

PAL to Lend a Helping Hand: Towards Building an Emotion Adaptive Polite and Empathetic Counseling Conversational Agent

Kshitij Mishra* and Priyanshu Priya* and Asif Ekbal

Department of Computer Science and Engineering

Indian Institute of Technology Patna, Bihta, Bihar, India

{kshitij_1921cs23, priyanshu_2021cs26, asif}@iitp.ac.in

Abstract

The World Health Organization (WHO) has significantly emphasized the need for mental health care. The social stigma associated with mental illness prevents individuals from addressing their issues and getting assistance. In such a scenario, the relevance of online counseling has increased dramatically. The feelings and attitudes that a client and a counselor express towards each other result in a higher or lower counseling experience. A counselor should be friendly and gain clients' trust to make them share their problems comfortably. Thus, it is essential for the counselor to adequately comprehend the client's emotions and ensure client's welfare, i.e. s/he should adapt and deal with the clients politely and empathetically to provide a pleasant, cordial and personalized experience. Motivated by this, in this work, we attempt to build a novel Polite and empathetic counseling conversational agent PAL. To have client's emotion-based polite and empathetic responses, two counseling datasets laying down the counseling support to substance addicts and crime victims are annotated. These annotated datasets are used to build PAL in a reinforcement learning framework. A novel reward function is formulated to ensure correct politeness and empathy preferences as per client's emotions with naturalness and non-repetitiveness in responses. Thorough automatic and human evaluation showcases the usefulness and strength of the designed novel reward function. Our proposed system is scalable and can be easily modified with different modules of preference models as per need¹.

1 Introduction

The global prevalence of mental health disorders is estimated to affect 20% of the world population (Holmes et al., 2018). The COVID-19 pandemic

*Jointly first authors.

¹Data and codes can be accessed at <https://github.com/Mishrakshitij/PAL.git> and [AI-NLP-ML-PAL](https://github.com/AsifEkbal/ML-PAL).

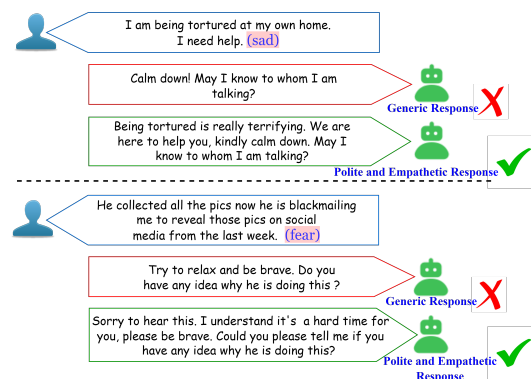


Figure 1: Example demonstrating the variation in polite and empathetic response of the agent depending on the client's emotion. The text highlighted in pink color shows the client's emotion.

has further jeopardized the mental well-being of individuals (Paredes et al., 2021). Further, the limited availability of experts (Organization et al., 2021) and the overwhelming healthcare burden (Huffman et al., 2021) have created a substantial demand for counseling services. Counselors are faced with a diverse audience, including individuals struggling with substance addiction and crime victims, who often experience mental health challenges such as anxiety, depression, and post-traumatic stress (Eisen et al., 1992; Ross and Peselow, 2012; Kilpatrick et al., 1985; Kilpatrick and Acierno, 2003). Therefore, pertinent and timely counseling support is crucial. In light of this, it is imperative to explore the potential benefits that counseling could derive from the advancements in artificial intelligence (AI) and natural language processing (NLP).

In order to ensure effective counseling, active listening, combined with accurate interpretation of the client's meaning and emotions is vital (Katz and McNulty, 1994). This demonstrates care and understanding, fostering a sense of being heard and valued, hence, facilitating deeper self-exploration for the client. The essential elements of effective support include politeness and empathy (Norcross,

2002; Robert et al., 2011; Newbold et al., 2019). Previous studies indicate that politeness plays a critical role in facilitating supportive interactions (Feng et al., 2016) and aids in the disclosure of sensitive information (Bickmore and Picard, 2004; Lucas et al., 2014). Quantitative data supports the idea that empathetic interactions play a significant role in reducing issues in mental health support (Elliott et al., 2018) and are essential for building a strong therapeutic bond and connection (Robert et al., 2011).

Thus, a counseling conversational agent should adapt its support to users' mental state whilst being polite and empathetic towards them. In Figure 1, the agent's polite and empathetic response in the green box fosters higher user engagement and satisfaction compared to the counseling response in the red box. Furthermore, the agent's response should align with the client's emotions - offering reassurance and compassion when they are sad and showing understanding and providing encouragement when they are afraid.

This paper introduces a novel client's emotion-adaptive **Polite and empathetic counseling dialogue system**, named **PAL** to create a comfortable, engaging and cordial atmosphere during counseling. For this, we create a novel counseling conversational dataset in English utilizing the recently released EmoInHindi (Singh et al., 2022) dataset and annotate it with defined emotion, politeness, and empathy labels. We also annotate the High-quality and Low-quality Counseling Conversations dataset (Pérez-Rosas et al., 2019) with emotion, politeness and empathy label for our work. Finally, to build our system, we first warm-start with a supervised fine-tuning (SFT) of a language model on these datasets. Then, this trained SFT model is fine-tuned in a reinforcement learning (RL) framework using preference models to ensure preferences of utterance-emotion politeness (and empathy) consistency, adaptiveness, and correctness by designing six preference rewards.

The *key* contributions and/or the distinguishing attributes of our current work can be summarized as: **(i)** Prepared two large-scale counseling conversational datasets annotated with three distinct labels, *viz.* emotion, politeness and empathy; **(ii)** Proposed **PAL** - a novel emotion-adaptive **Polite and empathetic counseling** conversational agent fine-tuned in an RL setting; **(iii)** Designed a novel reward function consisting of six preference re-

wards and two generic rewards to generate engaging, fluent and interactive client's emotion-adaptive polite and empathetic responses; **(iv)** Trained six transformer-based robust preference models comprising of two generation models and four classifiers to compute preference rewards; **(v)** Extensive automatic and human evaluation is performed with novel preference-specific and language quality-specific metrics to demonstrate the effectiveness of the proposed system.

2 Related Work

Recent research has focused on developing dialogue systems for online mental health support, investigating approaches such as context-specific adaptation, response diversity (Pérez-Rosas et al., 2019; Zhang and Danescu-Niculescu-Mizil, 2020; Sharma et al., 2020b, 2021; Saha et al., 2022a), linguistic development of counselors (Zhang et al., 2019), conversational engagement patterns (Sharma et al., 2020a), moderation analysis (Wadden et al., 2021), therapeutic action detection (Lee et al., 2019), and cognitive restructuring identification (Pruksachatkun et al., 2019).

In our current work, we prioritize two crucial aspects of conversation, namely politeness and empathy, which are fundamental components of counseling (Norcross, 2002; Bickmore and Picard, 2004; Robert et al., 2011; Lucas et al., 2014). Previous research has demonstrated that perceiving conversational agents as caring and polite can enhance empathy and encourage the disclosure of personal and sensitive information (Kim et al., 2018; Mishra et al., 2022a; Samad et al., 2022; Mishra et al., 2023). For instance, Newbold et al. (2019) and Mishra et al. (2022b) applied Brown and Levinson's politeness strategies (Brown et al., 1987) to create dialogue templates for mental health "mood log" and develop a polite persuasive dialogue system, respectively. The agent's polite demeanor reflects a sense of care and personal investment, resembling a human companion.

Prior works (Ghandeharioun et al., 2019; Mielewszczenko-Kowszewicz et al., 2022) advocate incorporating emotion as an extra input in generation frameworks. This can be achieved by including emotion either during decoding phase or using it as a reward to assist models in generating responses that align with the user's emotions. The role of empathy in therapy has been explored in (Morris et al., 2018; Sharma et al., 2020b), while studies

1.	Original (Hindi)	ठीक है कृपया इस विकल्प के साथ मेरी मदद करें।
	Translation (English)	Ok please help me with this option .
2.	Original (Hindi)	क्या आप कृपया हमें बताएंगे कि क्या आपने कभी उन गतिविधियों को महसूस किया है जो आपको उसके बारे में संदेह में रखती हैं?
	Translation (English)	Would you please tell us that have you ever felt the activities that kept you in doubt about him.

Figure 2: Sample translated utterances of EmoInHindi.

in (Sharma et al., 2021; Saha et al., 2022b) have investigated methods to induce empathy in mental health support conversations. Further, several research studies have endeavoured to understand and create computational techniques for recognizing empathy in various contexts, such as face-to-face therapy (Gibson et al., 2016) and text-based peer-to-peer support systems (Sharma et al., 2020b).

Existing studies primarily focus on analyzing social media posts and activities to detect different mental health concerns and provide support to users. They often emphasize incorporating empathy into the agent’s responses to instill motivation and hope in the user. In contrast, our current work aims to develop a counseling dialogue agent that can adapt its politeness and empathy based on the user’s emotions.

3 Dataset

We perform the experiments on two different counseling datasets, viz. English translation of EmoInHindi (Singh et al., 2022) dataset and High-quality and Low-quality Counseling Conversations dataset (Pérez-Rosas et al., 2019).

Metrics	Train	Validation	Test
# of Dialogues	1,179	181	454
# of Utterances	28,837	4,464	10,946
Avg. Utterances per Dialogue	24.46	24.66	24.11

Table 1: *en-EmoInHindi* dataset statistics.

3.1 English Translation of EmoInHindi

EmoInHindi is a Hindi conversational dataset comprising 1,814 dialogues focused on mental health and legal counseling for women and children victims of crimes like stalking, harassment, domestic violence, etc. For our present work, we translate each EmoInHindi dialogue to English, preserving *Fluency* (F) and *Adequacy* (A). The resulting dataset is named “*en-EmoInHindi*”. The translation is conducted in two phases with four experienced human translators who are native Hindi speakers

with post-graduate qualifications and relevant expertise.

In the first phase, we use Google Translate to translate the utterances from Hindi to English. The translators are asked to manually check and mark every translated utterance for both *F* and *A* on a scale of 1-5² (White and O’Connell, 1996). The average scores of *F* and *A* are found to be 3.23 and 3.10 respectively. In the second phase, to further

Metrics	Train	Validation	Test
# of Dialogues	167	26	65
# of Utterances	7,082	1,110	2,767
Avg. Utterances per Dialogue	42.41	42.69	42.57

Table 2: HLCC dataset statistics.

enhance translation quality, we select utterances with *F* or *A* scores of 1, 2, or 3 and instruct translators to post-edit them by referring to the original Hindi utterance. Subsequently, all four translators rate each corrected utterance on the 1-5 scale for both *F* and *A*. This process yields average *F* and *A* scores of 4.65 and 4.48, respectively, indicating good translation quality. In the scoring of fluency and adequacy, the inter-translator agreement has been found to be <69.4%, 66.4%> and <74.8%, 71.3%> in phases 1 and 2, respectively. The dataset statistics are given in Table 1. A couple of sample translations from EmoInHindi are shown in Figure 2. A few examples of erroneous utterances observed during translation are presented in Section A.1 of the Appendix.

3.2 High-quality and Low-quality Counseling Conversations Dataset

The High-quality and Low-quality Counseling Conversations (HLCC) dataset consists of 258 counseling interactions, with 154 interactions labeled as high-quality counseling and 104 labeled as low-quality counseling. The counseling is conducted using Motivational Interviewing (MI) and each counseling conversation has annotations on two micro-skills frequently used in the evaluation of MI counseling, namely reflective listening, and questions. The counseling sessions primarily cover different topics on behavior change, including smoking cessation, alcohol consumption, substance abuse, weight management, and medication adherence. The dataset statistics are given in Table 2.

²*Fluency* - 5: Flawless, 4: Good, 3: Non-native, 2: Disfluent, 1: Incomprehensible; *Adequacy* - 5: All, 4: Most, 3: Much, 2: Little, 1: None

3.3 Dataset Annotation

In order to build **PAL**, we annotate *en-EmoInHindi* and *HLCC* datasets with three distinct aspects, *viz.* emotion, politeness, and empathy, and name these datasets as *EPE-enEIH* and *EPE-HLCC*, respectively. Each task is carried out with the help of three annotators, two with Ph.D. degree in Linguistics and one with a Master’s degree³. All three annotators are proficient in English, sufficiently acquainted with labeling tasks and well-versed with the concepts of emotions, politeness and empathy.

3.3.1 Annotation for emotion

We annotate the client’s utterances in both datasets with emotion labels utilizing the *EMPATHETIC-DIALOGUES* dataset (Rashkin et al., 2019) comprising of 25K conversations grounded on 32 different emotions in two phases. In the first phase, we begin by reducing the number of emotions in the *EMPATHETIC-DIALOGUES* dataset from 32 to 11 by consolidating similar emotion labels into one emotion to obtain a better class distribution. This dataset with consolidated emotion labels will be called *EmpDial-merged*, hereafter. Table 3 shows the specifics of the merged emotion labels.

Then, a pre-trained RoBERTa (Liu et al., 2019) based model is fine-tuned on *EmpDial-merged* dataset for emotion classification. We notice that the emotion classifier trained on 11 emotions outperformed the classifier built on 32 emotion labels by a substantial margin⁴. This trained emotion classifier is used to label each utterance in both the datasets with an appropriate emotion class.

In the second phase, we employ three annotators and brief them about each emotion class along with a few examples. Thereafter, we ask them to manually cross-check the predicted emotion label of the utterances in both the datasets and correct it in case of any error. This gives our gold-standard datasets. We observe a reliable multi-rater Kappa (McHugh, 2012) agreement ratio of 71.6% and 73.4% for *HLCC* and *en-EmoInHindi* datasets, respectively.

3.3.2 Annotation for politeness

Politeness annotation for the agent’s utterances in both *HLCC* and *en-EmoInHindi* datasets is carried out in two phases. In the first phase, each utterance is passed through the *Stanford Politeness Classifier* trained on Wikipedia requests data (Danescu-

Niculescu-Mizil et al., 2013). The classifier outputs a politeness score between 0 and 1 and classifies the utterance as polite if the score is ≥ 0.5 , else impolite. In order to effectively adapt the conversation to the user’s mood and emotions during an ongoing dialogue, we fine-grain the politeness label as *impolite*, *low_polite* and *highly_polite*.

The utterances with a score < 0.5 are marked as *impolite*. The polite utterances with a score ≥ 0.5 and ≤ 0.65 are marked as *low_polite* and the utterances with a score > 0.65 are marked as *highly_polite*. The politeness score range for *low_polite* and *highly_polite* is determined empirically through inspection on a small subset of samples. In the second phase, we ask the same three annotators to manually cross-check the labels and correct the labels if found to be incorrect. A multi-rater Kappa (McHugh, 2012) agreement ratio of approximately 84% is observed among these annotators. Finally, majority voting is applied to decide the final politeness class of the given utterance.

3.3.3 Annotation for empathy

The agent’s utterances in both datasets are annotated with one of the ternary empathy labels, *viz.* *non-empathetic*, *low_empathetic* and *highly_empathetic*. To lessen the manual efforts, the empathy annotation proceeds in two stages. In the first stage, we randomly sample 100 and 600 dialogues from *HLCC* and *en-EmoInHindi* datasets, respectively and then ask all three annotators to manually annotate with the required empathy label. In the second stage, pre-trained RoBERTa-large (Liu et al., 2019) is fine-tuned on manually annotated samples to build an empathy classifier. Then, the remaining utterances in both datasets are passed through this classifier to predict the corresponding empathy label. Lastly, same annotators are asked to cross-verify the predicted labels and correct them, if needed. A reliable multi-rater Kappa (McHugh, 2012) agreement ratios of 78.2% and 77.1% in the first phase, and 86.5% and 84.3% in the second phase are observed for *HLCC* and *en-EmoInHindi* datasets, respectively.

4 Proposed Methodology

To build our proposed system, first, to foster natural language interaction between the counseling agent and the client, we get a warm-start by training a client’s emotion-informed dialogue model (*EIDM*) in a supervised learning framework. Then, to generate the client’s emotion-adaptive polite and

³Annotators are paid as per institute norms.

⁴The accuracy of the classifiers trained on 32 and 11 emotion labels was found to be 58.17% and 77.46%, respectively.

Distinct Emotion Class	Merged Emotion Class
caring, sentimental	caring
hopeful	hopeful
proud, content	content
surprised	surprised
afraid, apprehensive, anxious	afraid
confident, prepared	confident
grateful	proud
trusting	proud
joyful, impressed	joy
disappointed, guilty, embarrassed, sad, lonely	sad
angry, annoyed, jealous	angry

Table 3: 32 distinct emotions in EMPATHETICDIALOGUES dataset and their corresponding merged emotion class.

empathetic responses, *EIDM* is fine-tuned with a Proximal Policy Optimization (PPO) (Schulman et al., 2017) loss utilizing novel designed eight rewards. This end model is our proposed **Polite and empathetic counSELing** conversational system **PAL**. To generate polite and empathetic responses as per user’s varying emotion, we use six preference models *viz.* two sequence-to-sequence generation models - Utterance-emotion-politeness (G_{uep}) and Utterance-emotion-empathy (G_{uee}) and four classifiers - client’s utterance+emotion politeness (C_{uep}) and empathy (C_{uee}) and agent’s utterance politeness (C_{aep}) and empathy (C_{aee}). These preference models are used to design rewards that are maximized to force the agent to generate the client’s emotion-adaptive polite and empathetic responses. The overall architecture of the proposed system is shown in Figure 3.

A multi-turn dialogue with a client’s and counseling agent’s utterances can be represented as $d = \{ca_0, cl_0, \dots, ca_i, cl_i, \dots, ca_{T-1}, cl_{T-1}\}$, where ca_i and cl_i give the counseling agent’s and client’s i^{th} response out of T turns. Let the client’s current utterance with corresponding emotion, $e_k \in E = \{e_1, \dots, e_m\}$ (m is the number of emotion classes), be represented as $cl_{e_i} = [cl_i + cl_{i_{e_k}}]$. Further, let ρ_{ca} and ρ_{cl_e} representing agent’s and client’s utterances with corresponding emotions distributions, respectively. Now for a given context, an emotion-informed dialogue model (*EMID*) $\rho_\theta(d)$ is trained by minimizing the KL-divergence (Kullback and Leibler, 1951) between the predicted and ground truth utterances probability distributions. $\rho_\theta(d)$ can be expressed as:

$$\rho_\theta(d) = \prod_{T=0}^{T-1} \rho_{cl_e}(cl_i | cl_{e < i}, ca_{< i}) \rho_{ca}(ca_i | cl_{e < i}, ca_{< i}) \quad (1)$$

4.1 PAL

$\rho_\theta(d)$ is further fine-tuned in an RL framework with a PPO loss using a novel reward function $R = \beta_1 R_1 + \beta_2 R_2 + \beta_3 R_3 + \beta_4 R_4 + \beta_5 R_5 + \beta_6 R_6 + \beta_7 R_7 + \beta_8 R_8$, comprising of eight rewards $R_1, R_2, R_3, R_4, R_5, R_6, R_7$, and R_8 , to generate client’s emotion-adaptive polite and empathetic counseling responses while keeping informative content intact. All rewards values are between 0 and 1 and $\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + \beta_8 = 1$. These rewards are obtained using six preference models to ensure politeness and empathy preferences as per the client’s emotion. At each time step of training, for a given context, n -candidate responses are generated and are scored in terms of reward using these preference models. This score is further propagated in the policy loss as feedback to force the agent toward generating preferred high-quality candidates. RL Policy and PPO loss details can be referred to in Section A.2 of the Appendix.

4.1.1 Preference Models and Rewards

The proposed reward function R consists of two types of rewards *viz.* (i) Preference rewards (R_1, R_2, R_3, R_4, R_5 , and R_6) - to adapt politeness and empathy as per client’s emotion, and (ii) Generic rewards (R_7 , and R_8) - to account for dialog flow consistency and diversity in generated responses. To design preference rewards, six preference models are trained and generic rewards are obtained using the BERT-based cosine similarity model.

Preference Models. R_1 and R_2 are designed by training two BART-large (Lewis et al., 2019) based seq2seq models - (i) *Utterance-emotion-politeness* (G_{uep}): It takes client’s utterance with its corresponding emotion cl_{e_i} as input and outputs agent’s utterance with its corresponding politeness $[ca_i + ca_{i_{po_j}}]$ (here $po_j \in Po = \{po_0, po_1, po_2\}$), (ii) *Utterance-emotion-empathy* (G_{uee}): It takes client’s utterance with its corresponding emotion cl_{e_i} as input and outputs agent’s utterance with its corresponding empathy $[ca_i + ca_{i_{em_l}}]$ (here, $em_l \in Em = \{em_0, em_1, em_2\}$). R_3, R_4, R_5 , and R_6 are designed by training four RoBERTa-large (Liu et al., 2019) based classifiers - (i) *client’s utterance+emotion politeness* (C_{uep}): Input - cl_{e_i} , Output - $ca_{i_{po_j}}$ (ii) *client’s utterance+emotion empathy* (C_{uee}): Input - cl_{e_i} , Output - $ca_{i_{em_l}}$ (iii) *agent’s utterance politeness* (C_{aep}): Input - ca_i ,

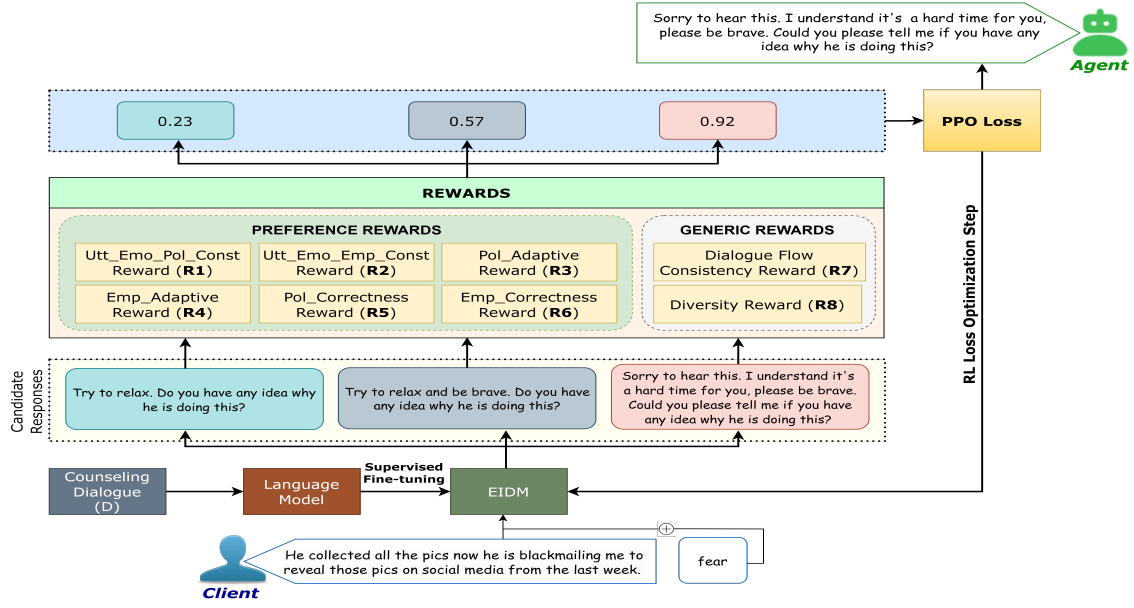


Figure 3: Overall architecture of the proposed system.

Output - $ca_{i_{po_j}}$ (iv) agent's utterance empathy (C_{aee}): Input - ca_i , Output - ca_{iem_i} .

Utterance-emotion-politeness and Utterance-emotion-empathy consistency rewards. It is a challenging task to approximate the relationship between client's dynamically changing emotion and the counselor's politeness or empathy, as for each client's utterance and its corresponding emotion, a counseling agent should change the way of counseling by applying correct politeness and empathy in its responses. Hence, to force the agent to employ emotion adaptive utterance-wise correct politeness (R_1) at turn T , the difference between negative likelihood losses (NLL) produced by G_{uep} when given with ground truth responses input-output pair ($NLL(G_{gr_{uep}})$) and predicted responses input-output pair ($NLL(G_{pr_{uep}})$) is taken. Similarly, for empathy, R_2 is computed using G_{uee} . The responses, deviating from ground truth response pair loss are penalized. R_1 and R_2 can be formulated as:

$$R1 = \tanh(NLL((G_{gr_{uep}}) - \alpha \times NLL(G_{pr_{uep}})) \quad (2)$$

$$R2 = \tanh(NLL((G_{gr_{uee}}) - \alpha \times NLL(G_{pr_{uee}})) \quad (3)$$

Politeness and Empathy-adaptive Rewards. To adapt varying politeness and empathy levels as per the client's utterance and emotion, the agent needs to be forced towards correct politeness and empathy. Using C_{uep} and C_{uee} , we get the agent's politeness and empathy probabilities for client's utterance with emotion (cl_{e_i}) for ground truth ($C_{gr_{uep}}$

and $C_{gr_{uee}}$) as well as agent's predicted responses ($C_{pr_{uep}}$ and $C_{pr_{uee}}$), and take the difference between them. The greater the deviation of the predicted response politeness or empathy, the greater will be the penalization.

$$R3 = \text{prob}(C_{gr_{uep}}) - \gamma \times \text{prob}(C_{pr_{uep}}) \quad (4)$$

$$R4 = \text{prob}(C_{gr_{uee}}) - \gamma \times \text{prob}(C_{pr_{uee}}) \quad (5)$$

Politeness and Empathy Correctness Rewards. Generated response deviation from true politeness and empathy labels is restricted by computing politeness and empathy probabilities of ground truth agent's utterances ($C_{gr_{aep}}$ and $C_{gr_{aee}}$) and agent's predicted utterances ($C_{pr_{aep}}$ and $C_{pr_{aee}}$) and taking the difference between them.

$$R5 = \text{prob}(C_{gr_{aep}}) - \delta \times \text{prob}(C_{pr_{aep}}) \quad (6)$$

$$R6 = \text{prob}(C_{gr_{aee}}) - \delta \times \text{prob}(C_{pr_{aee}}) \quad (7)$$

where, α , γ , and δ work as the penalization factors, i.e. greater respective values will result in hard penalization⁵.

Dialogue Flow Consistency Reward. In achieving all six preferences, generated responses should be consistent with the context of the dialogue. Therefore, we take the cosine similarity of generated response (pr_T) with context at turns $T - 1$ ($ctxt_{T-1}$) and T ($ctxt_T$). To avoid similar response generation, the threshold value of 0.75

⁵The value of α , γ , and δ is taken as greater than or equal to 1.

Dataset	PPL	BLEU	NIST	MET	R-2 F1
EPE-enEIH	3.87	0.23	0.313	0.553	0.340
EPE-HLCC	7.16	0.12	0.261	0.461	0.273

Table 4: Evaluation results of utterance+emotion politeness and empathy models.

is taken. Above the threshold value similarity, the generated responses are penalized.

$$R_7 = \frac{(\min(\cos(pr_T, ctxt_T) + \frac{\cos(pr_T, ctxt_{T-1})}{2}), 0.75)}{2} \quad (8)$$

Diversity Reward. To ensure an interactive and engaging counseling dialogue, diverse responses are crucial as the length of the dialogue increases. In this regard, we introduce the diversity reward R_8 , which takes into account two factors influencing dialogue quality: *diversity*, measured as the Jaccard similarity between pr_T and pr_{T-1} at turns T and $T - 1$ (Jaccard, 1912), and the *dialogue length*.

$$R_8 = (\text{len}(d)) - \left(\frac{pr_{T-1} \cap pr_T}{pr_{T-1} \cup pr_T} \right) \quad (9)$$

5 Experiments

We compare the performance of our PPO loss-based model to three baselines: GPT-2-medium (Radford et al., 2019), a strong baseline ARDM (Wu et al., 2021) and *EIDM*. Due to space restrictions implementation details are given in Section A.3 of the Appendix.

Evaluation Metrics. To evaluate the performance of the proposed system, **PAL**, both automatic and human evaluations are conducted. The generation preference models are assessed using Bleu score (BLEU) (Papineni et al., 2002), Perplexity (PPL) (Brown et al., 1992), METEOR score (MET) (Banerjee and Lavie, 2005), NIST score (NIST) (Doddington, 2002) and Rogue-2 F-1 score (R-2-F1) (Lin, 2004). All four classifier preference models are evaluated in terms of Weighted Accuracy (W-ACC) and Macro-F1. In automatic evaluation, **PAL** efficacy is evaluated in terms of two aspects: (i) preference success - Emotion-politeness consistency (**EPC**), Emotion-empathy consistency (**EEC**), Politeness correctness (**PC**), and Empathy Correctness (**EC**)⁶, and (ii) response quality - Perplexity (**PPL**) and Response-length (**R-LEN**).

For human evaluation, we recruit three evaluators with postgraduate qualifications and relevant

⁶EPC, EEC, PC, and EC are computed by four respective classifiers in Table 5.

Classifier	EPE-enEIH		EPE-HLCC	
	W-ACC	Macro-F1	W-ACC	Macro-F1
Utt+Emo Politeness	0.811	0.803	0.664	0.651
Utt+Emo Empathy	0.835	0.816	0.652	0.631
Politeness	0.857	0.848	0.694	0.673
Empathy	0.873	0.859	0.686	0.654

Table 5: Evaluation results of the classifiers. Utt+Emo Politeness (or Empathy) classifier takes user’s utterance+emotion and outputs agent’s politeness (or Empathy). Politeness (classifier) takes agent’s utterance and outputs politeness (or empathy).

experience in the field⁷. Initially, each evaluator engages with the system seven times, with a different set of responses each time. These 21 human-evaluated interactions are then cross-verified by experts from government-run institutions to ensure evaluation quality. Upon passing verification, an additional 30 interactions (10 per evaluator) are evaluated, resulting in a total of 51 human-evaluated dialogues. Human evaluation metrics include Emotion-politeness consistency (**EPC**), Emotion-empathy consistency (**EEC**), Politeness correctness (**PC**), Empathy correctness (**EC**), Fluency (**Flue**), Non-Repetitiveness (**N-Rep**) and Consistency (**Cons**). All dialogue interactions are evaluated on an integer Likert scale of 1-5^{8 9}.

6 Results and Analysis

First, all six preference models results are shown i.e., two generation models and four classifiers. Thereafter, the results of the proposed system **PAL** are stated and compared with three baselines, viz. LM (GPT-2-medium), ARDM, *EIDM* and three variants of **PAL**, **PAL-R** (**PAL** with zero reward), **PAL-GR** (**PAL** without generic rewards) and **PAL-PR** (**PAL** without preference rewards).

For both the datasets - *EPE-enEIH* and *EPE-HLCC*, Tables 5 and 4 show the evaluation results of all the preference models. Scores of PPL, BLEU, MET and R-2-F1 in Table 4 depicts that BART-large performed well enough to generate fluent and diverse sentences. It is also seen that all four RoBERTa-large based classifiers gave well scores of W-ACC and Macro-F1.

6.1 Automatic evaluation

Table 6 shows that proposed **PAL** gives better performance than the baselines: LM, ARDM, and

⁷Human evaluators are remunerated according to our institute’s policies.

⁸1-5 denotes low to high.

⁹An inter-evaluator agreement ratio of 75.1% is observed in evaluations.

Model	Dataset	EPC	EEC	PC	EC	PPL	R-LEN
LM	<i>EPE-enEIH</i>	62.3%	64.6%	65.8%	66.6%	3.91	15.11
	<i>EPE-HLCC</i>	54.1%	52.7%	59.4%	58.6%	16.19	18.14
ARDM (Wu et al., 2021)	<i>EPE-enEIH</i>	64.4%	68.2%	67.2%	69.8%	3.21	16.24
	<i>EPE-HLCC</i>	55.8%	54.3%	60.6%	58.7%	15.01	19.41
EIDM	<i>EPE-enEIH</i>	69.3%	72.6%	68.4%	71.2%	2.65	17.63
	<i>EPE-HLCC</i>	57.7%	56.4%	62.3%	59.2%	14.26	22.30
PAL-R	<i>EPE-enEIH</i>	68.1%	70.9%	67.8%	69.7%	2.87	16.87
	<i>EPE-HLCC</i>	56.8%	56.9%	61.5%	57.9%	14.90	21.14
PAL-GR	<i>EPE-enEIH</i>	72.3%	74.1%	70.6%	73.3%	2.18	19.22
	<i>EPE-HLCC</i>	59.3%	58.6%	65%	62.3%	13.55	25.16
PAL-PR	<i>EPE-enEIH</i>	71%	71.8%	69.5%	71.8%	2.38	22.17
	<i>EPE-HLCC</i>	58.8%	57.3%	63.9%	61.7%	13.63	24.5
PAL	<i>EPE-enEIH</i>	73.7%	75.9%	71.6%	75.2%	2.03	21.09
	<i>EPE-HLCC</i>	61.4%	60.7%	65.8%	63.1%	12.19	25.81

Table 6: Results of automatic evaluation.

Model	Dataset	EPC	EEC	PC	EC	Flue	N-Rep	Cons
LM	<i>EPE-enEIH</i>	2.68	2.86	2.92	3.04	3.10	2.90	3.02
	<i>EPE-HLCC</i>	2.11	2.32	2.18	2.14	2.88	2.64	2.71
ARDM (Wu et al., 2021)	<i>EPE-enEIH</i>	2.75	2.94	3.02	3.10	3.41	3.03	3.20
	<i>EPE-HLCC</i>	2.34	2.58	2.42	2.37	2.97	2.73	2.85
EIDM	<i>EPE-enEIH</i>	2.96	3.15	3.33	3.46	3.87	3.10	3.62
	<i>EPE-HLCC</i>	2.47	2.71	2.58	2.54	3.11	2.95	3.03
PAL-R	<i>EPE-enEIH</i>	2.86	3.22	3.30	3.49	3.74	3.19	3.54
	<i>EPE-HLCC</i>	2.52	2.59	2.63	2.42	3.12	3.04	3.12
PAL	<i>EPE-enEIH</i>	3.43	3.69	3.85	3.98	4.14	3.68	4.03
	<i>EPE-HLCC</i>	2.79	3.02	2.92	2.81	3.42	3.37	3.39

Table 7: Results of human evaluation.

EIDM as well as variants **PAL-R**, **PAL-GR**, **PAL-PR** in terms of all metrics. It can be observed that **PAL** achieves well scores in terms of preferences **EPC**, **EEC**, **PC**, and **EC** with a significant difference of $\langle 11.4, 9.3, 4.4, 5.6, 1.4, 2.7 \rangle$, $\langle 11.3, 7.7, 3.3, 5, 1.8, 4.1 \rangle$, $\langle 5.8, 4.4, 3.2, 3.8, 1, 2.1 \rangle$, and $\langle 8.6, 5.4, 4, 5.5, 1.9, 3.4 \rangle$, respectively as compared to LM, ARDM, *EIDM*, **PAL-R**, **PAL-GR**, and **PAL-PR**¹⁰ on *EPE-enEIH*. Hence, the design of our reward function in **PAL** is justified by its ability to encourage the agent to generate emotion-adaptive, correct, polite, and empathetic responses aligned with the client’s preferences. Similarly, **PAL** outperforms all other models in terms of **PPL** and **R-LEN**. Hence, the influence of generic rewards can be attributed to the model’s ability to establish a natural and seamless connection with the user, leading to interactive and engaging interactions. Additionally, it is noteworthy that there is minimal disparity between the results of ARDM and GPT-2, while *EIDM* demonstrates substantial improvement over both of them. It could be due to the use of emotion aspects as well during training.

¹⁰All results are statistically significant. We perform a statistical significance test, Welch’s t-test (Welch, 1947), and it is conducted at 5% (0.05) significance level.

Further, scores of **EPC**, **EEC**, **PC**, **EC** for **PAL-PR** are lower than **PAL-GR** suggesting the importance of preference rewards. Similarly, scores of **PPL** and **R-LEN** are lower in absence of generic rewards, hence, showcasing their requirement as well. The obtained results strengthen our hypothesis that incorporating politeness and empathy aspects into the counseling dialogue system enables to create a more engaging, comfortable, cordial, and client-understanding environment.

6.2 Human evaluation

Human assessments are recorded in Table 7. It can be seen that as compared to the LM, ARDM, *EIDM*, and **PAL-R**, the proposed **PAL** yields better scores in terms of **EPC**, **EEC**, **PC**, **EC**, **Flue**, **N-Rep** and **Cons** with a difference of $\langle 0.75, 0.68, 0.47, 0.57 \rangle$, $\langle 0.83, 0.75, 0.54, 0.47 \rangle$, $\langle 0.93, 0.83, 0.52, 0.55 \rangle$, $\langle 0.94, 0.88, 0.52, 0.49 \rangle$, $\langle 1.04, 0.73, 0.27, 0.40 \rangle$, $\langle 0.78, 0.65, 0.58, 0.49 \rangle$, and $\langle 1.01, 0.83, 0.41, 0.49 \rangle$, respectively on *EPE-enEIH*. The significance of all eight rewards is evident in generating accurate, natural, non-repetitive, emotion-adaptive, polite, and empathetic utterances. This indicates that the inclusion of preference and generic rewards contributes to the ability of **PAL** to establish a rap-

port with the client by generating engaging and interactive responses. A detailed discussion with interaction examples is given in Section A.5 of the Appendix.

6.3 Error Analysis

Our proposed model **PAL** performs well in terms of both preference and generic metrics but it is also seen that scores of **EPC** = 73.7% and **EEC** = 75.9% in automatic evaluation can be better. It can be due to the fact that **PAL** is trained as a dialogue model and due to finite memory capacity, it may not be able to memorize the old information and drop it with the increase in context as per turns. Similarly, low and high scores of **PPL** on *EPE-enEIH* and *EPE-HLCC*, respectively, can be due to the low and high variance in the datasets. As *EPE-enEIH* dataset is focused on a fixed topic of harassment of women and children, it has low variance, whereas *EPE-HLCC* dataset utterances have been taken from the transcripts of YouTube conversations between different clients and counselors covering a wide range of topics, resulting in high variance.

7 Conclusion and Future Direction

In this work, we proposed an emotion-adaptive polite and empathetic counseling conversational agent **PAL** to offer higher engagement in e-counseling sessions and resolve the issues faced by the clients. To this end, we prepared two counseling conversational datasets: *EPE-enEIH* and *EPE-HLCC* annotated with emotion, politeness and empathy labels. These datasets are used to train **PAL** in an RL framework. A novel reward function is designed to ensure the preferences of emotion-adaptive politeness and empathy in generated responses. Our extensive automatic and human evaluation results conclude the requirement and effectiveness of designed rewards considering all three aspects - emotion, politeness and empathy over supervised fine-tuned based models. It is found that **PAL** is able to efficiently generate engaging emotion-adaptive polite and empathetic responses. In future, we would like to incorporate domain and common-sense knowledge into the responses to facilitate more realistic and engaging counseling conversations.

Limitations

Our proposed model has also some limitations. First, the requirement of large memory power of

GPU (here, 40 GB) due to the use of GPT-2-medium in the training of **PAL**. Further, weight optimization for each of the possible combinations of different rewards may lead to model training and validation time to months. Hence, some heuristic is adopted to choose some sets of combinations of reward weights. In case of continuous, short and direct responses during interaction like ‘Yes’, ‘I don’t know’, ‘No’, ‘2’, ‘Yeah’, the system first tries to counsel client by inquiring about their issue but after three or four turns it starts deviating and may generate repetitive or inconsistent responses. This can be due to the fact that the datasets which are used to train the **PAL** mostly consists of interactive dialogues with long utterances, hence model gets confused when treated with short and direct responses. Lastly, it is also observed that sometimes, model asks too many questions to the user. This may dissatisfy the user. Hence, the model should be forced to generate only relevant inquiries by discriminating the irrelevant inquiries. This opens up the door for future studies to build a counseling dialogue system.

Ethics Statement

We recognize that there are substantial stakes involved in developing computational models for use in counseling applications and this fact necessitates careful attention to ethical issues. It is crucial for a counseling dialogue system to render support and suggestions to the clients while maintaining their privacy. Thus, proper care has been taken to maintain the client’s privacy. Despite the fact that we have used publicly available counseling conversational datasets, the annotators pledged not to contact or deanonymize any of the users or share the data with others. Further, it is to be noted that this paper does not make any therapy recommendations or clinical diagnostic claims.

Acknowledgements

Kshiti Mishra acknowledges the support received from the Indian government’s “Prime Minister’s Research Fellowship (PMRF) Program”. Priyanshu Priya acknowledges the financial support provided by the Department of Science and Technology, Ministry of Science and Technology, Government of India, through the Innovation in Science Pursuit for Inspired Research (INSPIRE) Fellowship. The authors also acknowledge partial support from the project titled "An Empathetic Knowl-

edge Grounded Conversational System for Mental Health Counseling and Legal Assistance," sponsored by IHUB Anubhuti, TIH, IIT Delhi.

References

- Satanjeev Banerjee and Alon Lavie. 2005. Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, pages 65–72.
- Timothy W Bickmore and Rosalind W Picard. 2004. Towards caring machines. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1489–1492.
- Penelope Brown, Stephen C Levinson, and Stephen C Levinson. 1987. *Politeness: Some universals in language usage*, volume 4. Cambridge university press.
- Peter F Brown, Stephen A Della Pietra, Vincent J Della Pietra, Jennifer C Lai, and Robert L Mercer. 1992. An estimate of an upper bound for the entropy of english. *Computational Linguistics*, 18(1):31–40.
- Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. In *Proceedings of ACL*.
- George Doddington. 2002. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the second international conference on Human Language Technology Research*, pages 138–145.
- Susan V Eisen, Deborah J Youngman, Mollie C Grob, and Diana L Dill. 1992. Alcohol, drugs, and psychiatric disorders: A current view of hospitalized adolescents. *Journal of Adolescent Research*, 7(2):250–265.
- Robert Elliott, Arthur C Bohart, Jeanne C Watson, and David Murphy. 2018. Therapist empathy and client outcome: An updated meta-analysis. *Psychotherapy*, 55(4):399.
- Bo Feng, Siyue Li, and Na Li. 2016. Is a profile worth a thousand words? how online support-seeker's profile features may influence the quality of received support messages. *Communication Research*, 43(2):253–276.
- Asma Ghandeharioun, Daniel McDuff, Mary Czerwinski, and Kael Rowan. 2019. Towards understanding emotional intelligence for behavior change chatbots. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 8–14. IEEE.
- James Gibson, Dogan Can, Bo Xiao, Zac E Imel, David C Atkins, Panayiotis Georgiou, and Shrikanth Narayanan. 2016. A deep learning approach to modeling empathy in addiction counseling. *Commitment*, 111(2016):21.
- Emily A Holmes, Ata Ghaderi, Catherine J Harmer, Paul G Ramchandani, Pim Cuijpers, Anthony P Morrison, Jonathan P Roiser, Claudi LH Bockting, Rory C O'Connor, Roz Shafran, et al. 2018. The lancet psychiatry commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry*, 5(3):237–286.
- Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2019. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*.
- Elizabeth M Huffman, Dimitrios I Athanasiadis, Nicholas E Anton, Lindsay A Haskett, Dominique L Doster, Dimitrios Stefanidis, and Nicole K Lee. 2021. How resilient is your team? exploring healthcare providers' well-being during the covid-19 pandemic. *The American Journal of Surgery*, 221(2):277–284.
- Paul Jaccard. 1912. The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2):37–50.
- Neil Katz and Kevin McNulty. 1994. Reflective listening. *Retrieved February*, 4:2021.
- Dean G Kilpatrick and Ron Acierno. 2003. Mental health needs of crime victims: Epidemiology and outcomes. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, 16(2):119–132.
- Dean G Kilpatrick, Connie L Best, Lois J Veronen, Angelynne E Amick, Lorenz A Villepontoux, and Gary A Ruff. 1985. Mental health correlates of criminal victimization: a random community survey. *Journal of consulting and clinical psychology*, 53(6):866.
- Junhan Kim, Yoojung Kim, Byungjoon Kim, Sukyung Yun, Minjoon Kim, and Joongseek Lee. 2018. Can a machine tend to teenagers' emotional needs? a study with conversational agents. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–6.
- Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.
- Fei-Tzin Lee, Derrick Hull, Jacob Levine, Bonnie Ray, and Kathleen McKeown. 2019. Identifying therapist conversational actions across diverse psychotherapeutic approaches. In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, pages 12–23.
- Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2019. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.

- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Ilya Loshchilov and Frank Hutter. 2018. Decoupled weight decay regularization. In *International Conference on Learning Representations*.
- Gale M Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. 2014. It’s only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37:94–100.
- Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.
- W Mieszczewicz-Kowszewicz, K Warpechowski, K Zieliński, R Nielek, and A Wierzbicki. 2022. Tell me how you feel: Designing emotion-aware voicebots to ease pandemic anxiety in aging citizens. *arXiv preprint arXiv:2207.10828*.
- Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal. 2022a. Please be polite: Towards building a politeness adaptive dialogue system for goal-oriented conversations. *Neurocomputing*, 494:242–254.
- Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal. 2023. Genpads: Reinforcing politeness in an end-to-end dialogue system. *Plos one*, 18(1):e0278323.
- Kshitij Mishra, Azlaan Mustafa Samad, Palak Totala, and Asif Ekbal. 2022b. Pepds: A polite and empathetic persuasive dialogue system for charity donation. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 424–440.
- Robert R Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen M Schueller. 2018. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. *Journal of medical Internet research*, 20(6):e10148.
- Joseph Newbold, Gavin Doherty, Sean Rintel, and Anja Thieme. 2019. Politeness strategies in the design of voice agents for mental health.
- John C Norcross. 2002. *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. Oxford University Press.
- World Health Organization et al. 2021. Mental health atlas 2020.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.
- Mario R Paredes, Vanessa Apaolaza, Cristóbal Fernandez-Robin, Patrick Hartmann, and Diego Yañez-Martinez. 2021. The impact of the covid-19 pandemic on subjective mental well-being: The interplay of perceived threat, future anxiety and resilience. *Personality and Individual Differences*, 170:110455.
- Verónica Pérez-Rosas, Xinyi Wu, Kenneth Resnicow, and Rada Mihalcea. 2019. What makes a good counselor? learning to distinguish between high-quality and low-quality counseling conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 926–935.
- Yada Pruksachatkun, Sachin R Pendse, and Amit Sharma. 2019. Moments of change: Analyzing peer-based cognitive support in online mental health forums. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–13.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. **Towards empathetic open-domain conversation models: A new benchmark and dataset**. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5370–5381, Florence, Italy. Association for Computational Linguistics.
- Elliot Robert, Arthur C Bohart, JC Watson, and LS Greenberg. 2011. Empathy. *Psychotherapy*, 48(1):43–49.
- Stephen Ross and Eric Peselow. 2012. Co-occurring psychotic and addictive disorders: neurobiology and diagnosis. *Clinical neuropharmacology*, 35(5):235–243.
- Tulika Saha, Vaibhav Gakhreja, Anindya Sundar Das, Souhitya Chakraborty, and Sriparna Saha. 2022a. Towards motivational and empathetic response generation in online mental health support. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2650–2656.
- Tulika Saha, Saichethan Reddy, Anindya Das, Sriparna Saha, and Pushpak Bhattacharyya. 2022b. A shoulder to cry on: Towards a motivational virtual assistant for assuaging mental agony. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2436–2449.
- Azlaan Mustafa Samad, Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal. 2022. Empathetic persuasion: Reinforcing empathy and persuasiveness in dialogue systems. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 844–856.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Ashish Sharma, Monojit Choudhury, Tim Althoff, and Amit Sharma. 2020a. Engagement patterns of peer-to-peer interactions on mental health platforms. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pages 614–625.

Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. 2021. Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference 2021*, pages 194–205.

Ashish Sharma, Adam S Miner, David C Atkins, and Tim Althoff. 2020b. A computational approach to understanding empathy expressed in text-based mental health support. *arXiv preprint arXiv:2009.08441*.

Gopendra Vikram Singh, Priyanshu Priya, Mauajama Firdaus, Asif Ekbal, and Pushpak Bhattacharyya. 2022. **EmoInHindi: A multi-label emotion and intensity annotated dataset in Hindi for emotion recognition in dialogues**. In *Proceedings of the Thirtieth Language Resources and Evaluation Conference*, pages 5829–5837, Marseille, France. European Language Resources Association.

David Wadden, Tal August, Qisheng Li, and Tim Althoff. 2021. The effect of moderation on online mental health conversations. *ICWSM*.

Bernard L Welch. 1947. The generalization of ‘student’s’ problem when several different population variances are involved. *Biometrika*, 34(1-2):28–35.

John S White and Theresa A O’Connell. 1996. Adaptation of the darpa machine translation evaluation paradigm to end-to-end systems. In *Conference of the Association for Machine Translation in the Americas*.

Qingyang Wu, Yichi Zhang, Yu Li, and Zhou Yu. 2021. Alternating recurrent dialog model with large-scale pre-trained language models. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1292–1301.

Justine Zhang and Cristian Danescu-Niculescu-Mizil. 2020. Balancing objectives in counseling conversations: Advancing forwards or looking backwards. *arXiv preprint arXiv:2005.04245*.

Justine Zhang, Robert Filbin, Christine Morrison, Jaclyn Weiser, and Cristian Danescu-Niculescu-Mizil. 2019. Finding your voice: The linguistic development of mental health counselors. *arXiv preprint arXiv:1906.07194*.

A Appendix

A.1 Errors in English Translation of EmoInHindi

The translation of Hindi utterances to English causes several errors. Some of them are illustrated in Figure 4. One of the errors that we observe is the translation of Hindi phrases and idioms into English. For instance, as shown in example 1 in Figure 4, the text “*ye bhains ke aage bin bajaana jaisa h*” is translated to “*it’s like barking in front of a buffalo*”. Also, we observe that few named entities are translated into English as shown in Example 4, the named entity “*Rakshak*” is translated into “*keeper*” or “*protector*”. These errors are manually post-edited with the help of human translators to create a high-quality dataset.

A.2 RL Policy and PPO Loss

RL policy is a probability mapping function \mathcal{P}_θ predicting a feasible response pr with t tokens (an action) for a given context (a state).

$$\mathcal{P}_\theta(pr_{1:t}|x) = \prod_{k=0}^t \mathcal{P}_\theta(gr_k|y_{<k}, x) \quad (10)$$

As only minor preferred *meta-communicative* language aspects (politeness and empathy) updates are needed keeping informative content same, PPO loss is selected to update the policy which results in low variance with respect to the old policy. Policy optimization can be formulated in three steps. First, gradient ascent is applied on loss function $J(\theta)$ to maximize the expected reward:

$$\nabla_\theta J(\theta) = \mathbb{E}_{r \sim \mathcal{P}_\theta} [\nabla_\theta \log \mathcal{P}_\theta(r) \hat{A}_R] \quad (11)$$

Second, large deviations are restricted by replacing the *log* term with an importance sampling term and catastrophic forgetting is prevented using clipping:

$$L^{\text{CLIP}}(\theta) = \hat{\mathbb{E}}[\min(pr_r(\theta) \hat{A}_r, \text{clip}(pr_y(\theta), 1-\varepsilon, 1+\varepsilon) \hat{A}_r)] \quad (12)$$

Here, the probability ratio between the new and old policies is calculated as $pr_r(\theta) = \mathcal{P}_\theta^{\text{new}} / \mathcal{P}_\theta^{\text{old}}$. ε and \hat{A}_y gives the clipping range and estimated advantage (normalized rewards), respectively. Finally, parameter updating is done as follows:

$$\theta_{k+1} = \underset{\theta}{\text{argmax}} \mathbb{E}_{s, a \sim \mathcal{P}_{\theta_k}} [L^{\text{CLIP}}] \quad (13)$$

1.	Original (Hindi)	हां, मैंने अपने परिवार के साथ इस पर चर्चा की है, लेकिन ये भैंस के आगे बीन बजाने जैसा है हर कोई मुझसे कहता है कि बाहर जाना बंद करो और अपना व्हाट्सएप हटा दो।
	Translated (English)	Yes, I have discussed it with my family, but it's like barking in front of a buffalo , everyone tells me to stop going out and delete my WhatsApp.
	Post-edited (English)	Yes, I have discussed it with my family, but it's totally useless , everyone tells me to stop going out and delete my WhatsApp.
2.	Original (Hindi)	ठीक है, लेकिन तुम मुझे परेशान करने की हिम्मत मत करो, मैं पहले से ही सटकी हुई हूँ।
	Translated (English)	Okay, but don't you dare bother me, I'm already hooked .
	Post-edited (English)	Okay, but don't you dare annoy me, I'm already in a bad mood .
3.	Original (Hindi)	नहीं, बिलकुल नहीं।दिमाग मत खाओ अब
	Translated (English)	No, not at all. Don't eat your brains now
	Post-edited (English)	No, not at all. Get lost now .
4.	Original (Hindi)	अगर आपको कोई चिंता है तो कृपया रक्षक से संपर्क करने में संकोच न करें। रक्षक से संपर्क करने के लिए धन्यवाद। आपका दिन अच्छा रहे !
	Translated (English)	Please don't hesitate to contact the keeper if you have any concerns. Thank you for contacting Protector . Have a nice day!
	Post-edited (English)	Please don't hesitate to contact the Rakshak if you have any concerns. Thank you for contacting Rakshak . Have a nice day!
5.	Original (Hindi)	पुलिस यह कहते हुए शिकायत दर्ज नहीं कर रही है कि वे इस प्रकार की शिकायत दर्ज करने के लिए अधिकृत नहीं हैं। मेरा मतलब है क्या सच में वे मेरी शिकायत दर्ज नहीं कर सकते कमीने। वो बस टहला रहे हैं।
	Translated (English)	Police is not registering complaint saying that they are not authorized to register such complaint. I mean really they can't register my complaint you bastard. They are just taking a walk .
	Post-edited (English)	Police are not lodging complaints stating that they are not authorized to lodge this type of complaint . I mean seriously they can't lodge my complaint , bastards. They are just killing my time .

Figure 4: Errors in translation of EmoInHindi dataset

A.3 Implementation Details

All classifiers are built by fine-tuning the RoBERTa-large (Liu et al., 2019) model. To train generation preference models, BART-large (Lewis et al., 2019) having a bidirectional (BERT-like) encoder and an autoregressive (GPT-like) decoder is used. The Language model GPT2-medium (Radford et al., 2019), ARDM (Wu et al., 2021) and EIDM are trained in a supervised learning setting with AdamW optimizer (Loshchilov and Hutter, 2018), learning rate $\alpha = 2e^{-05}$ and $seed_value = 10$. Fine-tuning of trained EIDM is done in an RL setting by experimenting with the different number of candidate responses i.e., $n = 2, 3, 4, 5, 10$. $n = 3$ is selected as the final value as per the loss obtained. The generated candidates are decoded using nucleus sampling (Holtzman et al., 2019) with temperature $\mathbb{T} = 0.8$ and probability $p = 0.9$.

The proposed system PAL is trained considering the $seed_value = 10$, $human_reward = 10$, $max_candidate_length = 50$, and AdamW optimizer (Loshchilov and Hutter, 2018) with a learning rate of $\alpha = 2e^{-05}$, $\varepsilon = 0.2$ and $epochs = 20$. The reward weight combination of 0.1, 0.1, 0.2, 0.2, 0.1, 0.1, 0.1, 0.1 are chosen as the final weights for $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$, and β_8 , respectively. Lastly, for R_1, R_2, R_3, R_4, R_5 , and R_6 , the penalization factors α, γ and δ are set to 2.

A.3.1 Hardware Configuration

The device configurations that we use for the experiment are: (a) GPU: A100-PCIE-40GB, (b) CUDA Support: CUDA 11.x (or later.) (c) GPU clocks: Base: 765 MHz, Boost: 1410 MHz (d) Memory Size: 40 GB (e) Memory Type: HBM2 (f) Memory clock: 1215 MHz (g) Bus Width: 5120 bits (h) Total board power: 250 W.

A.4 Weight Optimization

We experimented with the various combination of weights ($\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$ and β_8) to find the correct combination of weights for the reward function. We validate these rewards on 10% hold-out data and finally, the combination of weights that gives the highest perplexity score is selected to train PAL. Table 8 shows the weights considered for optimization with the dataset EPE-enEIH. It can be inferred from the table that a better perplexity score is obtained when all the rewards are considered. Further, we observe that when any reward is eliminated, the perplexity score is dropped, which, in turn, demonstrates the significance of each reward.

A.5 Discussion

A few examples of client’s interaction with our proposed system PAL and the baseline model EIDM is shown in Table 9 and Table 10. We notice from the examples that the user gave a better score to our

WEIGHT OPTIMIZATION								
β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	PPL
0.6	0.1	0	0.1	0.1	0	0	0.1	2.412
0	0.7	0	0	0.1	0.1	0	0.1	0 2.533
0	0	0	0.6	0.1	0.1	0.1	0.1	0 2.556
0.1	0.1	0.1	0.6	0	0.1	0	0	2.523
0	0	0	0	0	0	0	0	2.8734
0.1	0.1	0.15	0.1	0.2	0.15	0.1	0.1	2.289
0.1	0.1	0.1	0.3	0.2	0.1	0	0.1	2.346
0.2	0.1	0.1	0.2	0.1	0.2	0	0.1	0 2.1865
0.15	0.15	0.1	0.15	0.15	0.1	0.1	0.1	2.113
0.3	0.3	0.3	0.1	0	0	0	0	2.312
0.2	0.2	0.2	0.1	0.1	0.1	0	0.1	2.10
0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1	2.0312

Table 8: Weight optimisation using different reward weight combination

model **PAL** than *EIDM* in terms of all human evaluation metrics, viz. **EPC, EEC, PC, EC, Flue, N-Rep** and **Cons**. Thus, it can be concluded that the proposed system is able to incorporate politeness and empathy in the generated responses depending on the client’s underlying emotion, thereby showcasing friendliness and a caring attitude towards the client. Also, these results qualitatively show that **PAL** is able to generate more fluent, coherent, and engaging responses.

However, some limitations are also observed in the responses generated by **PAL**. For example, in Table 9, for the client’s utterance like “*I am probably not going to stop smoking; I haven’t even considered it, and until today’s fantastic news at work, I had no intention of doing so.*”, the model again responds with *Okay, so where do you put yourself on a scale of 1 to 10?*. This may be due to the fact that our designed reward function is unable to cover the significant aspects of common-sense knowledge that can probably map “no intention to quit” with the being on “1” on the scale of 1 to 10. Moreover, the counselor/agent also needs some domain knowledge to generate more informative response instead of generic one. For instance, in the counselor’s utterance “*That’s really great! Well, there are now a lot of products available that can help you fight the urge to smoke. For example, there are products that can help you replace nicotine, which can help ease the withdrawal symptoms you might feel when quitting or cutting down. You could also have some medications that might take away even some of the desire to smoke. You can look ahead to whichever is feasible for you.*”, a better response would be suggesting some appropriate medication like “Chantix is a medication for smoke cessation that you might consider if you want to quit.” instead of vaguely saying “You could also have some medications that might take away even some of the

desire to smoke.”. We may need relevant domain and/or common-sense knowledge to generate more realistic responses in counseling conversations.

Interlocutor	Utterances generated by our model - PAL (EPC = 4, EEC = 4, PC = 5, EC = 5, Flue = 5, N-Rep = 4, Cons = 4)
Counselor	Hello! How are you doing today?
Client	Not so good.
Counselor	Really sorry to hear this! I am wondering what made you feel bad, let me know so I can help you.
Client	Yeah, I hope so! Um... my workplace is just gone smoke-free. I'm not not very happy about it, I'm going to keep smoking but I was worrying if there's anything that can help me get through the day - it's like you know go get my beer and smoke cigarettes after work.
Counselor	Oh, I see! Your workplace just went smoke-free and this is what bothering you, my dear.
Client	Yes. I am not going to quit smoking. I like it very much, it is something that gives me solace, I feel relaxed. I don't know what's a big deal if anyone smoke or if I like smoking.
Counselor	Please don't panic! We will together figure out something for you. Can you tell a little bit about how smoking fits into your day, perhaps it will help.
Client	I suppose a little less than a pack a day I probably smoked for the last five years it's not too bad. I mean I could quit if I wanted to but I don't really want to right now.
Counselor	Okay, so you feel like right now smoking is a pretty important part of your life it's not something you're ready to change but if the time came when you did decide to quit you feel pretty confident that you could go ahead and do that.
Client	Of course, I can quit.
Counselor	Wonderful! Could you tell what makes you so confident?
Client	Well I mean I don't see why I couldn't although I have never attempted to quit. It's simply something I do for enjoyment during break time in the morning and when I have to work in the afternoon.
Counselor	Mm-hmm, I see! So, in terms of a readiness to actually go ahead and make a quit attempt where do you think you fall on a scale of one to ten where one is I'm not considering it at all and ten is I'm definitely ready I want to quit.
Client	I am probably not going to stop smoking; I haven't even considered it, and until today's fantastic news at work, I had no intention of doing so.
Counselor	Okay, so where do you put yourself on a scale of 1 to 10?
Client	That's what I am saying, I am on one.
Counselor	Relax, my friend! Well, I would love to know if you have ever thought about what could make you say two or three instead of one.
Client	Actually anything is conceivable, so I suppose it's feasible that I might be persuaded to stop smoking if there is anything I could do that would make me quit.
Counselor	That's really great! Well, there are now a lot of products available that can help you fight the urge to smoke. For example, there are products that can help you replace nicotine, which can help ease the withdrawal symptoms you might feel when quitting or cutting down. You could also have some medications that might take away even some of the desire to smoke. You can look ahead to whichever is feasible for you.
Client	Well, this seems useful information. I'm not going to stop smoking today; I just needed to know my alternatives. However, it's reassuring to know that if I do decide to quit, I do have resources to assist me.
Counselor	Yes, and back to that talent of yours, you stated you could quit if you want... What do you believe it would take to raise you to a four or five?
Client	I don't know exactly but I suppose if one of my friends chose to quit with me it would probably help too, but I and all my friends love to smoke, so I don't see that happening any time soon.
Counselor	Sure, I wonder if now that your workplace is going smoke-free if there will be some colleagues at work who are planning to quit.
Client	That's a good point, I hadn't really thought about that. That's a good idea.
Counselor	We are glad that you liked our suggestion. Anything else that I could do for you?
Client	Not as such.
Counselor	Okay, well it sounds like you're really here today just for information. I'm happy to help you pick out some of these products if and when you decide that the time is right for you.
Client	Sure, I would let you know.
Counselor	Yes, please feel free to contact if you need any help. We are here for you 24x7. Take care and have a wonderful day!
Interlocutor	Utterances generated by the baseline model - EIDM (EPC = 3, EEC = 3, PC = 2, EC = 3, Flue = 4, N-Rep = 4, Cons = 3)
Counselor	Hey! What's up?
Client	Hi, not good.
Counselor	What made you feel bad?
Client	My workplace is just gone smoke-free, I'm not very happy about it, I'm going to keep smoking but I was worrying if there's anything that can help me to get through the day like get some beer or smoke after work.
Counselor	So your workplace just went smoke free you're not too pleased huh
Client	Yes, I'm not going to quit smoking oh my god I like it if it helps me relax so I mean I don't see what the big deal is.
Counselor	Okay, tell me a little about your smoking pattern during the day, maybe I can help?
Client	Less than a pack, I can quit if I want to but not right now.
Counselor	What makes you so confident?
Client	Nothing as such but I think I can.
Counselor	On a scale of one to ten, where do you think you are in terms of being ready to quit smoking?
Client	Maybe 2.
Counselor	okay, what made you say 2?
Client	Well, anything is possible. I mean if there're things I can try that would make me like want to quit smoking.
Counselor	There are a lot of products available now that can actually help you curb that urge to smoke.
Client	That's good!
Counselor	Yes. Well, I think you're here only for the information today.
Client	Yes, thanks for the info.
Counselor	My pleasure! Thank you for reaching out to us.

Table 9: Example1: Client (substance addict) interaction with our proposed model **PAL** and the baseline *EIDM*. Orange and blue texts show polite and empathetic responses, respectively.

Interlocutor	Utterances generated by our model - PAL (EPC = 5, EEC = 4, PC = 4, EC = 4, Flue = 5, N-Rep = 5, Cons = 4)
Counselor	Hello there! You are connected to Rakshak. How may I assist you today?
Client	I am mentally disturbed and getting suicidal thoughts.
Counselor	Stay calm, my dear, suicide is not a solution to any problem. You are not alone. Before moving ahead, would you like to share your name and age?
Client	I can't.
Counselor	It's not a problem at all, Rakshak is always there to provide assistance whenever it's needed. Could you kindly explain the difficulty that you're having?
Client	Someone is sending me offensive messages on Facebook while claiming to be my actual ex-boyfriend. The jerk is using my ex-boyfriend's ID and sending messages from that account. But the fact is he passed away due to COVID around two months ago.
Counselor	Oh!! I am really saddened to learn this. May I inquire as to the nature of the offensive messages that you are receiving?
Client	That jerk has morphed my photos with my ex-boyfriend in the wrong way and blackmailing me.
Counselor	Please have some patience with me, as I completely understand the frustration you are feeling. Could you kindly inform me whether you have any clue as to who could be engaged in this?
Client	One of our friends, a close friend of my ex, loved me. He may be responsible.
Counselor	Why would he do this?
Client	He wants to marry me despite my constant rejection.
Counselor	Really sad, I understand it's a difficult time for you but I assure you everything will be fine soon. May I know when this problem started?
Client	About two weeks ago.
Counselor	I'm sorry for your week-long struggle, wishing you well. Let me know if you require in-person counselling.
Client	Yes, please. that would be great.
Counselor	Wonderful! Please feel free to contact the National Commission for Women helpline number 7827-170-170 anytime for seeking help or the Women-in-distress helpline number 1091.
Client	I will, thanks for the help.
Counselor	It's my pleasure, anything else that I can do for you?
Client	That's all. It was nice talking to you.
Counselor	I'm pleased that we might be of assistance. It is much appreciated that you contacted Rakshak. Stay safe and healthy!
Interlocutor	Utterances generated by the baseline model - EIDM (EPC = 2, EEC = 2, PC = 1, EC = 2, Flue = 4, N-Rep = 4, Cons = 3)
Counselor	Hello! Is there anything I can do for you?
Client	Yes, I am depressed and having suicidal feelings.
Counselor	Why are you feeling depressed?
Client	Someone is sending me vulgar messages on Facebook while claiming to be my actual ex-boyfriend. The jerk is using my ex-boyfriend's ID and sending messages from that account. But the fact is he passed away due to COVID around two months ago.
Counselor	What kind of vulgar messages you are getting?
Client	What do you mean what kind of vulgar, vulgar means vulgar.
Counselor	I want to know exactly what kind of messages you are receiving to help you better.
Client	It doesn't seem you want to help.
Counselor	It's not like that. I really want to help.
Client	Let it be.
Counselor	Thanks and visit again if you need anything. Take care!

Table 10: Example1: Client (crime victim) interaction with our proposed model **PAL** and the baseline **EIDM**. Orange and blue texts show polite and empathetic responses, respectively.

ACL 2023 Responsible NLP Checklist

A For every submission:

A1. Did you describe the limitations of your work?

8

A2. Did you discuss any potential risks of your work?

9

A3. Do the abstract and introduction summarize the paper's main claims?

1

A4. Have you used AI writing assistants when working on this paper?

Left blank.

B Did you use or create scientific artifacts?

3

B1. Did you cite the creators of artifacts you used?

3

B2. Did you discuss the license or terms for use and / or distribution of any artifacts?

9

B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?

Not applicable. Left blank.

B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?

9

B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?

3

B6. Did you report relevant statistics like the number of examples, details of train / test / dev splits, etc. for the data that you used / created? Even for commonly-used benchmark datasets, include the number of examples in train / validation / test splits, as these provide necessary context for a reader to understand experimental results. For example, small differences in accuracy on large test sets may be significant, while on small test sets they may not be.

3

C Did you run computational experiments?

4

C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?

No response.

The Responsible NLP Checklist used at ACL 2023 is adopted from NAACL 2022, with the addition of a question on AI writing assistance.

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

5

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

6

- C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?

Not applicable. Left blank.

D Did you use human annotators (e.g., crowdworkers) or research with human participants?

3

- D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

3

- D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

3

- D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating? For example, if you collected data via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?

Not applicable. Left blank.

- D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

Not applicable. Left blank.

- D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?

3