Model (AM):**
    - Outputs the new log probabilities for the sampled batch **using PPO loss calculation** and finally optimisation is performed.

Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Evaluation

Automatic Evaluation

- **PerStr:** Percentage of utterances generated with persuasive strategy.
- **EmoPr:** Percentage of empathetic utterances generated.
- **PPL:** Evaluate the generated response quality.
- **LEN:** Evaluate the average number of tokens generated in an utterance

Human Evaluation Setup

- **Per**, **Emp:** Checking persuasiveness and empathy factor in the dialogue based on one-five positive integer scale.
- **DonPr:** Calculating percentage of time people donated.
- **Cons**, **Fluen** and **Rep:** Check the *consistency* (with the dialogue context), linguistic fluency and non-repetitiveness of generated utterance in the dialogue.

Samad et al. Empathetic Persuasion: Reinforcing Empathy and Persuasiveness in Dialogue Systems. In NAACL Findings 2022.

Problem Definition and Motivation

Build a polite and empathetic dialogue system for persuading the users for charity donation

I am not ready to donate right now.

Do you reconsider for 10?

Only a little help may save the children as a whole. Would you like to reconsider for 10?

Persuasive conversations:

- Influence other person's attitude or intention.
- Identified by ***cause or stimulus*** and ***attitude.***

Its characteristics:

May fail even with ***compelling arguments.***

Use of empathy and polite tone

may evoke *better connection*, *cognitive* and *emotional processing* conducive to persuasion.

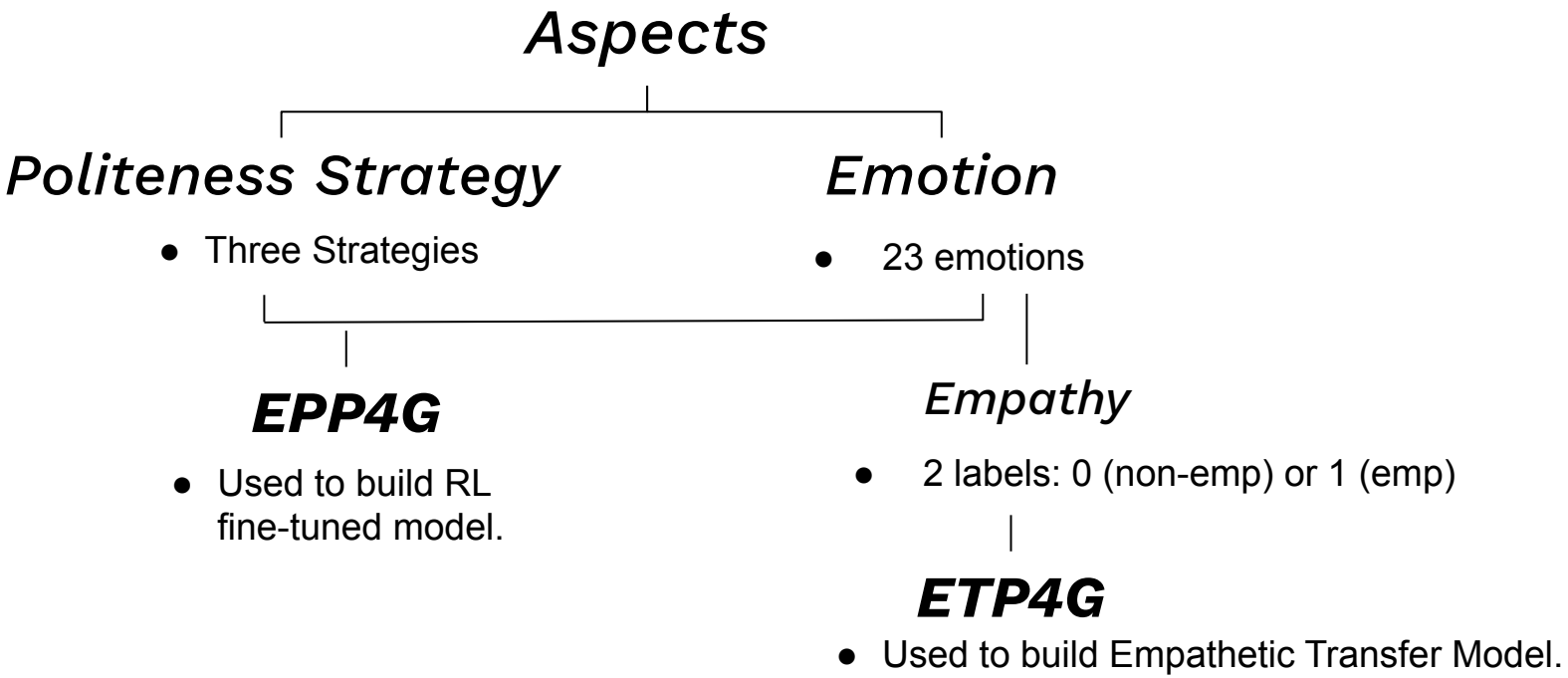
Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Dataset

- PERSUASION FOR GOOD dataset [6]
  - 1,017 persuasive conversations
    - For donation to a charity organization “*Save the Children*”
  - Between two humans
    - One acts as a *Persuader* and the other acts as a *Persuadee*
  - 11 Persuasion strategies
    - Persuader’s utterances are grounded in one of the 11 persuasion strategies

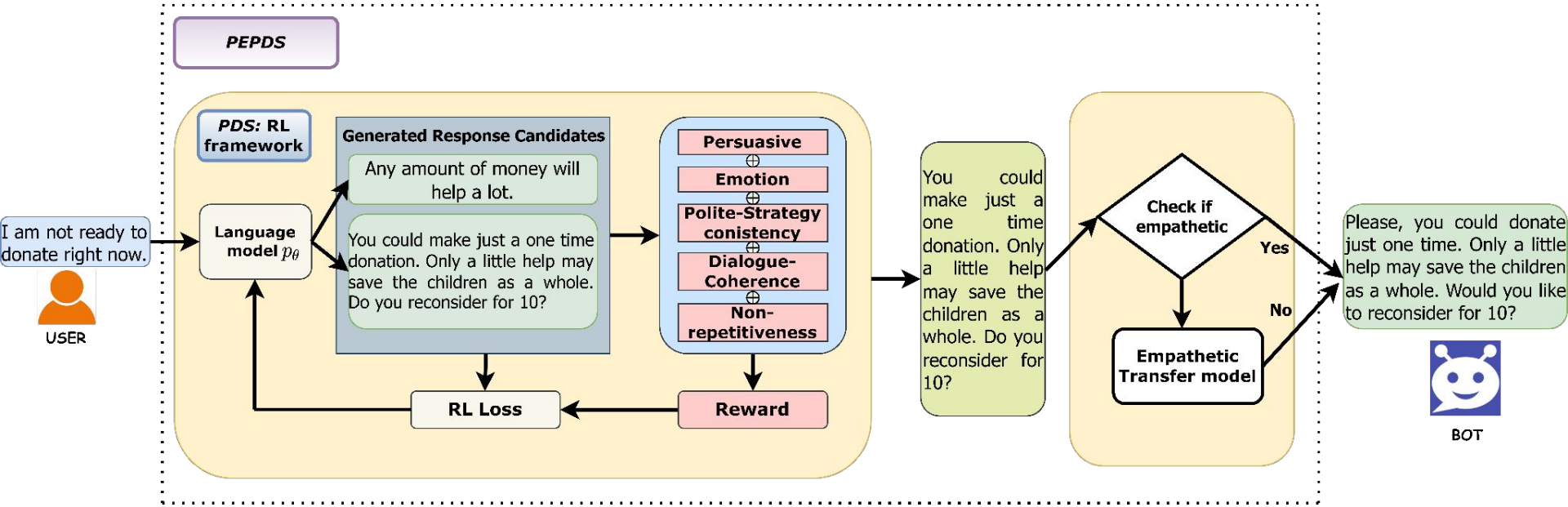
Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Dataset Annotation



Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

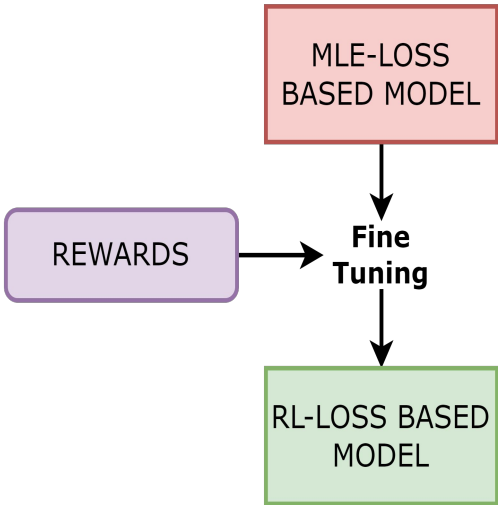
Polite and Empathetic Persuasive Dialogue System (*PEPDS*)



Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Polite and Empathetic Persuasive Dialogue System (PEPDS)

- Fine-tuned a Maximum Likelihood Estimation loss based model
  - Designed an efficient reward function consisting of sub rewards: *Task-specific rewards viz.* Persuasion, Emotion, Politeness-Strategy Consistency; *Generic rewards viz.* Dialogue-Coherence and Non-repetitiveness.
- Then, to generate empathetic utterances for non-empathetic ones, an Empathetic transfer model is built upon the RL fine-tuned model



Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Evaluation Metrics

Classifiers

- **Weighted Accuracy (W-ACC)** - measures weighted accuracy of a classifier, considering all classes
- **Macro-F1** - to account for imbalanced class distribution

Empathetic Transfer Model

- **Empathy Accuracy (EM-ACC)**
- **Perplexity (PPL)**
- **Bleu score (BLEU)**
- **METEOR score (MET)**
- **Rogue-2 F-1 score (R-2-F1) and**
- **NIST score (NIST)**

Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Evaluation Metrics

Automatic Evaluation

- **PerStr** - percentage of the utterances generated with persuasion strategy
- **PolSt** - percentage of utterances generated with consistent politeness strategy
- **Emp** - percentage of empathetic utterance generated,
- **PPL** - perplexity of the dialogue agent and
- **LEN** - number of tokens generated in an utterance

Human Evaluation

- **Per, Emp** - checking persuasiveness and empathy of the generated dialogue
- **DonPr** - computing percentage of time people donated
- **Const, Adeq, Fluen and N-Rep -** to evaluate if the generate utterances are consistent adequate, linguistically fluent and non-repetitive in nature

Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Mental Health: Need vs. Access

Access to **mental health care** is **poor** across the globe

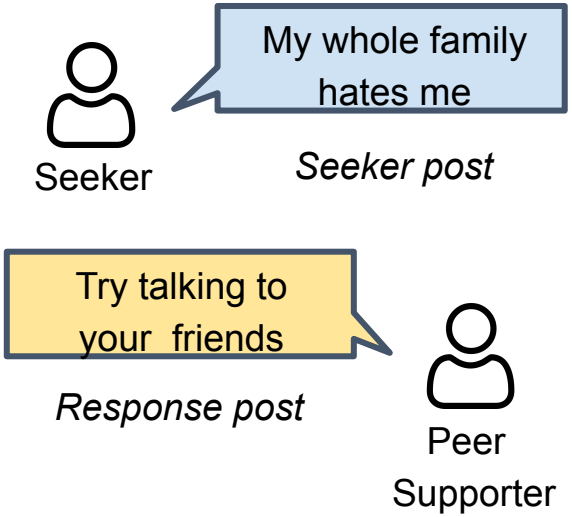
- Low-income, middle-income countries
  - **1** psychiatrist per **100k** individuals
- United States
  - **60%** of the counties do not have a single psychiatrist

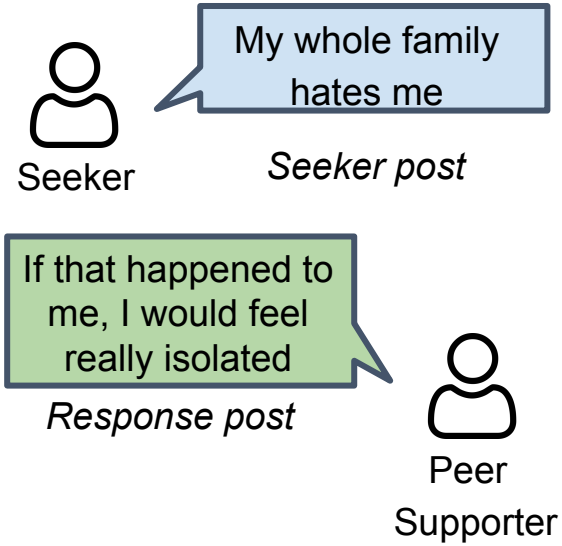
Key: We may **never** have enough mental health professionals to meet the **need**

Online peer support platforms can help!

- E.g. TalkLife, 7cups, Mental health subreddits
- Millions of users **seek** and **provide support** through **conversations**

Text-based, asynchronous conversations on peer-support platforms



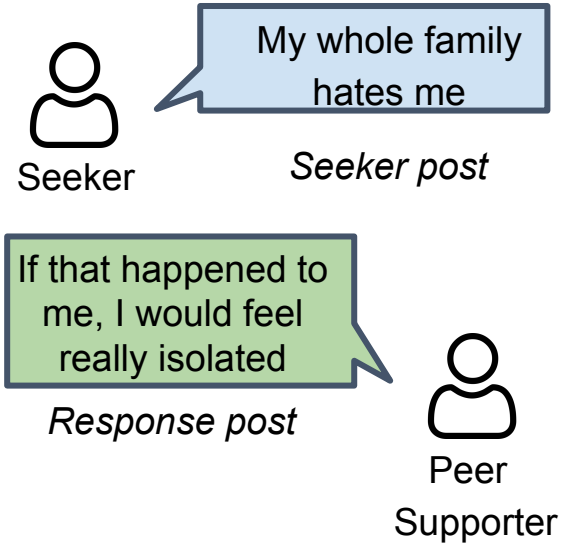


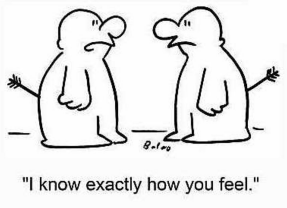
For online mental health platforms to be helpful, peer-supporters should provide effective support

Communication of Empathy in Conversation

Empathy

- **Core intervention component** of therapeutic counseling
- Quantitative evidence shows **strong associations with positive counseling outcomes** (Bohart et al., 2002; Elliot et al., 2011)
  - Symptom improvement
  - Alliance and rapport




"I know exactly how you feel."

Empathy: Current limitations and challenges

- Computational methods are **limited to face-to-face, speech- based therapy** (Gibson et al., 2016, Perez-Rosas et al., 2017)

Audio-video signals: speech, prosody

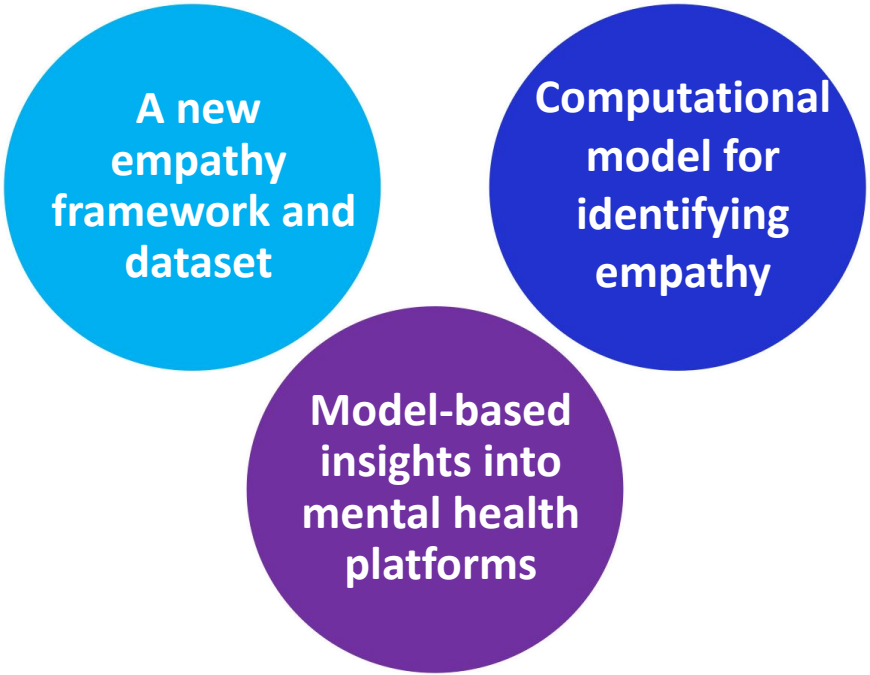
- Previous NLP research focuses on **empathy as reacting with emotions of warmth and compassion** (Buechel et al., 2018) or as **emotionally-grounded conversations** (Rashkin et al., 2019).

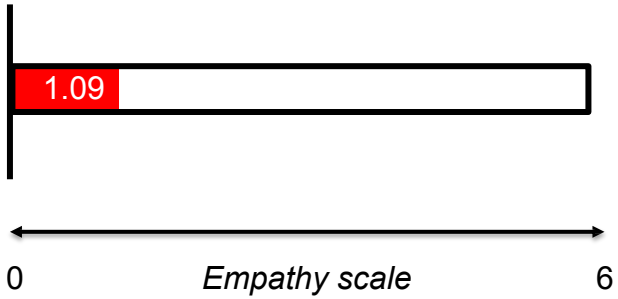
I feel sorry for you

- **Communicating cognitive understanding of feelings and experiences** of others is key in mental-health support (Selman, 1980)

If that happened to me, I would feel really isolated

Empathy expressed in Text-based Mental Health Support





Our analysis suggests that **highly empathic conversations are rare** in text-based mental health support

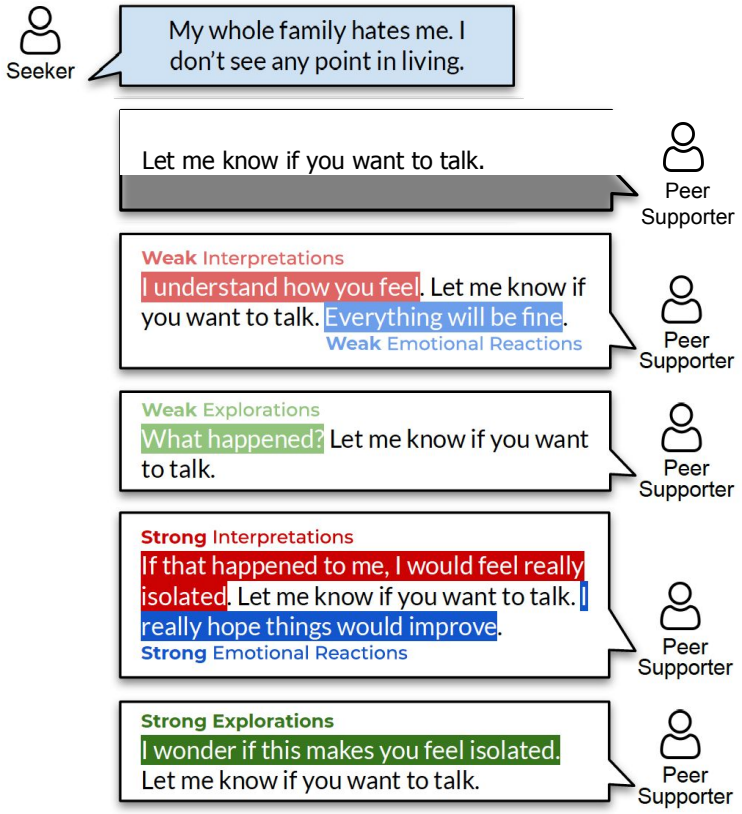
A new framework and dataset of empathy expressed in text-based mental health support

Adapted most prominent empathy scales from psychology/psychotherapy research to text-based mental health support in collaboration with clinical psychologists

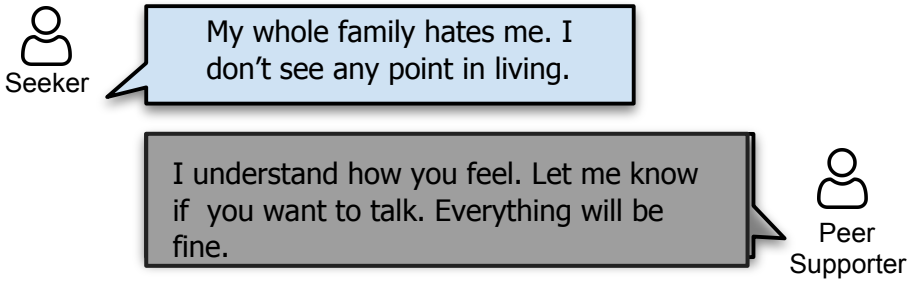
Framework of empathy expressed in conversations

Three communication mechanisms of empathy

- **Emotional Reactions:** communicating the emotions experienced after reading a post
- **Interpretations:** communicating understanding of the inferred feelings / experiences
- **Explorations:** improving one's understanding by exploring feelings / experiences



Tasks and Dataset



Task 1: Empathy Identification

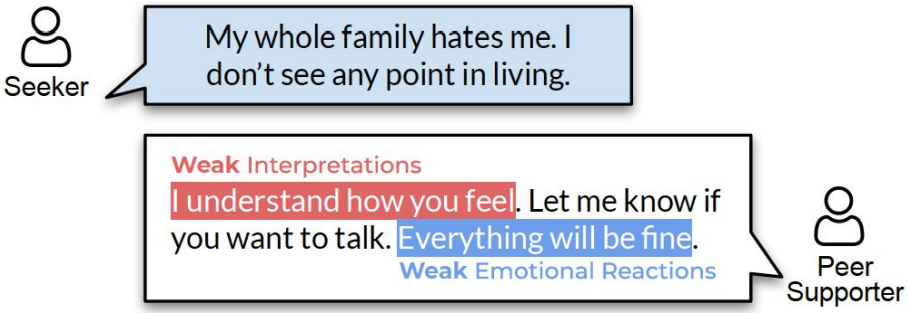
How empathic is response post in the context of seeker post?

Emotional Reactions – 1 out of 2
Interpretations – 1 out of 2
Explorations – 0 out of 2

Task 2: Rationale Extraction

What is the supporting rationale for the identified empathy levels?

Tasks and Dataset



Task 1: Empathy Identification

How empathic is response post in the context of seeker post?

Emotional Reactions – 1 out of 2
Interpretations – 1 out of 2
Explorations – 0 out of 2

Task 2: Rationale Extraction

What is the supporting rationale for the identified empathy levels?

Tasks and Dataset

Task 1: Empathy Identification

How empathic is response post in the context of seeker post?

Task 2: Rationale Extraction

What is the supporting rationale for the identified empathy levels?

- Dataset of 10k (post, response) pairs annotated on our framework of empathy with supportive evidences (*rationales*)
  - 7k from TalkLife, 3k from mental health subreddits
  - Hired and trained freelancers on Upwork
  - Series of phone calls and manual/automated feedback on sample posts
  - Kappa = **0.6865**

Model for identifying empathy
with supportive rationales

Multi-task, RoBERTa-based bi-encoder model

Computational Model

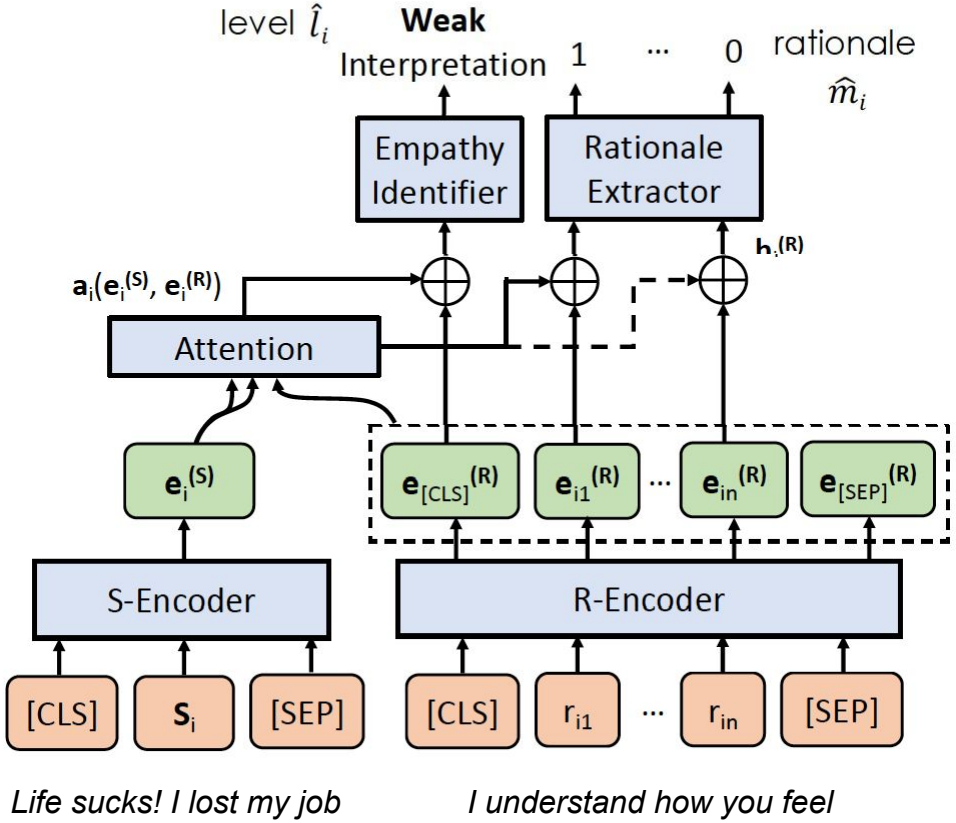
Multi-task, RoBERTa based bi-encoder model

Linear layer for input-level and token-level predictions

Attention between the two encodings

Two independently pretrained RoBERTa-based encoders

Tokenize seeker post and response post



(1) **Empathy identification**

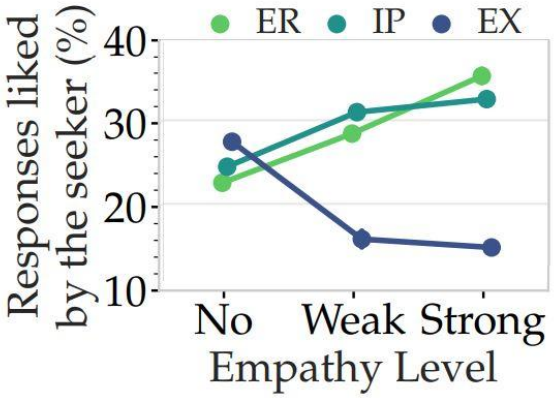
(2) **Rationale extraction**

(Seeker post, response post)

Model-based Insights
into Mental Health Platforms

Applied our model to a carefully filtered dataset of 235k interactions of significant mental health challenges on TalkLife(based on categories and triggering posts)

Validation: Positive feedback from seekers

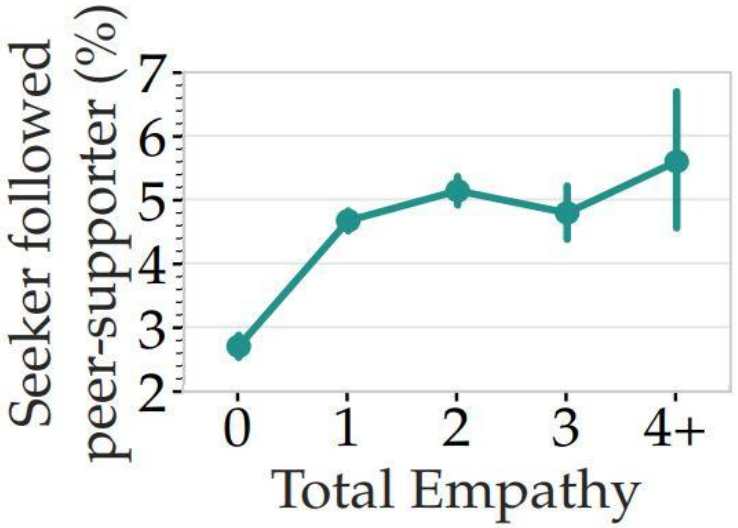


Strong communications of emotional reactions and interpretations receive **45% more likes** than their no communication

Stronger explorations get **47% more replies**

High empathy interactions are received positively by seekers

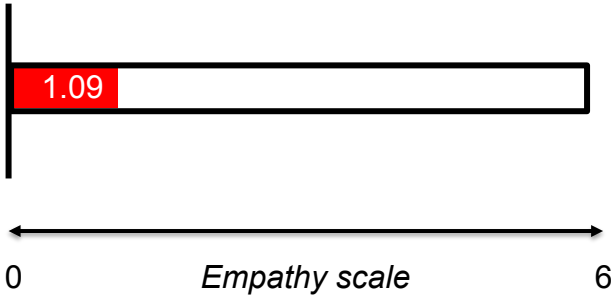
Validation: Forming of relationships



Seekers are 79% more likely to "follow" peer supporters after an empathic interaction than after a non-empathic one

Relationship forming more likely after empathic interactions

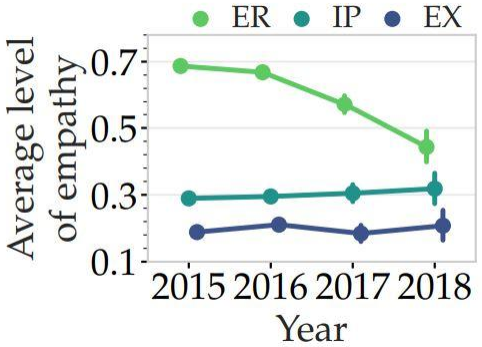
How empathic are peer supporters?



Empathy expressed is low

Does it improve over time?

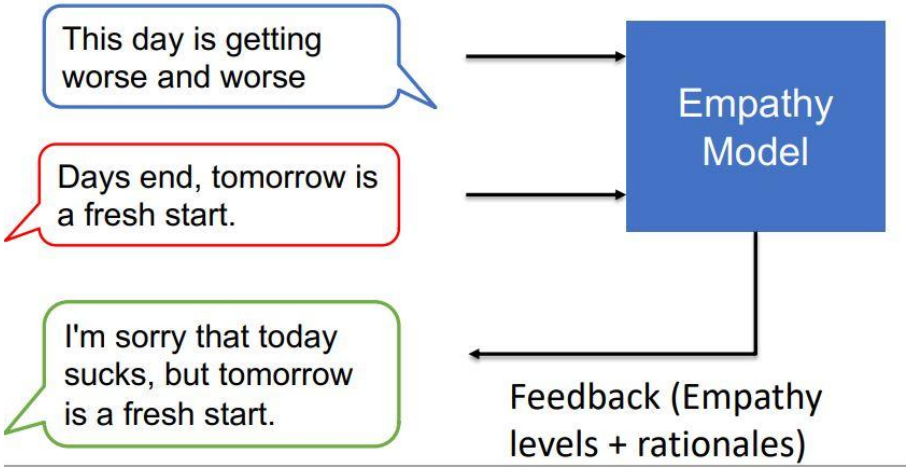
Peer-supporters do not self-learn empathy over time



- This is also true for therapists!
  - Without **deliberate practice** and **specific feedback**, even trained therapists often diminish in skills over time (Goldberg et al., 2016)

Implications for empathy-based feedback

- We can measure empathy successfully, and the measured components are important to mental health platforms
- However, highly empathic conversations are rare
- How can you help people express more empathy?
- **Simple Proof-of-concept using model-based feedback**



- Three participants were asked to rewrite responses using model-based feedback
- Empathy increased from 0.8 to 3

Evaluation Metrics (To be changed)

Automatic Evaluation

- **PerStr** - percentage of the utterances generated with persuasion strategy
- **PolSt** - percentage of utterances generated with consistent politeness strategy
- **Emp** - percentage of empathetic utterance generated,
- **PPL** - perplexity of the dialogue agent and
- **LEN** - number of tokens generated in an utterance

Human Evaluation

- **Per, Emp** - checking persuasiveness and empathy of the generated dialogue
- **DonPr** - computing percentage of time people donated
- **Const, Adeq, Fluen and N-Rep -** to evaluate if the generate utterances are consistent adequate, linguistically fluent and non-repetitive in nature

Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Problem Definition

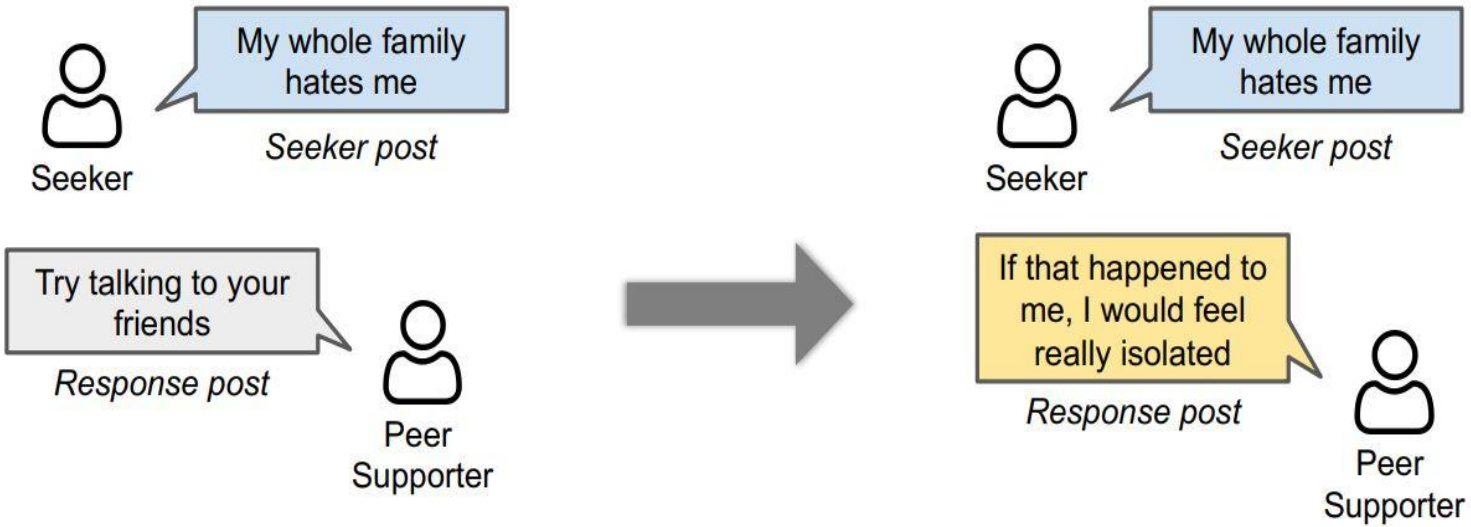
- ***Empathic Rewriting:*** Computationally transform low-empathy conversational posts to higher empathy

Motivation

- Improve empathy in online mental health support conversations
- High empathy interactions
  - Strong associations with symptom improvement in mental health []
  - Received positively by users on online peer support platforms []

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

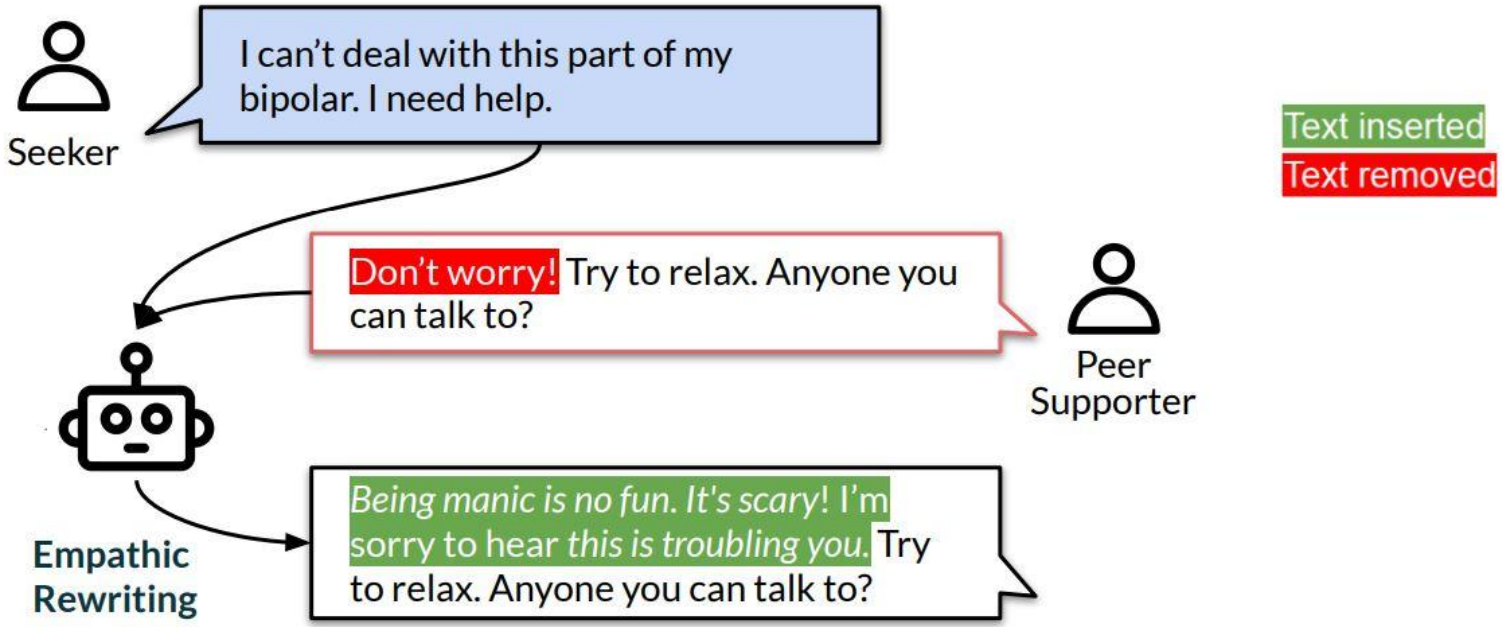
Highly empathic conversations are rare []



Key Question: How to improve empathy in peer-to-peer support

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Solution: Empathic Rewriting



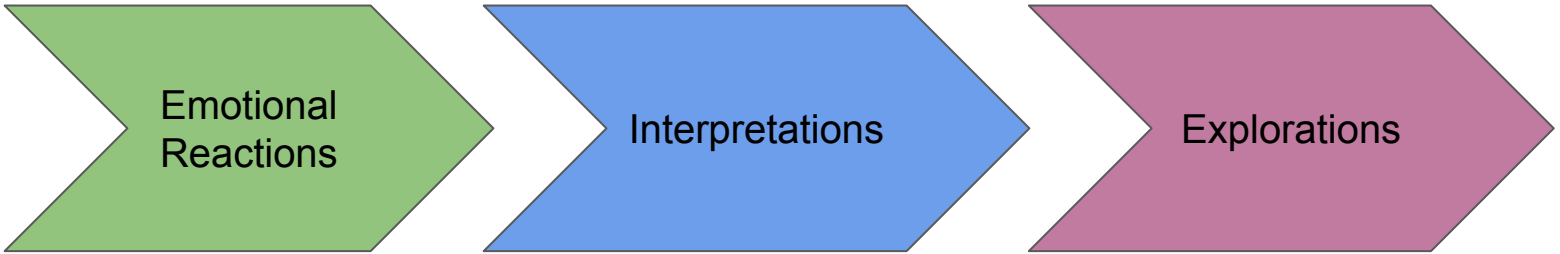
Transform low-empathy conversational posts on online peer-to-peer support platforms to higher empathy

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Challenges: Empathy is Complex

- Empathy is complex, conceptually-nuanced, multi-dimensional
  - Much more than sympathy or reacting with positive sentiment
  - Clinically relevant perspective – Understanding of hidden feelings and experiences

Theoretically-grounded framework of empathy [7]



Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Challenges: Why existing approaches fail?

- Can we transform every response to a generic, empathic response?
  - This must have been really hard for you
  - May not be specific to the emotions and experiences o Affects response diversity on the platform [8]
- Style transfer approaches may not work!
  - Requires changes beyond simple word-level transformations
  - **Sentiment Transfer:** *"The movie was bad"* → *"The movie was good"*
  - **Empathic Rewriting:** *"Being manic is no fun. It's scary! I'm sorry to hear …"* (3 new sentences)
- No parallel dataset exists and creating one is expensive
  - We will need domain-experts!

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Dataset

1

TalkLife Dataset

- 10.9M seeker posts
- 26.9M response posts
- 642K users

2

Curated mental -health related

- TalkLife hosts a large no. of social media interactions (“Happy mother’s day”)
- Curated mental health-related conversations using a BERT-based classifier
- 1.48M seeker posts and 3.33M response posts

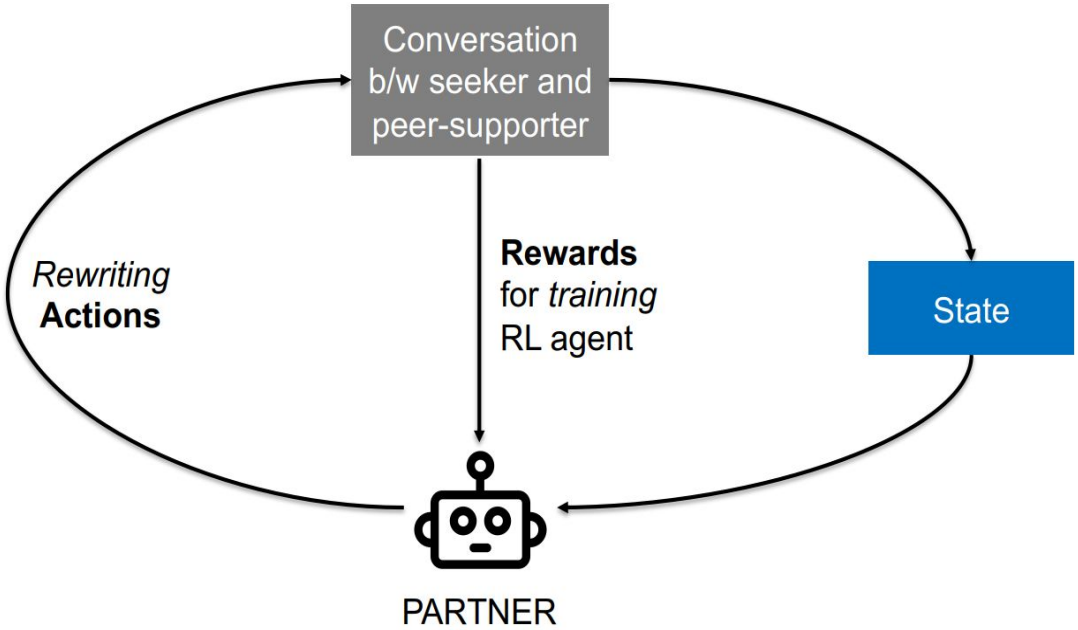
3

Computational labeling with empathy

- Label with empathy using the empathy classifier developed by [7]
- Used for a supervised warm-start training of our RL model

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: Empathic Rewriting using Reinforcement Learning (RL)

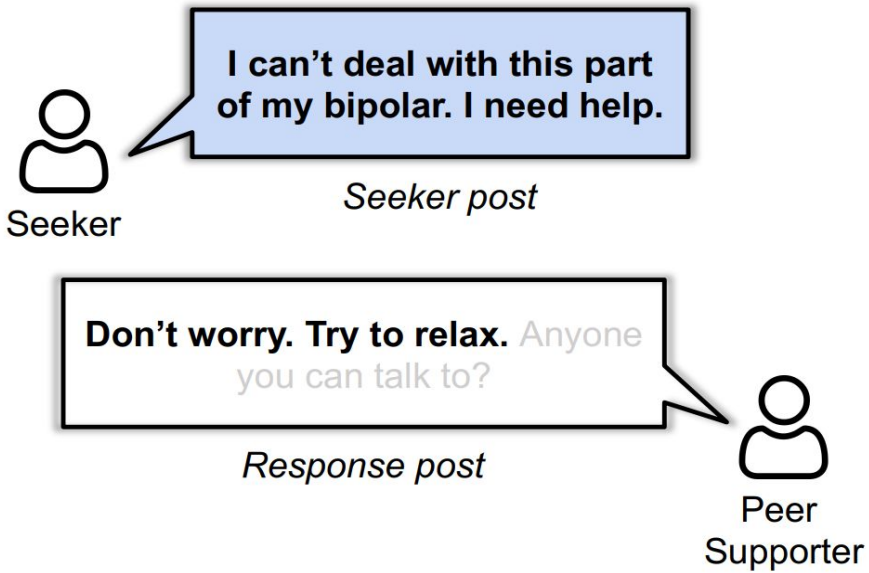


PARTNER: An RL Agent for the task of Empathic Rewriting

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: State

Seeker post & Fixed-length contiguous spans of response post



Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: Actions

- Sentence-level edits
  - Insert empathic sentences
  - Replace with empathic sentences

[Action 1] Select a position in the response span for insertion or replacement

[Action 2] Generate candidate empathic sentences

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: Actions

[Action 1] Select a position in the response span for insertion or replacement

Select the first sentence for replacement (**Don't worry**)

~~**Don't worry.**~~ **Try to relax.** Anyone you can talk to.

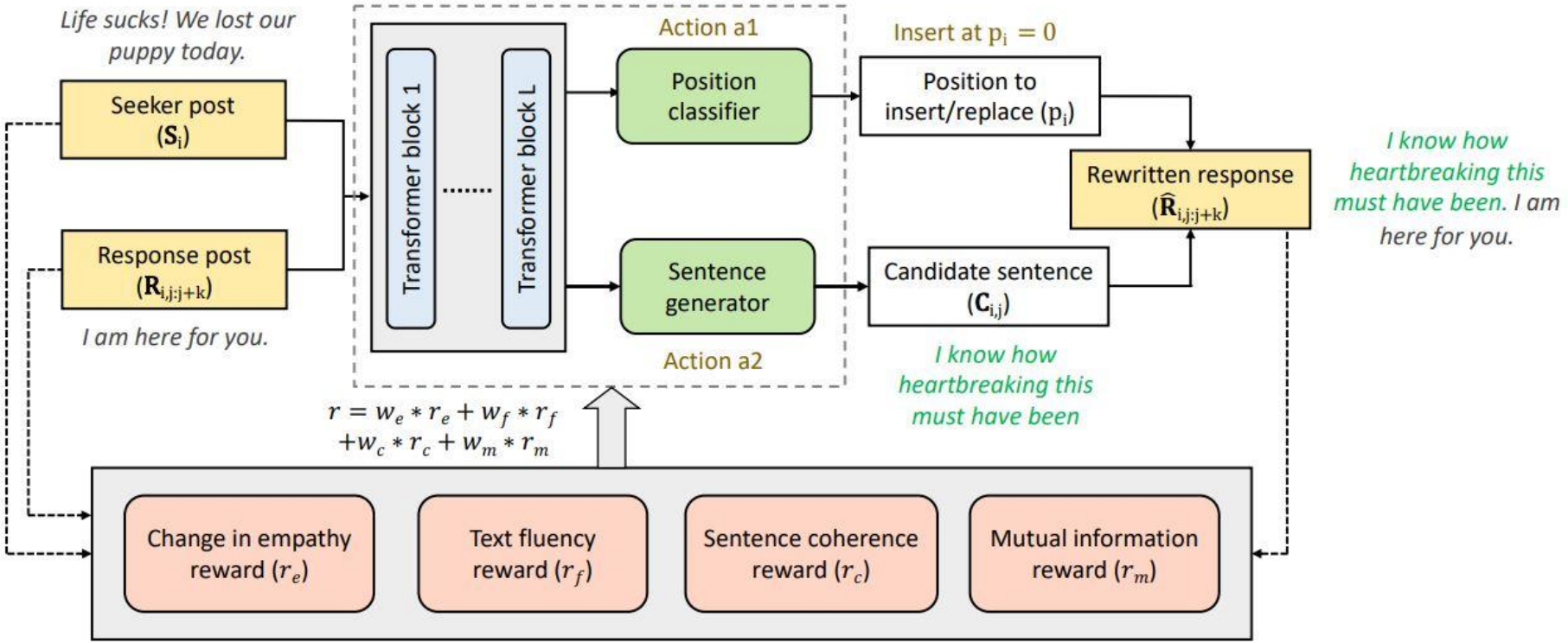
[Action 2] Generate candidate empathic sentences

Replace selected span (**Don't worry**) with "**Being manic is no fun**"

Being manic is no fun. **Try to relax.** Anyone you can talk to.

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: Policy



Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

PARTNER: Rewards

Ensuring highly empathic rewritings while maintaining *fluency, specificity,* and *diversity*

(1) **Change in empathy**

- Empathy of the rewritten response – Empathy of the original response
- Empathy scores b/w 0 to 6 based on a theoretically-grounded empathy framework [1]

(2) **Sentence Coherence**

- Average sentence coherence probability between a candidate sentence and existing sentences in the response.

(3) **Text Fluency**

- Measured using Perplexity of the rewritten response

(4) **Mutual information for specificity and diversity**

- To ensure specificity to the seeker post and diversity of responses, exploit the idea of maximizing mutual information between seeker post and the rewritten response post

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Evaluation Metrics

Automatic Evaluation

- **Change in empathy:** Measures is how much the empathy has changed from the original response to the rewritten responses
- **Perplexity:** Quantify fluency of the rewritten response
- **Sentence coherence:** Measure the sentence coherence in rewritten response
- **Specificity:** Measure how specific rewritten response is, to the seeker post
- **Diversity:** Measure diversity in the rewritten response
- **Edit rate:** Measure the number of changes between the original response and the rewritten respo

Human Evaluation

- **A/B Testing:** Compare PARTNER outputs against baseline models
  - Choose the output which is more
    - empathic
    - fluent
    - specific

Mishra et al. PEPDS: A Polite and Empathetic Persuasive Dialogue System for Charity Donation. In COLING 2022.

Conclusion and Future Directions

(15 minutes)

Conclusion

- Introduce conceptual models of empathy followed by need for empathy in conversational AI systems
    - An ideal ECAI system is expected to exhibit emotional and social competence
- Introduce and discuss various concepts related to empathy
    - Discussed the most recent and representative works utilizing these concepts
- Current trends in ECAI systems
    - Usefulness of empathy in persuasion
    - Relevance of empathy in therapy

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

Future Research Directions

- Combining target-dependent emotion with user modeling
  - Emotion is a particular dimension affixed to the speaker and other conversational participants.
  - Emotion and personality should be correlated dimensions of the user, and should therefore be modeled jointly
- Utilizing the existing knowledge base containing sentimental or emotional knowledge
  - Aid in detecting the emotional states of the user
  - Help in understanding background information beyond the context

Sharma et al. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. In WWW 2021.

References

[1] Ong, Desmond C., Jamil Zaki, and Noah D. Goodman. "Affective cognition: Exploring lay theories of emotion." *Cognition* 143 (2015): 141-162.

[2] Rashkin, Hannah, et al. "Towards Empathetic Open-domain Conversation Models: A New Benchmark and Dataset." *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019.

[3] Sharma, Ashish, et al. "A computational approach to understanding empathy expressed in text-based mental health support."

[4] Peskov, Denis, et al. "Multi-domain goal-oriented dialogues (multidogo): Strategies toward curating and annotating large scale dialogue data." *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 2019.

[5] Danescu-Niculescu-Mizil, Cristian, et al. "A computational approach to politeness with application to social factors."

[6] Wang, Xuewei, et al. "Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good." *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019.

[7] Sharma, Ashish, et al. "Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach." *Proceedings of the Web Conference 2021*. 2021.

[8] Tim Althoff, Kevin Clark, and Jure Leskovec. 2016. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. TACL (2016).

References

[9] C. K. Joshi, F. Mi, and B. Faltings, “Personalization in goal-oriented dialog,” 2017, arXiv:1706.07503.

[10] Hitesh Golchha, Mauajama Firdaus, Asif Ekbal, and Pushpak Bhattacharyya (2019). Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 851-860. 2019.

Acknowledgement

- Indian government's *"Prime Minister's Research Fellowship (PMRF) Program"*
- *"Innovation in Science Pursuit for Inspired Research (INSPIRE) Fellowship"* implemented by the Department of Science and Technology, Ministry of Science and Technology, Government of India
- Partial support from the project titled "An Empathetic Knowledge Grounded Conversational System for Mental Health Counseling and Legal Assistance", Sponsored by IHUB Anubhuti, TIH, IIIT Delhi.

PMRF
Prime Minister's Research Fellowship

inspire

THANK YOU!