

# Generating Personalized Dialogue via Multi-Task Meta-Learning

Lee Jing Yang<sup>1</sup>, Lee Kong Aik<sup>2</sup>, Gan Woon Seng<sup>3</sup>

School of Electrical and Electronic Engineering, Nanyang Technological University<sup>1,3</sup>

Institute for Infocomm Research, A\*STAR<sup>2</sup>

jingyang001@e.ntu.edu.sg<sup>1</sup>, lee\_kong\_aik@i2r.a-star.edu.sg<sup>2</sup>, ewsgan@ntu.edu.sg<sup>3</sup>

## Abstract

Conventional approaches to personalized dialogue generation typically require a large corpus, as well as predefined persona information. However, in a real-world setting, neither a large corpus of training data nor persona information are readily available. To address these practical limitations, we propose a novel multi-task meta-learning approach which involves training a model to adapt to new personas without relying on a large corpus, or on any predefined persona information. Instead, the model is tasked with generating personalized responses based on only the dialogue context. Unlike prior work, our approach leverages on the provided persona information only during training via the introduction of an auxiliary persona reconstruction task. In this paper, we introduce 2 frameworks that adopt the proposed multi-task meta-learning approach: the Multi-Task Meta-Learning (MTML) framework, and the Alternating Multi-Task Meta-Learning (AMTML) framework. Experimental results show that utilizing MTML and AMTML results in dialogue responses with greater persona consistency.

## 1 Introduction

Personalized dialogue generation involves generating dialogue responses which incorporates the personality of the interlocutors, leading to more natural and human-like dialogue. Thus far, approaches to personalized dialogue generation typically require a large number of persona-specific dialogue examples. Certain approaches also require persona information presented in the form of several predefined persona statements (eg. 'I love dogs', 'I am an engineering student.'). However, in a real-world system, large amounts of persona-specific dialogue are rarely available, and collecting descriptive persona statements from every interlocutor is intractable.

To address these practical issues, Persona Agnostic Meta-Learning (PAML) (Madotto et al., 2019), a framework which aims to train a model capable of rapid adaptation to new unseen personas, was proposed. The PAML framework was based on the popular Model-Agnostic Meta-Learning (MAML) framework (Finn et al., 2017). The recently proposed Customized Model Agnostic Meta-Learning (CMAML) (Song et al., 2020b) framework largely follows the PAML framework, with the exception of an additional network structure optimization component. Both the PAML and CMAML frameworks were benchmarked on the PersonaChat corpus (Zhang et al., 2018), a popular personalized dialogue generation corpus which provides persona statements describing each interlocutor in addition to the persona-specific dialogues. As it is unfeasible to collect persona statements from interlocutors in a real world setting, the PAML framework does not utilize the available persona statements during both meta-learning and inference. However, even though it is impractical to utilize the persona statements during inference, the persona statements can be used *during meta-learning* to further improve model performance.

Hence, we introduce a novel multi-task meta-learning approach which leverages predefined persona statements *only during meta-learning* via an additional persona reconstruction task. Essentially, this task involves generating all corresponding persona statements in its entirety given the dialogue context. We hypothesize that the introduction of the persona reconstruction task would result in parameters capable of effectively inducing the persona information from the dialogue context, which would lead to the generation of persona consistent dialogue. The persona statements are *not* used during inference. Prior usage of multi-task learning for personalized dialogue generation involved the addition of a persona classification task (Yang et al.,

2021; Su et al., 2019a) and a distraction utterance binary classification task (Na et al., 2021). To our knowledge, this is the first attempt at incorporating persona statement reconstruction.

Our contributions include 2 multi-task meta-learning frameworks which leverage the persona reconstruction task only during training. The Multi-Task Meta-Learning (MTML) framework, as well as a variant known as the Alternating Multi-Task Meta-Learning (AMTML) framework. While both MTML and AMTML involve the addition of a persona reconstruction task only during meta-learning, MTML involves combining the losses derived from generating the response and reconstructing the persona. AMTML, on the other hand, functions by constantly alternating between both tasks. Experimental results on the PersonaChat corpus reveal that utilizing MTML and AMTML result in responses which reflect the interlocutor’s persona to a larger extent compared to prior work.

## 2 Methodology

Our approach to personalized dialogue generation involves extending the PAML framework (Madotto et al., 2019) by introducing a persona reconstruction task only during the meta-learning stage. The PAML framework is essentially an adaptation of the general MAML framework for personalized dialogue generation. The MAML framework involves learning a parameter initialization capable of generalizing and adapting rapidly to new tasks unseen during the training process via gradient descent. Specifically, this involves obtaining the updated parameters by adjusting the model parameters by utilizing data from several related tasks. The original parameters are then optimized based on the updated parameters and a separate test set by computing the second order derivatives (Hessian matrix). In the PAML framework, each persona is viewed as a unique task. Consequently, the goal of the PAML framework is to obtain a parameter initialization capable of rapidly adapting to new personas. We hypothesize that the introduction of the persona reconstruction task would result in parameters which induce the persona from the dialogue context to a larger extent.

### 2.1 Multi-task Learning

Unlike traditional approaches that involve utilizing both the dialogue context and persona information as model inputs during training, our framework

involves reconstructing the persona statements as well as generating the dialogue response given the dialogue context. The primary task of generating the response and the corresponding loss function  $L_{res}$  can be expressed in the following equations:

$$f_{\phi}(x_t|x_{1:t-1}) = p(x_t|x_{1:t-1}; \phi) \quad (1)$$

$$L_{res}(\phi) = - \sum_{t=1}^T \log p(x_t|x_{1:t-1}; \phi) \quad (2)$$

where  $x_t$  and  $x_{1:t-1}$  refer to the response and dialogue context respectively, and  $\phi$  represents the model parameters. We hypothesize that the persona reconstruction task would result in model parameters capable of inducing the persona information from the dialogue context to a larger extent. This is due to the increased emphasis on persona consistency in the task of persona reconstruction. Also, rather than generating only selected keywords or phrases, the persona reconstruction task involves generating all corresponding persona statements in its entirety given the dialogue context. This is because generating complete sentences would also require the model to account for fluency in addition to persona consistency, which is also a vital aspect of dialogue generation. The persona statements  $\mathcal{P}_{1:N}$  are concatenated to form a single sequence  $\bar{\mathcal{P}}$ . The auxiliary persona reconstruction task and the corresponding loss function  $L_{rec}$  is formalized in the following expression:

$$\bar{\mathcal{P}} = \text{concat}(\mathcal{P}_{1:N}) \quad (3)$$

$$f_{\phi}(\bar{\mathcal{P}}|x_{1:t-1}) = p(\bar{\mathcal{P}}|x_{1:t-1}; \phi) \quad (4)$$

$$L_{rec}(\phi) = - \sum_{t=1}^T \log p(\bar{\mathcal{P}}|x_{1:t-1}; \phi) \quad (5)$$

During training, the persona reconstruction loss and the response generation loss will be weighted and summed. Hence, the total multi-task loss is expressed as:

$$L(\phi) = \alpha L_{res}(\phi) + (1 - \alpha) L_{rec}(\phi) \quad (6)$$

where  $\alpha$  determines the contribution of the persona reconstruction and response generation loss to the learning process.

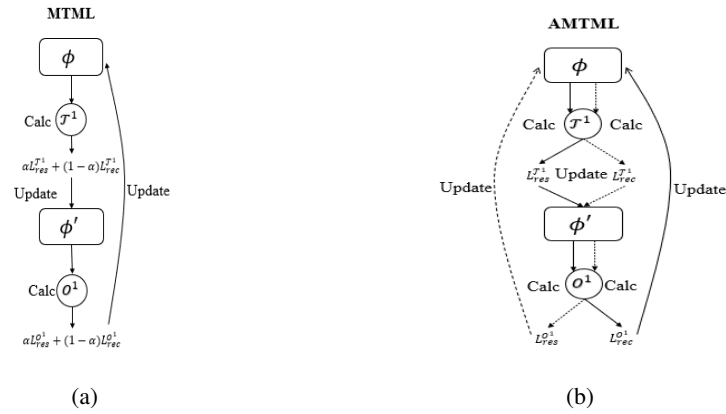


Figure 1: Diagrams depicting the computation path for MTML and AMTML. A batch size of 1 is assumed. (a): Diagram depicting the MTML framework. The weighted sum of the persona reconstruction loss and the response generation loss is used to update the parameters  $\phi$ . (b): Diagram depicting the AMTML framework. The computation flow alternates between the solid arrows and the dotted arrows after every iteration.

<b>Persona</b>	
i have three guns and love hunting my family lives down the street from me i go to church every sunday i drive a ford pickup truck i am very conservative	
<b>Dialogue Context</b>	
User1:	hey there how are you i am shy
User2:	well i am conservative so sounds like a match to me
User1:	i like that , did you go to college ?
User2:	no i just got my high school diploma and opened my owns shooting range
User1:	i went an got a computer science degree
User2:	wish i had continued school but i love what i do hunting , going to church and driving
User1:	gotcha , i build spaceships , the models
User2:	o ok my family lives down from me and they build garages and farms
User1:	that is cool , my mom is a doctor
<b>Responses</b>	
Ref	you ever eat at wendys ? i went to washington school before i quit .
<b>10-shot</b>	
<i>Std</i>	that is cool where are you from ?
<i>Std<sub>p</sub></i>	that is cool what do you do for work ?
<i>PAML</i>	he is good to church every week
<i>MTML<sub>0.8</sub></i>	sure , i pick him up for church every sunday with my ford pickup
<i>AMTML</i>	my parents live in the house i grew up in , just down the street .
<b>5-shot</b>	
<i>Std</i>	that is cool what kind of doctor do you do ?
<i>Std<sub>p</sub></i>	that is cool what kind of garages do you drive ?
<i>PAML</i>	that is good where do you live ?
<i>MTML<sub>0.8</sub></i>	that is good what do you do for work ?
<i>AMTML</i>	oh okay lol what do you do for work ?

Table 1: Dialogue responses generated by the implemented models.

## 2.2 Multi-Task Meta Learning Framework

Similar to PAML, MTML aims to learn a set of parameters  $\phi$  capable of quickly adapting to an unseen persona in order to generate contextually appropriate responses which reflect the persona, without relying on user persona statements. However, unlike PAML, which does not utilize the persona statements during both meta-learning and inference, we leverage the persona information *only during meta-learning*. The persona statements are not used during inference. The persona information is incorporated during meta-learning via the multi-task learning (Section 3.1).

We begin by dividing the corpus into a training set, validation set and a test set denoted by  $\mathcal{D}_{train}$ ,  $\mathcal{D}_{valid}$  and  $\mathcal{D}_{test}$  respectively. For each iteration  $i$ , a batch distinct of personas  $\mathcal{P}_{1:N}^{1:M}$  and a corresponding set of  $M$  dialogues are randomly sampled from the training set  $\mathcal{D}_{train}$  to form the meta training set  $\mathcal{T}^i$  (support set). Then, another set of  $M$  dialogues corresponding to the same batch of personas  $\mathcal{P}_{1:N}^{1:M}$  are randomly sampled from the training set  $\mathcal{D}_{train}$  to form the meta optimization set  $\mathcal{O}^i$  (query set). We refer to the meta training set and meta optimization set collectively as  $\mathcal{C}^i$  i.e.,  $\mathcal{C}^i = (\mathcal{T}^i, \mathcal{O}^i)$ .

During meta-training, the multi-task loss  $L_{mul}^{\mathcal{T}^i}$  is computed via a weighted sum between the response generation loss  $L_{res}^{\mathcal{T}^i}$  and the persona recreation loss  $L_{rec}^{\mathcal{T}^i}$ . To calculate  $L_{rec}^{\mathcal{T}^i}$ , the persona statements  $\mathcal{P}_{1:N}^{1:M}$  of each persona were concatenated into a single sequence, resulting in  $\bar{\mathcal{P}}^{1:M}$ . For an arbitrary persona in the meta training set, the equations for computing the multi-task loss are as follows:

$$L_{res}^{\mathcal{T}^i}(\phi) = - \sum \log p(x_t, |x_{1:t-1}; \phi) \quad (7)$$

$$\bar{\mathcal{P}}^{1:M} = \text{concat}(\mathcal{P}_{1:N}^{1:M}) \quad (8)$$

$$L_{rec}^{\mathcal{T}^i}(\phi) = - \sum \log p(\bar{\mathcal{P}} | x_{1:t-1}; \phi) \quad (9)$$

$$L^{\mathcal{T}^i}(\phi) = \alpha L_{res}^{\mathcal{T}^i}(\phi) + (1 - \alpha) L_{rec}^{\mathcal{T}^i}(\phi) \quad (10)$$

where  $\alpha$  accounts for the distribution between the 2 losses. Subsequently, the parameters  $\phi$  are updated via SGD. The updated model parameters  $\phi'$  can be expressed as:

$$\phi' = \phi - \eta_t \nabla_{\phi} L^{\mathcal{T}^i}(\phi) \quad (11)$$

where  $\eta_t$  refers to the inner loop learning rate and  $L^{\mathcal{T}^i}(\phi)$  represents the multi-task training loss attained.

During meta-optimization, the original model  $\phi$  is optimized on the multi-task loss attained on the meta optimization set  $\mathcal{O}^i$  and the updated parameters  $\phi'$ . This meta-objective function can be expressed as:

$$\begin{aligned} \min_{\phi} & \sum_{\mathcal{C}^i \sim \mathcal{D}_{train}} L^{\mathcal{O}^i}(\phi') \\ & = \sum_{\mathcal{C}^i \sim \mathcal{D}_{train}} L^{\mathcal{O}^i}(\phi - \eta_t \nabla_{\phi} L^{\mathcal{T}^i}(\phi)) \end{aligned} \quad (12)$$

where  $L^{\mathcal{O}^i}$  refers to the multi-task loss attained on the sampled meta optimization set  $\mathcal{O}^i$ . To obtain  $L^{\mathcal{O}^i}(\phi')$ , we first compute  $L_{res}^{\mathcal{O}^i}(\phi')$  and  $L_{rec}^{\mathcal{O}^i}(\phi')$  by applying Equation 7 - 9 on  $\mathcal{O}^i$ . Then, we compute the weighted sum between both losses:

$$L^{\mathcal{O}^i}(\phi') = \alpha L_{res}^{\mathcal{O}^i}(\phi') + (1 - \alpha) L_{rec}^{\mathcal{O}^i}(\phi') \quad (13)$$

where  $\alpha$  determines the contribution of the persona reconstruction and response generation loss respectively. When  $\alpha = 1$ , MTML is analogous to MAML/PAML. Next, we sum the  $L^{\mathcal{O}^i}(\phi')$  attained for every sampled persona  $\mathcal{P}_{1:N}^{1:M}$ . The original parameters  $\phi$  are then updated using the average of losses obtained by dividing the summed loss by the batch size  $M$ . This can be formalized as:

$$\phi = \phi - \eta_o \nabla_{\phi} \frac{1}{M} \sum_{\mathcal{O}^i} L^{\mathcal{O}^i}(\phi') \quad (14)$$

where  $\eta_o$  refers to the outer loop learning rate and  $L^{\mathcal{O}^i}(\phi')$  represents the multi-task training loss attained by the updated parameters  $\phi'$  on the sampled validation set  $\mathcal{O}^i$ . This computation involves obtaining the gradient of a gradient i.e., second order differentiation. A summary of MTML framework is provided in Algorithm 1. Additionally, an overview of the MTML framework is provided in Figure 1(a).

## 2.3 Alternating Multi-Task Meta-Learning Framework

Instead of combining the response generation loss and persona reconstruction loss, Alternating-MTML (AMTML) involves constantly alternating between the two loss functions. Essentially, at every iteration, meta-training and meta-optimization are conducted with different loss functions. For AMTML, the  $\alpha$  parameter would not be used. At

---

**Algorithm 1** MTML

---

**Require:** Hyperparameters  $\alpha, \eta_t, \eta_o$ **Require:** Dataset  $\mathcal{D}_{train}$ **for** iteration  $i = 1$  to  $n$  **do**  Sample  $\mathcal{C}^i = (\mathcal{T}^i, \mathcal{O}^i) \sim \mathcal{D}_{train}$   **for** each persona in  $\mathcal{T}^i$  **do**    Calc  $L^{\mathcal{T}^i}(\phi) = \alpha L_{res}^{\mathcal{T}^i}(\phi) + (1 - \alpha) L_{rec}^{\mathcal{T}^i}(\phi)$     Update  $\phi' = \phi - \eta_t \nabla_{\phi} L^{\mathcal{T}^i}(\phi)$     Calc  $L^{\mathcal{O}^i}(\phi') = \alpha L_{res}^{\mathcal{O}^i}(\phi') + (1 - \alpha) L_{rec}^{\mathcal{O}^i}(\phi')$   **end for**  Update  $\phi = \phi - \eta_o \nabla_{\phi} \frac{1}{M} \sum_{\mathcal{O}^i} L^{\mathcal{O}^i}(\phi')$ **end for**

---

every iteration, during meta-training, either the response generation loss  $L_{res}$  or the persona reconstruction loss  $L_{rec}$  will be used to update the parameters in the inner loop. Then, the alternate loss would be used to compute the loss for meta-optimization in the outer loop. For example, if the response generation loss is used to update the parameters during meta-training i.e.  $L_{res}^{\mathcal{T}^i}$ , the persona reconstruction loss would be used to derive the meta-optimization loss i.e.  $L_{rec}^{\mathcal{O}^i}$ .

In our implementation, this is achieved by utilizing the response generation loss during meta-training when the iteration count is even, and utilizing the persona reconstruction loss when during meta-optimization when the iteration count is odd. Due to the alternating loss functions, the computational complexity and memory requirements for AMTML is lower than MTML, which requires computing both loss functions during both meta-training and meta-optimization. A summary of the AMTML framework is provided in Algorithm 2. Additionally, an overview of the AMTML framework is provided in Figure 1(b). The training set  $\mathcal{D}_{train}$  is used to train the model and the validation set  $\mathcal{D}_{valid}$  is used to facilitate early stopping.

### 3 Experiment and Results

#### 3.1 Corpus

The proposed MTML and AMTML frameworks were evaluated on the PersonaChat dialogue corpus (Zhang et al., 2018). The corpus comprises 1155 distinct personas, each consisting of several persona statements. The PersonaChat dialogue corpus was chosen due to the available persona statements which are used to compute the persona reconstruction loss. In our experiment, the corpus is divided into a training, validation and test set. The val-

---

**Algorithm 2** Alternating MTML

---

**Require:** Hyperparameters  $\eta_t, \eta_o$ **Require:** Dataset  $\mathcal{D}_{train}$ **for** iteration  $i = 1$  to  $n$  **do**  Sample  $\mathcal{C}^i = (\mathcal{T}^i, \mathcal{O}^i) \sim \mathcal{D}_{train}$   **for** each persona in  $\mathcal{T}^i$  **do**    **if**  $i$  is even **then**      Calc  $L_{res}^{\mathcal{T}^i}(\phi)$       Update  $\phi' = \phi - \eta_t \nabla_{\phi} L_{res}^{\mathcal{T}^i}(\phi)$       Calc  $L^{\mathcal{O}^i}(\phi') = L_{rec}^{\mathcal{O}^i}(\phi')$     **else if**  $i$  is odd **then**      Calc  $L_{rec}^{\mathcal{O}^i}(\phi')$       Update  $\phi' = \phi - \eta_t \nabla_{\phi} L_{rec}^{\mathcal{O}^i}(\phi')$       Calc  $L^{\mathcal{T}^i}(\phi') = L_{res}^{\mathcal{T}^i}(\phi')$     **end if**  **end for**  Update  $\phi = \phi - \eta_o \nabla_{\phi} \frac{1}{M} \sum_{\mathcal{O}^i} L^{\mathcal{O}^i}(\phi')$ **end for**

---

idation and test sets each consist of 100 unique personas. For our experiments, we utilize the training and validation sets during the meta-learning stage and the test set during the testing stage.

#### 3.2 Implementation

Following Madotto et al., we adopt the standard Transformer architecture (Vaswani et al., 2017) consisting of 6 encoder layers, 6 decoder layers and 4 attention heads is used along with the GloVe embedding (Pennington et al., 2014). The dimensions of the word embedding and hidden dimension of the Transformer are fixed at 300. We use SGD ( $\eta_t = 0.005$ ,  $M = 16$ ) during meta-training and Adam ( $\eta_o = 0.003$ ,  $M = 16$ ) during meta-optimization. For MTML, we define an additional hyperparameter,  $\alpha$ , which accounts for the distribution between the persona recreation loss and the response generation loss.

#### 3.3 Evaluation

**Automatic Metrics** Similar to Madotto et al., we compute the BLEU score (Papineni et al., 2002), the perplexity (PPL) and the C score (Madotto et al., 2019) to evaluate the quality of the generated response. The BLEU score measures the similarity between the generated response and the reference response. PPL is the negative log of the generated response. The C-score reflects the amount of persona information present in the generated response by measuring the persona consistency with respect to the corresponding persona statements via

a BERT-based Natural Language Inference (NLI) model, which is finetuned to indicate if the generated response entails or contradicts the corresponding persona statements. A high C score would imply greater persona consistency.

**Human Evaluation** We engaged 3 graduated individuals to evaluate 50 responses for each model using 3 criteria: persona consistency, fluency and contextual coherence. Consistency reflects the amount of persona information corresponding to the persona statements are reflected in the generated response. Fluency accounts for any grammatical, spelling and phrasing issues, while Coherence reflects the appropriateness of the dialogue with respect to the dialogue history. Responses were assigned a rating from -1 to 1 for each criteria. For consistency, -1 indicates contradiction, 0 indicates neutrality and 1 indicates persona consistency (i.e. the response accurately reflects information from the corresponding persona statements). For coherence, the individuals were told to assign a score of -1 for incoherent, contextually illogical responses, 0 for a moderate coherent responses and 1 for coherent, contextual responses. Finally, for fluency, -1 is assigned to responses with several fluency issues, 0 for responses with one or two fluency errors and 1 for perfectly fluent responses.

### 3.4 Experimental Settings

We benchmark MTML/AMTML with the PersonaChat corpus. We train the following Transformer models:

*Std*: We pretrain a standard Transformer model with only the dialogue context as input.

*Std<sub>p</sub>*: We pretrain a standard Transformer model with both the dialogue context and persona statements as inputs.

*PAML*: We pretrain a standard Transformer via PAML (Madotto et al., 2019).

*MTML<sub>α</sub>*: We pretrain a standard Transformer trained via MTML (Section 3.2) where  $\alpha = [0.9, 0.8, 0.7, 0.6, 0.5]$ .

*AMTML*: We pretrain a standard Transformer via AMTML (Section 3.3).

During testing, the pretrained models described above were further trained, or finetuned, on dialogues corresponding to personas from the test set  $\mathcal{D}_{test}$ . As highlighted earlier, the models were finetuned using only the dialogue context, which was constructed by concatenating all previous utterances in the dialogue. For our experiment, the

length of the dialogue context in each dialogue example would vary according to the number of turns. No restriction was placed on the number of dialogue turns required. No persona statements were used during finetuning. In the 5-shot and 10-shot setting, 5 and 10 dialogues corresponding to a persona was used to finetune the model parameters respectively. Then, the model tasked with generating the response corresponding to the same persona. Samples of the dialogue responses generated by each model is provided in Table 1. Table 2 and 3 depicts the results of automatic evaluation in a 10-shot and 5-shot setting respectively, while Table 4 and 5 depicts the results of the human evaluation in a 10-shot and 5-shot setting respectively.

### 3.5 Results & Discussion

*MTML<sub>α</sub>* generally achieves higher C-scores and Consistency scores compared to PAML, indicating responses which incorporate a larger amount of persona information. This confirms our hypothesis that the introduction of the persona reconstruction task during meta-learning would result in a model which induces the persona from the dialogue context to a larger extent. However, it can be seen that PPL increases as  $\alpha$  decreases. Since PPL scores have been found to correlate negatively with human likeness (Adiwardana et al., 2020), a high PPL score is undesirable. This finding is supported by the human evaluation results, where the Fluency and Coherence scores drop as  $\alpha$  increases in both the 5-shot and 10-shot settings. This implies that there is a trade-off between general fluency and persona consistency in the generated response. During meta-learning, if  $\alpha$  is too large, the combined loss can be effectively reduced by minimizing the persona reconstruction loss. Hence, the model would be trained to generate responses which contain as much persona information as possible without considering fluency or context. While this would result in persona consistent responses, the responses would be largely incoherent and unnatural. Since  $\alpha = 0.8$  strikes a balance between the PPL and C-scores, we conclude that the optimal value of  $\alpha = 0.8$ .

Both *MTML<sub>0.8</sub>* and *AMTML* improved the persona consistency of the generated responses. In terms of C-score, *MTML<sub>0.8</sub>* demonstrated a 78.9%(10-shot) and 100%(5-shot) improvement over *PAML*. In terms of persona consistency, *MTML<sub>0.8</sub>* demonstrated a 64.0%(10-shot) and

	PPL	BLEU	C-score
<i>Std</i>	35.87	0.93	0.00
<i>Std<sub>p</sub></i>	38.72	1.66	0.10
<i>PAML</i>	41.80	0.71	0.19
<i>MTML</i> <sub>0.5</sub>	77.32	0.53	0.46
<i>MTML</i> <sub>0.6</sub>	57.10	0.53	0.41
<i>MTML</i> <sub>0.7</sub>	52.44	0.57	0.47
<i>MTML</i> <sub>0.8</sub>	43.28	0.42	0.34
<i>MTML</i> <sub>0.9</sub>	40.39	0.71	0.21
<i>AMTML</i>	48.66	0.48	0.29

Table 2: Automatic evaluation results (10-shot).

Method	PPL	BLEU	C-score
<i>Std</i>	36.75	1.02	-0.02
<i>Std<sub>p</sub></i>	38.78	1.79	0.09
<i>PAML</i>	40.46	0.65	0.15
<i>MTML</i> <sub>0.5</sub>	76.38	0.41	0.50
<i>MTML</i> <sub>0.6</sub>	55.19	0.53	0.48
<i>MTML</i> <sub>0.7</sub>	50.69	0.44	0.45
<i>MTML</i> <sub>0.8</sub>	41.42	0.38	0.30
<i>MTML</i> <sub>0.9</sub>	39.94	0.62	0.13
<i>AMTML</i>	44.90	0.42	0.26

Table 3: Automatic evaluation results (5-shot).

	Consistency	Fluency	Coherence
<i>Std</i>	0.16	0.92	0.24
<i>Std<sub>p</sub></i>	0.19	0.89	0.29
<i>PAML</i>	0.25	0.77	0.28
<i>MTML</i> <sub>0.5</sub>	0.47	0.24	0.20
<i>MTML</i> <sub>0.6</sub>	0.45	0.57	0.27
<i>MTML</i> <sub>0.7</sub>	0.45	0.60	0.30
<i>MTML</i> <sub>0.8</sub>	0.41	0.80	0.32
<i>MTML</i> <sub>0.9</sub>	0.39	0.88	0.17
<i>AMTML</i>	0.42	0.89	0.33

Table 4: Human evaluation results (10-shot).

	Consistency	Fluency	Coherence
<i>Std</i>	0.10	0.87	0.20
<i>Std<sub>p</sub></i>	0.09	0.89	0.21
<i>PAML</i>	0.13	0.84	0.27
<i>MTML</i> <sub>0.5</sub>	0.22	0.03	-0.12
<i>MTML</i> <sub>0.6</sub>	0.24	0.65	0.15
<i>MTML</i> <sub>0.7</sub>	0.29	0.65	0.13
<i>MTML</i> <sub>0.8</sub>	0.22	0.78	0.29
<i>MTML</i> <sub>0.9</sub>	0.15	0.77	0.19
<i>AMTML</i>	0.23	0.85	0.20

Table 5: Human evaluation results (5-shot).

Automatic	PPL	BLEU	C-score
<i>P<sup>2</sup>Bot</i>	18.1	0.61	0.33
Human	Consistency	Fluency	Coherence
<i>P<sup>2</sup>Bot</i>	0.39	0.91	0.43

Table 6: Automatic and human evaluation results attained by *P<sup>2</sup>Bot*

15.4%(5-shot) improvement over *PAML*. Similarly, compared to *PAML*, *AMTML* achieved a 52.6%(10-shot) and 73.3%(5-shot) improvement when it comes to C-score. In terms of Consistency, *AMTML* demonstrated a 68.0%(10-shot) and 76.9%(5-shot) improvement over *PAML*. Compared to *MTML*<sub>0.8</sub>, in both the 10-shot and 5-shot settings, *AMTML* achieved similar Consistency scores and slightly lower C-scores. However, while *MTML*<sub>0.8</sub> is comparable to *PAML* with regard to Fluency and PPL, *AMTML* outperformed *PAML*, *MTML*<sub>0.8</sub> and all other MTML variants in terms of Fluency. When it comes to coherence, *PAML*, *MTML*<sub>0.8</sub> and *AMTML* generally achieved similar results.

Based on the results attained, while responses generated via MTML has a slight edge in terms of persona consistency, responses generated via *AMTML* are more fluent. On a side note, it should also be highlighted that the BLEU score did not correlate with any aspect of human evaluation. This further emphasizes the unsuitability of the BLEU score as an evaluation metric for dialogue generation (Liu et al., 2016).

### 3.5.1 PersonaChat SOTA Comparison

Additionally, we compare our proposed frameworks with the current state-of-the-art framework for PersonaChat: *P<sup>2</sup>Bot* (Song et al., 2020a). *P<sup>2</sup>Bot* involves finetuning the GPT pretrained language model on the training set via a transmitter receiver framework. This framework models the user’s perception of the other party’s persona in addition to the user’s own persona. Hence, unlike MTML and *AMTML*, in the case of *P<sup>2</sup>Bot*, persona statements are provided to the model along with the dialogue context during inference and testing.

From Table 6, it can be observed that the C-score attained by *MTML*<sub>0.8</sub>, in the 10-shot setting, was comparable to *P<sup>2</sup>Bot*. When it comes to the Consistency score, in the 10-shot setting, both *MTML*<sub>0.8</sub> and *AMTML* outperformed *P<sup>2</sup>Bot*. This implies that the responses generated

by *MTML*<sub>0.8</sub> and *AMTML* generally reflect the corresponding persona information to a greater extent compared to *P<sup>2</sup>Bot* despite not being provided the persona statements during testing. However, in terms of Fluency and Coherence, *P<sup>2</sup>Bot* still outperforms both *MTML*<sub>0.8</sub> and *AMTML*. This could be partially attributed to the use of the GPT pretrained model, which enhanced the overall quality of the generated responses.

### 3.5.2 Persona Reconstruction

In this section, we will provide a brief discussion regarding the persona reconstruction task. Based on the results attained, it is clear that the introduction of the persona reconstruction task during meta-learning further incentivizes the model to incorporate more persona information in the generated responses. Under the proposed frameworks, it is challenging to evaluate the performance of the model solely on the persona reconstruction task. However, based on the observed responses and loss values, persona reconstruction tend to be more successful when a longer dialogue context  $x_{1:t-1}$  (greater number of turns) is provided. This is expected as a short dialogue context would not contain sufficient persona information for the model to reconstruct.

Persona reconstruction is a interesting and challenging task that could be explored in future work. For this task, the model has to successfully infer the persona from the dialogue context as well as ensure the fluency of the generated description. Finetuning pretrained language models would be a good starting point for future work. Also, to prevent a mix-up between the personas from each interlocutor, only the dialogue utterances of the corresponding interlocutor should be utilized during training and inference.

## 4 Related Work

**Multi-task Learning** Multi-task learning broadly refers to the process of learning more than one tasks/objectives concurrently with a shared model. In addition to personalized dialogue generation, multi-task learning has been applied to task-oriented dialogue subtasks including response generation (Zhu et al., 2019), dialogue state tracking (GM and Sengupta, 2019; Trinh et al., 2018; Rastogi et al., 2018), dialogue act selection (McLeod et al., 2019), as well as conditional open-domain dialogue generation (Zeng and Nie, 2021; Ide and Kawahara, 2021).

**Meta-learning** Meta-learning involves teaching models how to learn efficiently and quickly. There are 3 broad categories of meta-learning algorithms: optimization-based (Finn et al., 2017; Nichol et al., 2018), metric-based (Snell et al., 2017; Vinyals et al., 2016), and model-based (Mishra et al., 2018; Santoro et al., 2016). Optimization-based meta-learning approaches, which involve directly updating the model’s parameters to allow for rapid adaptation to unseen tasks, have been applied to various dialogue tasks. Examples of such applications include task-oriented dialogue generation (Mi et al., 2019; Dai et al., 2020; Peng et al., 2020), domain adaptation (Qian and Yu, 2019) and dialogue state tracking (Peng et al., 2020; Huang et al., 2020).

**Personalized Dialogue Generation** There are numerous forms of personalized dialogue generation. The form covered in this paper requires leveraging both the persona information and dialogue context. Another form of personalized dialogue generation involves conditioning the response on external profile/identity information. Thus far, many different architectures (Wu et al., 2020; Song et al., 2019; Wolf et al., 2019; Kottur et al., 2017; Joshi et al., 2017) and training frameworks (Song et al., 2020a; Liu et al., 2020; Zheng et al., 2019; Su et al., 2019b) which involve utilizing the encoded persona/ or personality information and dialogue context as input have been proposed. For certain dialogue corpora such as DailyDialog (Li et al., 2017) and PERSONALDIALOG (Zheng et al., 2020), the persona descriptions are not provided. Instead, a representation of the interlocutor’s personality should be inferred from the dialogue history.

## 5 Conclusion

In this work, we proposed *MTML* and *AMTML*, 2 meta-learning frameworks which adopt our multi-task learning approach involving the addition of a persona reconstruction task. Empirical results demonstrate that both *MTML* and *AMTML* effectively increases the amount of persona information reflected in the generated dialogue responses compared to prior work. However, there is still room for improvement when it comes to the fluency and contextual coherence of the generated responses. Future work could involve improving these aspects of the responses by incorporating pretrained language models in meta-learning framework.



## References

- Daniel Adiwardana, Minh-Thang Luong, David R. So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, and Quoc V. Le. 2020. [Towards a human-like open-domain chatbot](#).
- Yinpei Dai, Hangyu Li, Chengguang Tang, Yongbin Li, Jian Sun, and Xiaodan Zhu. 2020. [Learning low-resource end-to-end goal-oriented dialog for fast and reliable system deployment](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 609–618, Online. Association for Computational Linguistics.
- Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. [Model-agnostic meta-learning for fast adaptation of deep networks](#). In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1126–1135, International Convention Centre, Sydney, Australia. PMLR.
- Sushravya GM and Shubhashis Sengupta. 2019. [Unsupervised multi-task learning dialogue management](#). pages 196–202.
- Yi Huang, Junlan Feng, Min Hu, Xiaoting Wu, Xiaoyu Du, and Shuo Ma. 2020. [Meta-reinforced multi-domain state generator for dialogue systems](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7109–7118, Online. Association for Computational Linguistics.
- Tatsuya Ide and Daisuke Kawahara. 2021. [Multi-task learning of generation and classification for emotion-aware dialogue response generation](#). *CoRR*, abs/2105.11696.
- Chaitanya K. Joshi, Fei Mi, and Boi Faltings. 2017. [Personalization in goal-oriented dialog](#). *CoRR*, abs/1706.07503.
- Satwik Kottur, Xiaoyu Wang, and Vitor Carvalho. 2017. [Exploring personalized neural conversational models](#). In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 3728–3734.
- Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu. 2017. [DailyDialog: A manually labelled multi-turn dialogue dataset](#). In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 986–995, Taipei, Taiwan. Asian Federation of Natural Language Processing.
- Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. 2016. [How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation](#). In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2122–2132, Austin, Texas. Association for Computational Linguistics.
- Qian Liu, Yihong Chen, Bei Chen, Jian-Guang Lou, Zixuan Chen, Bin Zhou, and Dongmei Zhang. 2020. [You impress me: Dialogue generation via mutual persona perception](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1417–1427, Online. Association for Computational Linguistics.
- Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. 2019. [Personalizing dialogue agents via meta-learning](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5459, Florence, Italy. Association for Computational Linguistics.
- Sarah McLeod, Ivana Kruijff-Korabayova, and Bernd Kiefer. 2019. [Multi-task learning of system dialogue act selection for supervised pretraining of goal-oriented dialogue policies](#). In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 411–417, Stockholm, Sweden. Association for Computational Linguistics.
- Fei Mi, Minlie Huang, Jiyong Zhang, and Boi Faltings. 2019. [Meta-learning for low-resource natural language generation in task-oriented dialogue systems](#). In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 3151–3157. International Joint Conferences on Artificial Intelligence Organization.
- N. Mishra, Mostafa Rohaninejad, Xi Chen, and P. Abbeel. 2018. [A simple neural attentive meta-learner](#). In *ICLR*.
- Young Na, Junekyu Park, and Kyung-Ah Sohn. 2021. [Is your chatbot perplexing?: Confident personalized conversational agent for consistent chit-chat dialogue](#). In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART*, pages 1226–1232. INSTICC, SciTePress.
- Alex Nichol, Joshua Achiam, and John Schulman. 2018. [On first-order meta-learning algorithms](#). *ArXiv*, abs/1803.02999.
- 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, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Baolin Peng, Chenguang Zhu, Chunyuan Li, Xiujuan Li, Jinchao Li, Michael Zeng, and Jianfeng Gao. 2020. [Few-shot natural language generation for task-oriented dialog](#). In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 172–182, Online. Association for Computational Linguistics.
- Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. [GloVe: Global vectors for word](#)

- representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.
- Kun Qian and Zhou Yu. 2019. [Domain adaptive dialog generation via meta learning](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2639–2649, Florence, Italy. Association for Computational Linguistics.
- Abhinav Rastogi, Raghav Gupta, and Dilek Hakkani-Tur. 2018. [Multi-task learning for joint language understanding and dialogue state tracking](#). In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 376–384, Melbourne, Australia. Association for Computational Linguistics.
- Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, and Timothy Lillicrap. 2016. [Meta-learning with memory-augmented neural networks](#). In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1842–1850, New York, New York, USA. PMLR.
- Jake Snell, Kevin Swersky, and Richard Zemel. 2017. [Prototypical networks for few-shot learning](#). In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Haoyu Song, Wei-Nan Zhang, Yiming Cui, Dong Wang, and Ting Liu. 2019. [Exploiting persona information for diverse generation of conversational responses](#).
- Haoyu Song, Wei-Nan Zhang, Jingwen Hu, and Ting Liu. 2020a. [Generating persona consistent dialogues by exploiting natural language inference](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):8878–8885.
- Yiping Song, Zequn Liu, Wei Bi, Rui Yan, and Ming Zhang. 2020b. [Learning to customize model structures for few-shot dialogue generation tasks](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5832–5841, Online. Association for Computational Linguistics.
- Feng-Guang Su, Aliyah R. Hsu, Yi-Lin Tuan, and Hung yi Lee. 2019a. [Personalized dialogue response generation learned from monologues](#). In *INTERSPEECH*.
- Feng-Guang Su, Aliyah R. Hsu, Yi-Lin Tuan, and Hung-Yi Lee. 2019b. [Personalized Dialogue Response Generation Learned from Monologues](#). In *Proc. Interspeech 2019*, pages 4160–4164.
- Anh Duong Trinh, Robert J Ross, and John D Kelleher. 2018. [A multi-task approach to incremental dialogue state tracking](#). In *Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue - Full Papers*, Aix-en-Provence, France. SEMDIAL.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#). In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Oriol Vinyals, Charles Blundell, Timothy Lillicrap, koray kavukcuoglu, and Daan Wierstra. 2016. [Matching networks for one shot learning](#). In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc.
- Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019. [Transfertransfo: A transfer learning approach for neural network based conversational agents](#). *CoRR*, abs/1901.08149.
- Bowen Wu, Mengyuan Li, Zongsheng Wang, Yifu Chen, Derek Wong, Qihang Feng, Junhong Huang, and Baoxun Wang. 2020. [Guiding variational response generator to exploit persona](#).
- M. Yang, W. Huang, W. Tu, Q. Qu, Y. Shen, and K. Lei. 2021. [Multitask learning and reinforcement learning for personalized dialog generation: An empirical study](#). *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):49–62.
- Yan Zeng and Jian-Yun Nie. 2021. [A simple and efficient multi-task learning approach for conditioned dialogue generation](#). In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4927–4939, Online. Association for Computational Linguistics.
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. [Personalizing dialogue agents: I have a dog, do you have pets too?](#) In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia. Association for Computational Linguistics.
- Yinhe Zheng, Guanyi Chen, Minlie Huang, Song Liu, and Xuan Zhu. 2019. [Personalized dialogue generation with diversified traits](#). *CoRR*, abs/1901.09672.
- Yinhe Zheng, Rongsheng Zhang, Minlie Huang, and Xiaoxi Mao. 2020. [A pre-training based personalized dialogue generation model with persona-sparse data](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):9693–9700.
- Chenguang Zhu, Michael Zeng, and Xuedong Huang. 2019. [Multi-task learning for natural language generation in task-oriented dialogue](#). In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1261–1266, Hong Kong, China. Association for Computational Linguistics.