# *Help Me Heal*: A Reinforced Polite and Empathetic Mental Health and Legal Counseling Dialogue System for Crime Victims

**Kshitij Mishra*, Priyanshu Priya*, Asif Ekbal**

Department of Computer Science and Engineering, Indian Institute of Technology Patna, Bihta, India
{kshitij_1921cs23, priyanshu_2021cs26, asif}@iitp.ac.in

## Abstract

The potential for conversational agents offering mental health and legal counseling in an autonomous, interactive, and vitally accessible environment is getting highlighted due to the increased access to information through the internet and mobile devices. A counseling conversational agent should be able to offer higher engagement mimicking the real-time counseling sessions. The ability to empathize or comprehend and feel another person's emotions and experiences is a crucial quality that promotes effective therapeutic bonding and rapport-building. Further, the use of polite encoded language in the counseling reflects the nobility and creates a familiar, warm, and comfortable atmosphere to resolve human issues. Therefore, focusing on these two aspects, we propose a Polite and Empathetic Mental Health and Legal Counseling Dialogue System (**Po-Em-MHLCDS**) for the victims of crimes. To build **Po-Em-MHLCDS**, we first create a Mental Health and Legal Counseling Dataset (**MHLCD**) by recruiting six employees who are asked to converse with each other, acting as a victim and the agent interchangeably following a fixed stated guidelines. Second, the **MHLCD** dataset is annotated with three informative labels, *viz.* counseling strategies, politeness, and empathy. Lastly, we train the **Po-Em-MHLCDS** in a reinforcement learning framework by designing an efficient and effective reward function to reinforce correct counseling strategy, politeness and empathy while maintaining contextual-coherence and non-repetitiveness in the generated responses. Our extensive automatic and human evaluation demonstrate the strength of the proposed system. Codes and Data can be accessed at https://www.iitp.ac.in/ ai-nlp-ml/resources.html#MHLCD or https://github.com/Mishrakshitij/Po-Em-MHLCDS

## Introduction

With approximately 20% of the global population suffering from mental health issues (Holmes et al. 2018), availability of its immediate treatment has become the primary concern worldwide. However, according to the WHO's Mental Health ATLAS 2020[1], there is a global paucity of mental health professionals. Limited access to in-person treatment and other obstacles like social stigma and prejudice (White and Dorman 2001) cause millions of individuals who need emotional and mental health-related support (Eysenbach et al. 2004) to resort to text-based peer support forums such as TalkLife, Psychcentral[2], etc. Although the peer supporters on such forums are intended to help support seekers positively, they are usually untrained and unacquainted of the best practices in counseling, hence, are unable to deliver good and mutually engaging replies (Gage-Bouchard et al. 2018). Consequently, it becomes imperative to build intelligent conversational agents that can offer potential solutions to support seekers.

A polite behaviour of a dialogue agent facilitates natural, smooth, and engaging conversations (Coppock 2005) and can foster the interaction between the agent and users (Golchha et al. 2019). Politeness has been identified and evaluated as an essential component of support messages (Brown, Levinson, and Levinson 1987; Feng, Li, and Li 2016). Incorporation of politeness in a counseling agent's replies may create a warm atmosphere and enhance the user experience. Furthermore, to administer successful counseling and elicit positive-outcomes in support-based conversations, empathy is vital (Norcross 2002; Elliott et al. 2018) and has been substantially investigated.

The World Health Organization (WHO) states that approximately one-third (33.3%) of women worldwide have experienced physical and/or sexual violence at least once in their lifetime, and approximately one billion children aged 2 to 17 have experienced physical, sexual, or emotional violence. This affects their physical, emotional, and mental well-being. Further, owing to the lack of knowledge about their legal and human rights or their privacy, most people are less likely to disclose an assault or abuse. An immediately available secure mental health and legal counseling dialogue system may able to help these people (*victims* hereafter).

Though a few dialogue systems for mental health assistance (*WoeBot* (Fitzpatrick, Darcy, and Vierhile 2017), *Tess* (Fulmer et al. 2018)) and legal assistance (*DoNotPay*, *Convey Law*[3]) have been reported in the literature, none of these is designated to provide both mental health and legal counseling to the *victims*. Neither do they induce politeness in the responses while showcasing empathy. Thus, we propose

*These authors contributed equally.



[1]https://www.who.int/publications/i/item/9789240036703

[2]https://www.talklife.co/, https://psychcentral.com/

[3]https://donotpay.com/, https://www.conveylaw.com/

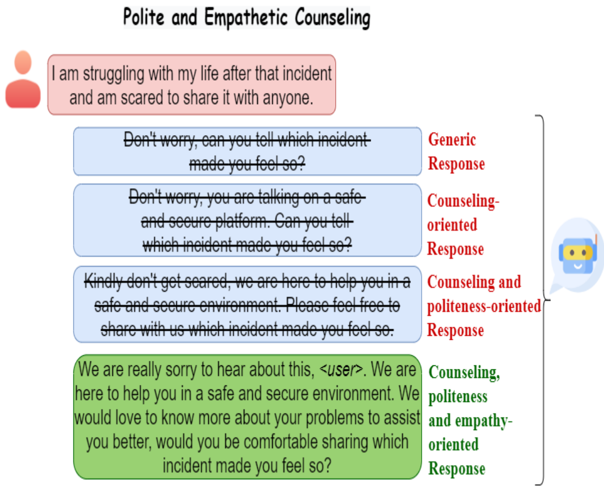


Figure 1: Example showcasing the use of politeness and empathy (compassionate emotion) in counseling

here a novel research direction of inducing both politeness and empathy in a counseling dialogue system's responses whilst being contextually consistent.

A counseling dialogue system should be able to employ different *counseling strategies* as per ongoing conversation and user's state. Further, it should create a familiar, warm and comfortable environment by employing *polite* tone in its responses. Lastly, to console the users and gain their trust, it should be able to make an *empathetic* connection with the user. For instance, in Figure 1, although the agent's responses in blue boxes try to counsel the victim, the response in the green box facilitates better user engagement, reflecting a sense of compassion and confidence with the user.

However, owing to the scarcity of available data, developing such a polite and empathetic personalized dialogue agent in a Supervised Learning (SL) framework that can generalize to diverse users and contexts is challenging.

Due to the potential of reinforcement learning (RL) to learn and improve as per some feedbacks received in the form of rewards by interacting with the environment, the research community explore RL-based techniques to build the dialogue systems (Casanueva et al. 2018; Mesgar, Simpson, and Gurevych 2020). Similarly, some research works have focused on adapting politeness in a task-oriented dialogue system's responses (Mishra, Firdaus, and Ekbal 2022), or improving empathy in mental health support dialogues (Sharma et al. 2021; Saha et al. 2022a) in an RL-framework. Therefore, to offer a familiar, warm and comfortable environment, thereby facilitating better user engagement, in our present work, we propose a polite and empathetic mental health and legal counseling dialogue system (**Po-Em-MHLCDS**) based on the RL framework.

Development of the proposed **Po-Em-MHLCDS** has been done in three stages. First, due to non-availability of mental health and legal counseling data, we collect and prepare a Mental Health and Legal Counseling Dataset (**MHLCD**). Further, **MHLCD** is annotated with defined counseling strategies, politeness information and empathy factor in agent's responses. Second, to achieve natural language interaction between counseling dialogue agent and the user, we train a Maximum Likelihood Estimation (MLE) loss based model on **MHLCD** dataset. Third, this MLE-loss based model is fine-tuned in an RL setting by designing a novel reward function to ensure right counseling-strategy, politeness, empathy, contextual-coherence, and non-repetitiveness in the generated responses. Further, RL-policy is optimized in such a way that it maximizes the reward value given by this function. Lastly, the proposed system's performance is assessed in terms of automatic and human evaluation. Thus, our current work has the following key attributes:

(i) Constructed a large-scale mental health and legal counseling dialogue dataset **MHLCD** and manually annotated it with three informative labels *viz.* counseling-strategy, politeness and empathy;
(ii) Built transformers-based robust counseling-strategy, politeness, and empathy classifiers;
(iii) Designed an efficient and effective reward function to generate non-repetitive and contextually-coherent responses with correct counseling-strategy, politeness, empathy imbibed in it;
(iv) Proposed a Polite and Empathetic Mental Health and Legal Counseling Dialogue System (**Po-Em-MHLCDS**) in an RL framework utilizing our newly designed reward function. To the best of our knowledge, this is the very first attempt towards this direction;
(v) Performed extensive experimental analysis employing automatic and human evaluation to demonstrate the strength of our proposed system.

## Related Work

There have been several attempts to build end-to-end dialogue systems (Wu, Martinez, and Klyen 2018; Zhong et al. 2022; Zhang et al. 2022) With the increasing concern for the psychological well-being of individuals, recently, efforts have been made to develop dialogue systems for mental health support (Sharma et al. 2020b, 2021; Saha et al. 2022a,b). Past works have explored diagnosing mental health issues from social media posts and activities (Yazdavar et al. 2018; Reis et al. 2019) The existing studies for mental health support are primarily centered on analyzing effective approaches to obtain context-specific adaptation and response diversity (Althoff, Clark, and Leskovec 2016; Pérez-Rosas et al. 2019; Zhang and Danescu-Niculescu-Mizil 2020). Besides, researchers have developed techniques for gauging linguistic development of counselors (Zhang et al. 2019), extracting conversational engagement patterns (Sharma et al. 2020a), analyzing moderation (Wadden et al. 2021), detecting therapeutic actions (Lee et al. 2019), and identifying cognitive restructuring (Pruksachatkun, Pendse, and Sharma 2019) in supportive talks. Some preliminary studies in legal sphere are also available (John et al. 2017; Do et al. 2017).

The role of empathy in building mental health support conversations has been explored in (Elliott et al. 2011; Castonguay and Hill 2017). The authors in (Morris et al. 2018) utilized a corpus-based approach for constructing nuanced and personalized empathetic responses. Several efforts have been made to understand and build computational methods for identifying empathy in face-to-face therapy (Gibson et al. 2016; Pérez-Rosas et al. 2017), and text-based peer-to-peer support system (Sharma et al. 2020b). The authors in (Sharma et al. 2021; Saha et al. 2022a,b) investigated ways to induce empathy in mental health support conversations. A very few research attempts have explored that the polite or caring behavior of the conversational agent can further improve the sense of empathy and facilitate the revelation of personal information (Lucas et al. 2014; Kim et al. 2018). Existing research works primarily utilized mental health-related posts from different social platforms to collect their conversational datasets. Our work differentiates from these works in four aspects: (i) We propose a novel mental health and legal counseling dialogue dataset created from scratch; (ii) To get a better view of the counseling dialogue, we define a set of counseling strategies, which help the dialogue agent to take next counseling action as per the context available; (iii) To facilitate understanding of victim's emotional state and create a comfortable atmosphere, two *meta-communicative* aspects of language, *viz.* politeness and empathy are also incorporated in the agent's responses; (iv) By designing an efficient reward function, we build a mental health and legal counseling dialogue system in an RL framework. As per our knowledge, this is the first step towards developing such a dialogue system.

## Dataset

To facilitate the development of a polite and empathetic goal-oriented dialogue system for mental health and legal counseling assistance, we create a novel, high-quality, and large-scale mental health and legal counseling conversational dataset, named **MHLCD** annotated with appropriate counseling strategy, politeness and empathy labels. **MHLCD** dataset statistics are shown in Table 1.

### Data Preparation

The **MHLCD** dataset comprises of dialogues centered on mental health and legal counseling aid for women and children who have been victims of any type of crimes, including domestic violence, rape, acid attacks, physical/cyber stalking, workplace harassment, online harassment, impersonation, trolling, matrimonial fraud, financial fraud, child pornography, women/child trafficking, non-consensual sexting, doxing/outing, and exclusion. A mental health and legal counseling dialogue system should advance support to individuals for good mental and emotional health with authentic information. To assure this, before the data preparation began, several websites *viz.* National Cybercrime Reporting Portal[4], National Commission for Women[5], Ministry of Women and Child Development[6]; and documents, *viz.* Criminal Law Amendment Act 2013[7], IT (Amendment) Act 2008[8] have been referred for providing authentic counseling services and legal assistance. A few other websites containing news items and case studies on crimes against women and children have been explored, which helped us to get real-life stories of such incidents. Now, to prepare the dataset, six employees are recruited and asked to converse with each other following these stories/content/information using the Wizard-of-Oz approach (Kelley 1984) in pairs, where one individual plays the role of the *counselor* (agent) and the other one as a *victim* (user). During the conversation, the two participants were randomly assigned the role of the agent and a victim to eradicate the correlation between the agent's counseling strategies and the targeted victim's characteristics.

### Data Preparation Guidelines

The mental health and legal experts from government-run institutions of national repute were consulted to comprehend the dialogue flows in victims' situations. The interaction with the experts helps in drafting the guidelines for preparing the dyadic conversations between the agent and victims, which are as follows:

(i) The participant playing the agent role first asks for a few basic information about the victims to assess their profile variables.
(ii) The agent should identify the victims' issues and evaluate their immediate psychological requirements.
(iii) During counseling, the agent should be patient, empathetic, and respectful towards the victims in order to boost up their morale and provide a judgement-free environment to share their feelings freely.
(iv) The agent should motivate the victims to lodge the complaint, seek medical care, or contact the support groups/organizations that can help them. If the victim agrees, then provide them with the pertinent and authentic legal, medical, and/or organizational information.
(v) The agent should make the victims aware of the basic safety measures that will eventually help them prevent such unwanted incidents.

### Mental Health and Legal Counseling Strategies

The *counselor* (or agent) can employ different counseling strategies to provide the counseling assistance to the victims. The participants (acting as the agent) are provided with tips on different counseling strategies along with a few example sentences. These different counseling strategies are designed using the counseling theories and a preliminary examination of 60 conversation samples. The three pairs of participants independently prepared 60 conversations, analyzed discrepancies, and modified the strategies accordingly under the supervision of experts. Specifically, 11 different counseling strategies are identified, which are as follows:

- **Problem assessment** refers to finding the answers to five W's and one H, i.e., What has happened?, Who is it

[4]https://cybercrime.gov.in/
[5]http://ncw.nic.in/
[6]https://www.wcd.nic.in/
[7]https://www.iitk.ac.in/wc/data/TheCriminalLaw.pdf
[8]https://www.meity.gov.in/writereaddata/files/itact2000/it_amendment_act2008.pdf

about?, When did it happen?, Where did it take place?, Why did it happen?, and How did it happen?

- **Confidentiality assurance** refers to making the victims believe that the agent respects their privacy and whatever they will share will not be disclosed to anyone.
- **Motivational directive** refers to encouraging the victims to have patience and be optimistic and motivate them to get involved in the things that give them relief and hope.
- **Emotional support** refers to providing a safe and non-judgmental environment along with emotional comfort to the victims to express their feelings.
- **Counseling support** refers to providing information of a professional counselor with whom the victim can talk and get support from professionals.
- **Reassurance** refers to making the victims understand that no one ever deserves to be abused or harassed; it is not their fault and assure them that they are not alone and can always seek support in such situations.
- **Legal awareness** refers to providing law and other relevant legal information to the victims in order to make them aware of their rights and seek justice lawfully.
- **Reporting assistance** refers to providing step-by-step guidance for reporting the assault if the victim wishes to do so.
- **Safety guidance** refers to providing a few safety measures/tips so that the victims can make themselves aware of the crimes and prevent such incidents.
- **Credibility assurance** refers to the use of credentials and citing organizations (governmental or NGOs) to establish credibility and earn the victim's trust.
- **No-Strategy** is designated to the utterances which do not employ any counseling strategy.

### Data Annotation

To annotate the agent's utterances with the correct counseling strategy (eleven classes), politeness (three classes: impolite, neutral, polite) and empathy (two classes: empathetic and non-empathetic), three annotators with post-graduate qualifications and significant experience in the related tasks are recruited. These annotators are briefed with appropriate annotation guidelines and provided with illustrative examples for each of the labels. To reduce the manual efforts, entire annotation process is carried out in two phases. In the first phase, 210 dialogues (with approx. 6K utterances) are manually annotated with the counseling strategy, empathy, and politeness labels as per the guidelines provided. In the second phase, the agent's utterances in the remaining 796 dialogues are passed through fine-tuned transformer-based counseling strategy, politeness and empathy classifiers to predict the respective labels. Then, these predicted labels are cross-verified for their correctness by the same three annotators in order to create a gold-standard dataset. We observe a reliable multi-rater Kappa (McHugh 2012) agreement ratio of 71.2%, 78.6%, and 81.2% in the first phase, and 77.2%, 82.6%, and 84.1% in the second phase for counseling strategy, politeness, and emotion labels, respectively.

| **Metrics** | **Train** | **Validation** | **Test** |
|---|---|---|---|
| *# of Dialogues* | 755 | 100 | 151 |
| *# of Utterances* | 20886 | 2795 | 4163 |
| *Avg. Utterances per Dialogue* | 27.66 | 27.95 | 27.57 |

Table 1: Dataset statistics.

## Proposed System: Po-Em-MHLCDS

To build our proposed system, we first alternatively train two different language models to learn the distribution of the user's (victim) and agent's (counselor) utterances to achieve natural language interaction between both the models. This MLE-loss based agent's trained model is then fine-tuned in an RL framework. In order to optimize the agent's behaviour, we employ an RL-loss considering an efficiently designed reward function to ensure politeness, empathy, contextual coherence and non-repetitiveness while counseling the victim. The RL-loss is optimized to investigate new and potentially better policies to generate agent's responses.

$$R = \alpha_1 R_1 + \alpha_2 R_2 + \alpha_3 R_3 + \alpha_4 R_4 + \alpha_5 R_5 \tag{1}$$

where $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 1$.

### MLE-Loss Based Dialogue Model

A multi-turn mental health and legal counseling dialogue can be represented as $d = \{c_0, v_0, .., c_i, v_i, .., c_{T-1}, v_{T-1}\}$, where $c_i$ denotes the counselor's $i^{th}$ response and $v_i$ denotes the victim's $i^{th}$ response, and $T$ denotes the total number of turns in the dialogue. The agent's (counselor, $c$) and user's (victim, $v$) utterances' distribution are modelled recurrently. In accordance with (Wu et al. 2021), two language models (LMs), here, GPT-2-medium (Radford et al. 2019), one for counselor $p_c$ and one for victim $p_v$ are considered to model the probability distributions over the dialogue $d$'s utterances. These LMs, $p_c$ and $p_v$ try to predict the next best possible token $r_j$ in a response $r = \{r_1, r_2, .., r_j, .., r_t\}$ with $t$ tokens. Given the context of the dialogue, the joint probability of the victim's and the counselor's utterance can be expressed as follows:

$$p_v(v_i|v_{<i}, c_{<i}) = \prod_{j=1}^{t_{v_i}} P(r_j|r_{<j}, v_{<i}, c_{<i}) \tag{2}$$

$$p_c(c_i|v_{<=i}, c_{<i}) = \prod_{j=1}^{t_{c_i}} P(r_j|r_{<j}, v_{<=i}, c_{<i}) \tag{3}$$

Finally, MLE-loss based trained dialogue model is obtained by maximising the likelihood $p_\theta(d)$, defined on dialogue $d$. It can be written as:

$$p_\theta(d) = \prod_{T=0}^{T-1} p_v(v_i|v_{<i}, c_{<i}) p_c(c_i|v_{<=i}, c_{<i}) \tag{4}$$

### RL-Loss Based Fine-Tuning

Once the trained dialogue model, $p_\theta(d)$ providing the natural language interaction between the agent and the victim is obtained, it is fine-tuned with an RL-loss. Given a dialogue context, first, $p_\theta(d)$ generates, *n*-candidate responses. Second, these candidates are quality checked in terms of counseling-strategy correctness, politeness, empathy, contextual-coherence and non-repetitiveness. Lastly, by employing Proximal Policy Optimization (PPO) (Schulman et al. 2017) method, RL-policy is optimized to generate a polite and empathetic counseling response.

**Classifiers** To formulate the counseling strategy, politeness, and empathy reward feedbacks, respective classifiers are needed. Additionally, to evaluate the presence or absence of counseling strategy in the generated response, a binary counseling classifier is required. All these four classifiers are built by fine-tuning a pre-trained RoBERTa-large (Liu et al. 2019) model.

**Rewards** A mental health and legal counseling dialogue system should be able to employ correct counseling strategy with politeness and empathy in the generated responses. These rewards are *Task-specific* rewards, trying to fulfil the trueness of some tasks. Further, to ensure the language quality and context adequacy, it should generate contextually coherent and non-repetitive responses. These rewards are *Generic* rewards trying to ensure the language quality in responses. Thus, we design an efficient compound reward function $R$ that captures all the five aspects *viz.* $R_1$ adapts true counseling strategy, $R_2$ incorporates politeness, $R_3$ ensures required empathy, $R_4$ accounts for contextual-coherence, and $R_5$ reinforces non-repetitiveness. Lastly, to formulate $R$ appropriately, a weighted sum of all these five rewards is considered, which can be expressed as:

**Counseling, Politeness, and Empathy Reward:** The agent is forced to employ correct counseling strategy, politeness and empathy in the generated responses by penalizing the generated responses deviating from the ground truth counseling strategy, politeness and empathy classes. To achieve this, first, the counseling strategy, politeness and empathy label for the generated response $r_T$ at turn $T$ are predicted. Then, these predictions are compared with their respective labels of the ground truth response. Counseling reward ($R_1$), Politeness Reward ($R_2$), and Empathy Reward ($R_3$) can be formulated as:

$$R_1 = \mathcal{P}_{cou}(c_T) - \beta \sum_{i \in C_{cs}} \mathcal{P}_{cou_i}(r_T) \tag{5}$$

$$R_2 = \mathcal{P}_{pol}(c_T) - \beta \sum_{i \in C_{po}} \mathcal{P}_{pol_i}(r_T) \tag{6}$$

$$R_3 = \mathcal{P}_{emp}(c_T) - \beta \sum_{i \in C_{em}} \mathcal{P}_{emp_i}(r_T) \tag{7}$$

where, $\mathcal{P}_{cou}(c_T)$, $\mathcal{P}_{pol}(c_T)$, and $\mathcal{P}_{emp}(c_T)$ denote the predicted probabilities of counseling strategy, politeness, and empathy for the ground truth utterance $c_T$ at turn $T$. Similarly, $\mathcal{P}_{cou_i}(r_T)$, $\mathcal{P}_{pol_i}(r_T)$, and $\mathcal{P}_{emp_i}(r_T)$ denote the predicted probabilities of counseling strategy, politeness and empathy for the generated response $r_T$ at turn $T$; $i \in C_{cs}$, $C_{po}$, and $C_{em}$ with $C_{cs} = \{c_1, c_2, ..., c_n\}$, $C_{po} = \{c_1, c_2, ..., c_m\}$ and $C_{em} = \{0, 1\}$ representing a set of $n$ counseling strategies, $m$ politeness labels, and two empathy classes, respectively. The $\beta$ serves as a penalization factor, i.e. greater the $\beta$ is, greater would be the penalization[9].

**Contextual-Coherence Reward:** Any dialogue system should maintain the context of the dialogue. Hence, the generated responses should not deviate from the dialogue context. Therefore, to assess the contextual coherence of the generated responses, the fourth reward $R_4$ is formulated as:

$$R_4 = \frac{1}{3}(\cos(r_T, c_T) + \cos(r_T, u_T) + \cos(r_T, c_{T-1})) \tag{8}$$

**Non-Repetitivenesss Reward:** A counseling dialogue system should not ask/suggest the same responses to the victims. It is required to engage the victim in the *subject* s/he is being counseled in. Therefore, the generated responses should be diverse and interactive. To account for this, a non-repetitiveness reward, $R_5$ is calculated as the Jaccard distance between the generated responses, $r_T$ and $r_{T-1}$ at turns $T$ and $T-1$, respectively (Jaccard 1912), which can be formulated as:

$$R_5 = 1 - \frac{r_{T-1} \cap r_T}{r_{T-1} \cup r_T} \tag{9}$$

**Policy** A dialogue agent's action selection can be modeled through a policy which can be defined as a probability mapping function $\mathcal{P}_\theta$ that represents the likelihood of generating an utterance $r$ consisting of $M$ tokens .

$$\mathcal{P}_\theta(r_{1:M}|x) = \prod_{m=0}^{M} \mathcal{P}_\theta(r_m|y_{<m}, x) \tag{10}$$

**Proximal Policy Optimization:** To have low variance from the old policies, policy updates are made at each step using the PPO method. It updates the current policy by seeking improvement on a certain parameters so that it is not too different from the previous policy. The policy optimization can be decomposed into three steps: First, using gradient ascent, the expected reward is maximized in the loss function $J(\theta)$:

$$\nabla_\theta J(\theta) = E_{r \sim \mathcal{P}_\theta}[\nabla_\theta \log \mathcal{P}_\theta(r) \hat{A}_r] \tag{11}$$

Second, to restrict large deviation from the old policy, log term in the above equation is replaced with an importance sampling term. Further, to prevent catastrophic forgetting, clipping is performed. The clipped variant of PPO does not take into account any KL-divergence term (Kullback and Leibler 1951) or any constraint in the objective function, rather it relies on a specialized clipping to ensure small deviation from the true distribution. It can be formulated as:

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

Here, $pr_r(\theta) = \mathcal{P}_\theta^{new}/\mathcal{P}_\theta^{old}$ gives the probability ratio of generating a response between the new and old policies. $\varepsilon$ is the clipping range and $\hat{A}_y$ is the estimated advantage which is the normalized rewards in our case. Lastly, parameters are updated using the following steps:

$$\theta_{k+1} = \underset{\theta}{\text{argmax}} \underset{s,a \sim \mathcal{P}_{\theta_k}}{E} [L^{\text{CLIP}}] \tag{12}$$

[9]The value of $\beta$ is taken as greater than or equal to 1.

| | BERT-large | | RoBERTa-large | |
|---|---|---|---|---|
| **Classifier** | **W-ACC** | **Macro-F1** | **W-ACC** | **Macro-F1** |
| Counseling strategy | 0.904 | 0.851 | 0.923 | 0.869 |
| Politeness | 0.978 | 0.964 | 0.990 | 0.989 |
| Empathy | 0.962 | 0.952 | 0.977 | 0.972 |

Table 2: Evaluation results of the counseling-strategy, politeness and empathy classifiers.

| Model | CoStr | Pol | Emp | PPL | R-LEN |
|---|---|---|---|---|---|
| ARDM (Wu et al. 2021) | 75.24% | 89.1% | 41.3% | 3.21 | 16.02 |
| Po-Em-MHLCDS-R | 77.13% | 90.1% | 42.6% | 2.87 | 16.91 |
| **Po-Em-MHLCDS** | **80.30%** | **92.54%** | **46.4%** | **1.91** | **18.71** |

Table 3: Results of automatic evaluation. Here, Po-Em-MHLCDS refers to our proposed system considering all rewards. Po-Em-MHLCDS-R refers to Po-Em-MHLCDS with no rewards.

## Experiments

### Implementation Details

To train the MLE-loss based dialogue model (MLE-DM), the pre-trained GPT-2 medium (Radford et al. 2019) is employed. The fine-tuning of trained MLE-DM is done in an RL setting by experimenting with different number of candidate responses, i.e., $n = 2, 3, 4, 5, 10$. As per the loss obtained, $n = 3$ is selected as the final value. To decode generated candidates, nucleus sampling (Holtzman et al. 2019) is adopted with temperature $T = 0.8$ and probability $p = 0.9$. The proposed system 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 $0.3, 0.2, 0.2, 0.2, 0.1$ are chosen as the final weights for $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, and $\alpha_5$, respectively. Lastly, for counseling, politeness and empathy rewards, the penalization factor, $\beta$ is set to 2.

For validation along with training, considering three candidate responses per utterance per dialogue, **Po-Em-MHLCDS** takes approximately 30 mins/epoch, hence, total time it took 600 minutes (10 hours) to train and validate the model. Finally, the testing of the proposed system takes approximately 5 minutes for 300 utterances.

### Evaluation Metrics

The proposed system, **Po-Em-MHLCDS** is evaluated using both automatic and human evaluation metrics. Counseling strategy, politeness and empathy classifiers are evaluated in terms of Weighted Accuracy (W-ACC) and Macro-F1 (to account for imbalanced class distribution).

The effectiveness of the proposed **Po-Em-MHLCDS** model is checked by evaluating it in terms of task success (here, counseling, politeness and empathy) and quality of the generated responses. Task success metrics are: **CoStr** that computes the number of utterance generated with counseling strategy, **Pol** which gives the number of polite utterances, and **Emp** that computes the number of empathetic utterances generated. To check the quality of the generated responses, Perplexity (**PPL**) score and response-length (**R-LEN**) are evaluated. **CoStr**, **Pol**, and **Emp** are evaluated using a counseling classifier[10], politeness and empathy classifiers. The accuracy of these classifiers on test set provides the respective metrics (**CoStr**, **Pol**, **Emp**) scores for our proposed system, **Po-Em-MHLCDS**.

Human evaluation are done by recruiting six evaluators with postgraduate qualification and proficiency in similar tasks[11]. To test the robustness of our system, each evaluator is asked to interact with our system 3 times, with a constraint that each time they would have to interact by using a different set of responses. Then, these 18 human-evaluated dialogues are sent to the experts from government-run institutions for cross-verification in terms of evaluation quality. After experts pass the evaluation process, further 42 dialogues are evaluated. Hence, we end up with total 60 human evaluated dialogues. All six evaluators are asked to rate each dialogue interaction in terms of counseling strategy correctness (**Con**), politeness (**Pol**), empathy (**Emp**), consistency (**Const**), fluency (**Fluen**), and non-repetitiveness (**N-Rep**) on an integer scale of 1-5[12].

## Results and Analysis

We performed experimental results analysis in two steps. First, we analyze the results of sub modules, i.e. counseling strategy, politeness and empathy classifiers used in our propose **Po-Em-MHLCDS** framework. Then, we state the results of the proposed system, and compare it with two baseline models, *viz.* ARDM (MLE-based model) (Wu et al. 2021) (one with which we initialized our proposed model) and Po-Em-MHLCDS-R (one with zero reward). Lastly, to study the effect and contribution of each of the task-specific rewards and generic rewards, we perform reward weight optimization. The automatic and human evaluation results are shown in Table 3 and Table 4, respectively.

Evaluation results of all three classifiers are shown in Table 2. It can be observed that all three classifiers achieved

[10]Accuracy of counseling classifier is 94.23%. It checks if a response incorporates a counseling strategy or not.

[11]Human evaluators were paid as per our university norms.

[12]The scale 1-5 denotes low to high intensity such as Con = 1 denotes non-counseling and Con = 5 denotes highly counseling.

| Model | Con | Pol | Emp | Const | Fluen | N-Rep |
|---|---|---|---|---|---|---|
| ARDM | 3.04 | 3.83 | 2.13 | 3.74 | 4.12 | 3.87 |
| Po-Em-MHLCDS-R | 3.39 | 3.96 | 2.28 | 3.91 | 4.31 | 4.11 |
| **Po-Em-MHLCDS** | **3.94** | **4.41** | **2.85** | **4.16** | **4.57** | **4.72** |

Table 4: Results of human evaluation.

significantly good scores in terms of both Weighted Accuracy (**W-ACC**) and **Macro-F1**. It can also be seen that RoBERTa-large (Liu et al. 2019) yields better scores for both the metrics as compared to BERT-large (Devlin et al. 2018). It could be due to the fact that RoBERTa-large consists of larger number of parameters constituting it to better approximate for all the three classifications.

### Automatic evaluation

In Table 3, it can be seen that our proposed model, **Po-Em-MHLCDS** performs better than the baselines: ARDM and without reward model (Po-Em-MHLCDS-R) in terms of all metrics. For task-specific metrics, *viz.* **CoStr**, **Pol**, and **Emp**, **Po-Em-MHLCDS** achieves good scores of 80.3%, 92.54%, and 46.4%, with a significant difference of 5.06, 3.44, and 5.1, respectively as compared to the baseline ARDM[13]. This justifies the design of reward function as it can be seen that **Po-Em-MHLCDS** is able to generate polite and empathetic responses while incorporating the correct counseling strategy. It is due to the fact that task-specific rewards forces the RL-agent to generate counseling strategy grounded responses. It can be also seen that as compared to the Po-Em-MHLCDS-R, our proposed system performs better in terms of **CoStr**, **Pol**, and **Emp** with a difference of 3.17, 2.44, and 3.8, respectively. This further strengthens our design of rewards to build a polite and empathetic mental health and legal counseling dialogue system. It can also be observed from Table 3 that **Po-Em-MHLCDS** obtains well scores of **PPL** - 1.91 and **R-LEN** - 18.71 with a significant difference of 1.3 and 2.69, respectively in comparison to the baseline ARDM. This could be due to the task-specific and contextual-coherence rewards driving the model to build a connection with the victim in smooth language to generate contextually adequate and fluent responses. This results in generation of interactive and engaging responses. Lastly, it can also be seen in Table 3 that **Po-Em-MHLCDS** performs better than its variant Po-Em-MHLCDS-R, hence, strengthening the requirement of both task-specific and generic rewards to generate fluent, non-repetitive, polite and empathetic responses grounded in appropriate counseling strategy.

### Human evaluation

Table 4 shows the human evaluation results. It can be observed that **Po-Em-MHLCDS** yields better scores in terms of **Con**, **Pol**, **Emp**, **Const**, **Fluen** and **N-Rep** with a difference of 0.9, 0.58, 0.72, 0.42, 0.45, and 0.85, respectively as compared to the the baseline ARDM scores of **Const**: 4.16, **Fluen**: 4.57, and **N-Rep**: 4.72, which implies that contextual-coherence and fluency rewards have played a crucial role in generating consistent, fluent and non-repetitive utterances. Further, in terms of **Con**, **Pol** and **Emp**, **Po-Em-MHLCDS** attains well scores of 3.94, 4.41, and 2.85, respectively. Therefore, it can be inferred that adding politeness and empathy rewards with counseling reward helps **Po-Em-MHLCDS** to build a rapport with the victim, by generating engaging and interactive responses.

[13]We perform statistical significance test, Welch's t-test (Welch 1947), and it is conducted at 5% (0.05) significance level.

## Conclusion and Future Work

In this work, we have built a polite and empathetic mental health and legal counseling dialogue system, **Po-Em-MHLCDS** to offer higher engagement in e-counseling sessions and resolve the issues faced by the *victims*. For this, we prepared a Mental Health and Legal Counseling Dataset (**MHLCD**) and further annotated it with three informative labels, *viz.* counseling strategies, politeness and empathy. Then, we trained the proposed system on this annotated dataset in a reinforcement learning framework. A novel reward function ensuring correct counseling strategy, politeness and empathy while reinforcing contextual-coherence and non-repetitiveness in the generated responses is designed to optimize the RL-loss. Automatic and human evaluation conclude that **Po-Em-MHLCDS** achieves promising results as compared to strong MLE-loss based baselines. Further, our results also supports the use of designed reward function to better facilitate the counseling, politeness and empathy in generated response. A counseling dialogue system can use the external knowledge to generate knowledge-grounded and more realistic responses. This gives future directions for our current research work.

## Acknowledgments

Kshtij Mishra gratefully acknowledges the assistance of the Indian government's "Prime Minister's Research Fellowship (PMRF) Program". Priyanshu Priya acknowledges the 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 for financial support. Authors acknowledge the 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.

## References

Althoff, T.; Clark, K.; and Leskovec, J. 2016. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics*, 4: 463–476.

Brown, P.; Levinson, S. C.; and Levinson, S. C. 1987. *Politeness: Some universals in language usage*, volume 4. Cambridge university press.

Casanueva, I.; Budzianowski, P.; Su, P.-H.; Ultes, S.; Rojas-Barahona, L.; Tseng, B.-H.; and Gašić, M. 2018. Feudal reinforcement learning for dialogue management in large domains. *arXiv preprint arXiv:1803.03232*.

Castonguay, L. G.; and Hill, C. E. 2017. *How and why are some therapists better than others?: Understanding therapist effects*. JSTOR.

Coppock, L. 2005. Politeness strategies in conversation closings. *Unpublished manuscript: Stanford University*.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Do, P.-K.; Nguyen, H.-T.; Tran, C.-X.; Nguyen, M.-T.; and Nguyen, M.-L. 2017. Legal question answering using ranking SVM and deep convolutional neural network. *arXiv preprint arXiv:1703.05320*.

Elliott, R.; Bohart, A. C.; Watson, J. C.; and Greenberg, L. S. 2011. Empathy. *Psychotherapy*, 48(1): 43.

Elliott, R.; Bohart, A. C.; Watson, J. C.; and Murphy, D. 2018. Therapist empathy and client outcome: An updated meta-analysis. *Psychotherapy*, 55(4): 399.

Eysenbach, G.; Powell, J.; Englesakis, M.; Rizo, C.; and Stern, A. 2004. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Bmj*, 328(7449): 1166.

Feng, B.; Li, S.; and Li, N. 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.

Fitzpatrick, K. K.; Darcy, A.; and Vierhile, M. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health*, 4(2): e7785.

Fulmer, R.; Joerin, A.; Gentile, B.; Lakerink, L.; Rauws, M.; et al. 2018. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health*, 5(4): e9782.

Gage-Bouchard, E. A.; LaValley, S.; Warunek, M.; Beaupin, L. K.; and Mollica, M. 2018. Is cancer information exchanged on social media scientifically accurate? *Journal of Cancer Education*, 33(6): 1328–1332.

Gibson, J.; Can, D.; Xiao, B.; Imel, Z. E.; Atkins, D. C.; Georgiou, P.; and Narayanan, S. 2016. A deep learning approach to modeling empathy in addiction counseling. *Commitment*, 111(2016): 21.

Golchha, H.; Firdaus, M.; Ekbal, A.; and Bhattacharyya, P. 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, Volume 1 (Long and Short Papers)*, 851–860.

Holmes, E. A.; Ghaderi, A.; Harmer, C. J.; Ramchandani, P. G.; Cuijpers, P.; Morrison, A. P.; Roiser, J. P.; Bockting, C. L.; O'Connor, R. C.; Shafran, R.; et al. 2018. The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry*, 5(3): 237–286.

Holtzman, A.; Buys, J.; Du, L.; Forbes, M.; and Choi, Y. 2019. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*.

Jaccard, P. 1912. The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2): 37–50.

John, A. K.; Caro, L. D.; Robaldo, L.; and Boella, G. 2017. Legalbot: A deep learning-based conversational agent in the legal domain. In *International Conference on Applications of Natural Language to Information Systems*, 267–273. Springer.

Kelley, J. F. 1984. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems (TOIS)*, 2(1): 26–41.

Kim, J.; Kim, Y.; Kim, B.; Yun, S.; Kim, M.; and Lee, J. 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*, 1–6.

Kullback, S.; and Leibler, R. A. 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1): 79–86.

Lee, F.-T.; Hull, D.; Levine, J.; Ray, B.; and McKeown, K. 2019. Identifying therapist conversational actions across diverse psychotherapeutic approaches. In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, 12–23.

Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; and Stoyanov, V. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Loshchilov, I.; and Hutter, F. 2018. Decoupled Weight Decay Regularization. In *International Conference on Learning Representations*.

Lucas, G. M.; Gratch, J.; King, A.; and Morency, L.-P. 2014. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37: 94–100.

McHugh, M. L. 2012. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3): 276–282.

Mesgar, M.; Simpson, E.; and Gurevych, I. 2020. Improving factual consistency between a response and persona facts. *arXiv preprint arXiv:2005.00036*.

Mishra, K.; Firdaus, M.; and Ekbal, A. 2022. Please be polite: Towards building a politeness adaptive dialogue system for goal-oriented conversations. *Neurocomputing*, 494: 242–254.

Morris, R. R.; Kouddous, K.; Kshirsagar, R.; and Schueller, S. M. 2018. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. *Journal of medical Internet research*, 20(6): e10148.

Norcross, J. C. 2002. *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients.* Oxford University Press.

Pérez-Rosas, V.; Mihalcea, R.; Resnicow, K.; Singh, S.; and An, L. 2017. Understanding and predicting empathic behavior in counseling therapy. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1426–1435.

Pérez-Rosas, V.; Wu, X.; Resnicow, K.; and Mihalcea, R. 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*, 926–935.

Pruksachatkun, Y.; Pendse, S. R.; and Sharma, A. 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*, 1–13.

Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8): 9.

Reis, J. C.; Correia, A.; Murai, F.; Veloso, A.; and Benevenuto, F. 2019. Supervised learning for fake news detection. *IEEE Intelligent Systems*, 34(2): 76–81.

Saha, T.; Gakhreja, V.; Das, A. S.; Chakraborty, S.; and Saha, S. 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*, 2650–2656.

Saha, T.; Reddy, S.; Das, A.; Saha, S.; and Bhattacharyya, P. 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*, 2436–2449.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Sharma, A.; Choudhury, M.; Althoff, T.; and Sharma, A. 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, 614–625.

Sharma, A.; Lin, I. W.; Miner, A. S.; Atkins, D. C.; and Althoff, T. 2021. Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference 2021*, 194–205.

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

Wadden, D.; August, T.; Li, Q.; and Althoff, T. 2021. The effect of moderation on online mental health conversations. *ICWSM*.

Welch, B. L. 1947. The generalization of 'STUDENT'S'problem when several different population variances are involved. *Biometrika*, 34(1-2): 28–35.

White, M.; and Dorman, S. M. 2001. Receiving social support online: implications for health education. *Health education research*, 16(6): 693–707.

Wu, Q.; Zhang, Y.; Li, Y.; and Yu, Z. 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*, 1292–1301.

Wu, X.; Martinez, A.; and Klyen, M. 2018. Dialog generation using multi-turn reasoning neural networks. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 2049–2059.

Yazdavar, A. H.; Mahdavinejad, M. S.; Bajaj, G.; Thirunarayan, K.; Pathak, J.; and Sheth, A. 2018. Mental health analysis via social media data. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, 459–460. IEEE.

Zhang, C.; D'Haro, L. F.; Friedrichs, T.; and Li, H. 2022. MDD-Eval: Self-Training on Augmented Data for Multi-Domain Dialogue Evaluation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 11657–11666.

Zhang, J.; and Danescu-Niculescu-Mizil, C. 2020. Balancing objectives in counseling conversations: Advancing forwards or looking backwards. *arXiv preprint arXiv:2005.04245*.

Zhang, J.; Filbin, R.; Morrison, C.; Weiser, J.; and Danescu-Niculescu-Mizil, C. 2019. Finding your voice: The linguistic development of mental health counselors. *arXiv preprint arXiv:1906.07194*.

Zhong, M.; Liu, Y.; Xu, Y.; Zhu, C.; and Zeng, M. 2022. Dialoglm: Pre-trained model for long dialogue understanding and summarization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 11765–11773.