

# PEDNet: A Persona Enhanced Dual Alternating Learning Network for Conversational Response Generation

Bin Jiang<sup>1\*</sup>, Wanyue Zhou<sup>1</sup>, Jingxu Yang<sup>1</sup>, Chao Yang<sup>1</sup>, Shihan Wang<sup>2</sup> and Liang Pang<sup>3</sup>

<sup>1</sup>College of Computer Science and Electronic Engineering, Hunan University, China

<sup>2</sup>Intelligent Systems, Utrecht University, The Netherlands

<sup>3</sup>CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences, China

{jiangbin, zhouwanyue, yangjingxu, yangchaoedu}@hnu.edu.cn  
s.wang2@uu.nl, pangliang@ict.ac.cn

## Abstract

Endowing a chatbot with a personality is essential to deliver more realistic conversations. Various persona-based dialogue models have been proposed to generate personalized and diverse responses by utilizing predefined persona information. However, generating personalized responses is still a challenging task since the leverage of predefined persona information is often insufficient. To alleviate this problem, we propose a novel Persona Enhanced Dual Alternating Learning Network (PEDNet) aiming at producing more personalized responses in various open-domain conversation scenarios. PEDNet consists of a Context-Dominated Network (CDNet) and a Persona-Dominated Network (PDNet), which are built upon a common encoder-decoder backbone. CDNet learns to select a proper persona as well as ensure the contextual relevance of the predicted response, while PDNet learns to enhance the utilization of persona information when generating the response by weakening the disturbance of specific content in the conversation context. CDNet and PDNet are trained alternately using a multi-task training approach to equip PEDNet with the both capabilities they have learned. Both automatic and human evaluations on a newly released dialogue dataset Persona-chat demonstrate that our method could deliver more personalized responses than baseline methods.

## 1 Introduction

Generating human-like conversations by chatbots has been a long-term goal in artificial intelligence. End-to-end neural generative models attract much attention as potential solutions to open-domain dialogue systems, which aim to conduct smooth and natural conversations and gain trust from users. The Sequence-to-Sequence (Seq2Seq) models (Sutskever et al., 2014; Vinyals and Le, 2015; Shang et al., 2015) have achieved success in generating semantic and syntactic responses for a given conversation context. Nevertheless, Seq2Seq tends to generate short and generic responses like “*I don’t know*” and “*I’m fine*” as they predict the response by using the maximum-likelihood estimation (MLE) objective function (Li et al., 2016a). And the generated responses are often lack of consistent personality since Seq2Seq is trained over collected conversations from different speakers (Li et al., 2016b; Zhang et al., 2018).

To tackle these issues, various methods have been proposed toward diverse and consistent response generation, by endowing open-domain chatbot with a coherent persona. For example, Zhang et al. (2018) introduced persona profiles in response generation where the persona information is used to enrich responses, and their experiments demonstrated improvements over those not leveraging persona information. Lian et al. (2019) further proposed a persona selection method, which focused on selecting the optimal persona to generate personalized responses efficiently. These efforts have obtained considerable attention and proven that leveraging the additional persona information can promote more user engagement in open-domain dialogue systems.

Table 1 gives several examples of persona-based conversation. The generated responses are supposed to be consistent with the predefined persona profile which specifies the personality that the chatbot needs to present in the conversation. However, expressing proper and rich persona information in responses is

\*Corresponding author

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>.

Persona profile	
$p_1$ : I enjoy playing tennis.	$p_3$ : I am good at playing violin.
$p_2$ : I have a dog.	$p_4$ : Green is my favorite color.
Conversation examples	
Context X1	Did you see the <u>football games</u> today?
Response Y1	Yes, I did. I caught it after my tennis match today. ( $p_1$ )
Context X2	Did you see the <u>rainbow</u> this evening?
Response X2	Yes, I did. I happened to be walking my dog at that time. ( $p_2$ )
Context X3	Did you see the <u>street band performing</u> yesterday?
Response X3	No, I did not. I am too busy practicing the violin. ( $p_3$ )

Table 1: Examples of persona-based conversation.

usually not so easy in practice. The model needs to carefully consider the specific content (e.g., “*football games*” in context  $X1$ , “*the street band performing*” in context  $X3$ ) in the context to select which persona to use as the character of the current conversation. On the other hand, from Table 1 we can see that responses are mainly dominated by the persona of the speaker in many conversation scenarios, which means once the persona is determined, the model needs to focus more on how to express this persona effectively. In that case, the specific information in the conversation context will become less significant for the response generation.

In this paper, we propose a novel Persona Enhanced Dual Alternating Learning Network (PEDNet) to achieve an appropriate and rich persona expression in responses. PEDNet is composed of a Context-Dominated Network (CDNet) and a Persona-Dominated Network (PDNet). They are built upon a common encoder-decoder backbone with shared parameters. Specifically, CDNet is a general persona-based dialogue model equipped with a Memory Network (Sukhbaatar et al., 2015) which is capable of reading and conditioning on persona profile. It is reinforced with a persona-aware representation to select persona properly. Different from CDNet, PDNet generates responses by conditioning on the optimal persona pre-labeled by BERT (Devlin et al., 2019) and a general representation of the conversation context that contains no specific content (e.g., “*Do you see...today?*” of context  $X1$  in Table 1). PDNet is specially designed to emphasize the contribution of persona in dialogue modeling by ignoring the specific content of the context, while CDNet ensures the semantical relevance between the context and the generated response. PEDNet can learn the capabilities of each sub-network by training CDNet and PDNet alternately with a multi-task training approach.

The main contributions of this work are summarized as follows:

- We propose a novel Persona Enhanced Dual Alternating Learning Network (PEDNet) to learn an appropriate and rich persona expression in the generated response, by explicitly distinguishing the effects of the specific content in conversation context on the persona selection and persona incorporation respectively.
- The proposed PEDNet consists of two backbone-shared sub-networks with different learning purposes: CDNet learns to select the persona properly and PDNet focuses more on enhancing the utilization of the persona information. CDNet and PDNet are trained alternately with a multi-task training approach to benefit PEDNet with the both capabilities they have learned.
- Experimental results on the Persona-chat dataset show that our model can generate more personalized responses than baselines. As a side contribution, we will release our code <sup>1</sup> to facilitate the research community.

## 2 Related Work

The success of Seq2Seq motivates the development of the research for improving the quality of response generation. However, the problem of tending to generate generic or inconsistent responses still remains in

<sup>1</sup>Code available at: <https://github.com/zwyccodes/PEDNet>.

this domain. One promising approach to alleviate such a challenge is to find a better objective function (Li et al., 2016a). Besides, learning the inherent attributes is another way to improve the diversity of responses and ensure the contextual consistency. Li et al. (2016b) first proposed a persona-based model by using user embeddings, which projects each user into a vector and feeds the vector to the decoder at each decoding step. Zhang et al. (2017) employed a two-phase training approach, which initialized the model using large scale data and then fine-tuned the model to generate personalized responses. Kottur et al. (2017) explored a response generation model conditioned on both speakers and context history. Nevertheless, these models cannot explain how the personality expressed in the response is captured since all the information about the user is encoded as a dense vector for decoding.

To maintain a coherent personality, Qian et al. (2018) designed a model explicitly expressing a profile value in the response according to its prespecified persona profile. Zhang et al. (2018) created a Persona-chat dataset, and they proposed two generative models to incorporate persona information into responses. Yavuz et al. (2019) used a copy mechanism which allowed the decoder to hierarchically attend and copy from the external persona profile in addition to the conversation context. A persona selection mechanism was proposed in Lian et al. (2019), where both prior and posterior distributions over persona are used to facilitate persona selection. Song et al. (2019) proposed a memory-augmented architecture to exploit persona information from context and incorporate a conditional variational autoencoder model together to generate diverse responses.

So far, the studies all generate responses conditioning on both the context and the persona without distinction and emphasis, which often result in insufficient utilization of persona information. Different from those previous works, we design a dual alternating learning network, which explicitly distinguishes the different effects of specific content in the context on the two aspects of persona selection and persona incorporation, so as to realize an effective persona expression.

### 3 Model

#### 3.1 Task Definition

We follow the definition of persona in Zhang et al. (2018) and work on the explicit textual persona. The task can be formulated as follows: given a context  $X = (x_1, \dots, x_n)$  with  $n$  words and a persona profile formed of  $M$  persona texts  $P = \{p_1, p_2, \dots, p_M\}$ , the system is supposed to generate a response  $Y = (y_1, \dots, y_m)$  with  $m$  words that is coherent with this persona profile. The response generation process can be briefly stated as:

$$\mathbf{P}(Y|X, P) = \prod_{t=1}^m \mathbf{P}(y_t|y_1, \dots, y_{t-1}, X, P) \quad (1)$$

#### 3.2 Encoder-Decoder Backbone

Our model is implemented with a GRU-based (Cho et al., 2014; Chung et al., 2014) Seq2Seq backbone, which consists of two components: an encoder and a decoder.

The encoder aims to encode a context  $X$  to a sequence of hidden states. We define  $u^{(X)}, \mathbf{h}^{(X)} = \text{Encode}(X)$ , where  $u^{(X)}$  denotes the final state of the GRU and  $\mathbf{h}^{(X)} = (h_1, \dots, h_n)$  denotes the outputs of the GRU cell at all steps. The decoder takes as input a context vector  $c_t$  and the embedding of a previously decoded word  $y_{t-1}$  to update its hidden state  $s_t$  with another GRU. The context vector  $c_t$  is an attention-weighted combination of  $\mathbf{h}^{(X)}$  and dynamically attends on context information during decoding (Bahdanau et al., 2015). We define this operation as  $c_t = \text{Attention}(s_{t-1}, \mathbf{h}^{(X)})$ . Once the state vector  $s_t$  is obtained, the decoder generates a token according to the output probability distribution over the vocabulary:

$$y_t \sim o_t = \text{softmax}([s_t; c_t]) \quad (2)$$

where  $[\cdot; \cdot]$  represents a vector concatenation.

#### 3.3 Persona Enhanced Dual Alternating Learning Network

PEDNet is proposed to improve the persona expression in responses and it consists of two sub-networks: PDNet and CDNet. The architecture overview of PEDNet is given in Figure 1. First, CDNet is used to

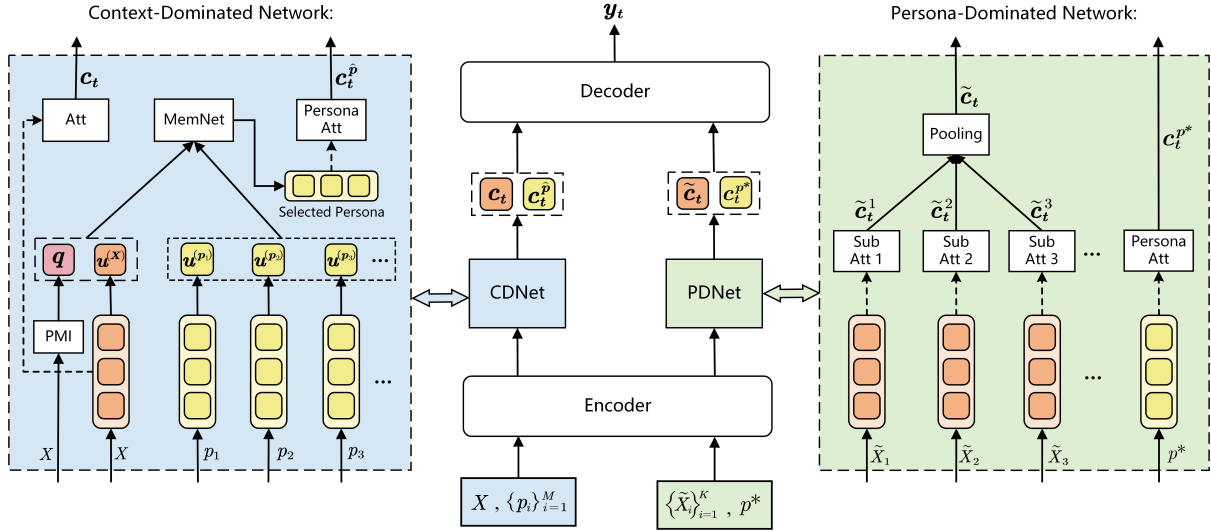


Figure 1: The architecture of the proposed PEDNet. CDNet and PDNet share a common encoder-decoder backbone and are trained alternately using a multi-task training approach.  $X$  is the conversation context,  $\{p_i\}_{i=1}^M$  is the predefined persona profile,  $\{\tilde{X}_i\}_{i=1}^K$  is the congeneric contexts set retrieved from the corpus,  $p^*$  is the optimal persona pre-labeled by BERT.

select a persona and ensure that the generated response is semantically relevant to the current context. Besides, in order to prevent the specific content in the context from dispersing the model’s attention in persona incorporation, PDNet takes as input a pre-given optimal persona and a general representation of the conversation context without the specific content. PEDNet can be benefited by both PDNet and CDNet to learn an effective persona expression through a multi-task training approach during the training phase. In the testing phase, PEDNet directly generates responses based on CDNet.

### 3.3.1 Specific Content Extraction

As mentioned in Mou et al. (2016), Pointwise Mutual Information (PMI) (Church and Hanks, 1990) is an appropriate statistic for specific words prediction, which is also adopted in our work to extract the context words which have high mutual information against the persona profile. Suppose that we have an input sample  $(X, Y, P)$ , given a word  $w_x$  in  $X$  and a word  $w_p$  in  $P$ , the  $\text{PMI}(w_x, P)$  for the word  $w_x$  against the persona profile  $P$  is calculated below:

$$\text{PMI}(w_p, w_x) = \log \frac{\mathbf{P}(w_p, w_x)}{\mathbf{P}(w_p)\mathbf{P}(w_x)} = \log \frac{\mathbf{P}(w_p|w_x)}{\mathbf{P}(w_p)} \quad (3)$$

$$\text{PMI}(w_x, P) = \sum_{w_p \in P} \text{PMI}(w_p, w_x) \quad (4)$$

Finally, we choose the specific words  $Q$  by setting a threshold<sup>2</sup> to filter out those words with low PMI scores and limit the maximum number of specific words to 5.

### 3.3.2 Context-Dominated Network

This network generates responses conditioning on both the context  $X$  and persona profile  $P = \{p_1, p_2, \dots, p_M\}$ . Figure 1 depicts the details of CDNet. Specifically, CDNet needs to select a persona  $p_i$  and then generates a response by integrating  $p_i$ . The details are as follows:

We convert  $X$  and each persona  $p_i$  into vector representations using a context encoder and a persona encoder respectively:

<sup>2</sup>In our experiments, the specific words extraction works best when the threshold is equal to 3.5, if there is no word in the context reaching the PMI threshold, the specific words will be none.

$$u^{(X)}, \mathbf{h}^{(X)} = \text{Encode}(X) \quad (5)$$

$$u^{(p_i)}, \mathbf{h}^{(p_i)} = \text{Encode}(p_i) \quad (6)$$

Then we compute a specific content representation  $q = \sum_{w \in Q} \Psi(w)$  by a summation over word embeddings in  $Q$ , where  $\Psi(\cdot)$  transfers a word to its embedding. After that, we obtain a persona-aware representation  $z$  by concatenating  $u^{(X)}$  and  $q$ . We select a persona by a multi-layer Memory Network:

$$\text{prob}_i^k = \text{softmax}(m^k u^{(p_i)}) \quad (7)$$

$$o^k = \sum_{i=1}^M \text{prob}_i^k u^{(p_i)} \quad (8)$$

$$m^{k+1} = o^k + m^k \quad (9)$$

where  $k$  denotes the layer number and  $m^1 = z^T$ . In this work, a 3-layer Memory Network is used<sup>3</sup>. We select the persona with the highest probability in the last layer of the memory network and fit it with the optimal persona:  $\hat{p} = p_j$ , where  $j = \text{argmax}_i(\text{prob } b_i^3)$ .

Finally, the decoder generates a response by the selected persona  $\hat{p}$ . The hidden state of the decoder at time  $t$  is:

$$c_t^{\hat{p}} = \text{Attention}(s_{t-1}, \mathbf{h}^{(\hat{p})}) \quad (10)$$

$$c_t = \text{Attention}(s_{t-1}, \mathbf{h}^{(X)}) \quad (11)$$

$$s_t = \text{GRU}([e(y_{t-1}); c_t; c_t^{\hat{p}}], s_{t-1}) \quad (12)$$

where  $e(y_{t-1})$  is the word embedding of  $y_{t-1}$ .

### 3.3.3 Persona-Dominated Network

Compared with CDNet, the inputs of PDNet are made some adjustments. Specifically, we remove the specific content in  $X$  by replacing the specific words  $Q$  with special symbols “<KEY>”, and  $X$  is changed into a revised form  $\tilde{X}$ . Since a single context may have difficulty in providing sufficient general content, we further retrieve extra  $K-1$  revised contexts which are similar to  $\tilde{X}$  from the corpus using Gensim<sup>4</sup>, to construct a congeneric contexts set  $\{\tilde{X}_i\}_{i=1}^K$ . Meanwhile, we use a pre-trained BERT model to calculate the cosine similarity between  $Y$  and each persona text in  $P$ , so as to select a text  $p^*$  with the highest similarity score as the optimal persona for the current conversation. In calculating cosine similarity, the sentence representations of  $Y$  and  $P$  are obtained by averaging the outputs of the second last hidden layer of BERT. Thereafter, we obtain a new model input  $\{\{\tilde{X}_i\}_{i=1}^K, Y, p^*\}$ , where  $p^*$  represents the optimal persona of this turn conversation. The details of PDNet is also given in Figure 1. Similar to CDNet,  $\{\tilde{X}_i\}_{i=1}^K$  and  $p^*$  are first encoded respectively:

$$u^{(\tilde{X}_i)}, \mathbf{h}^{(\tilde{X}_i)} = \text{Encode}(\tilde{X}_i) \quad (13)$$

$$u^{(p^*)}, \mathbf{h}^{(p^*)} = \text{Encode}(p^*) \quad (14)$$

These vectors  $\{u^{(\tilde{X}_i)}\}_{i=1}^K$  are transformed into a general representation of the revised context  $\tilde{X}$  through an average pooling operation:  $s_0 = \frac{1}{K} u^{(\tilde{X}_i)}$ . The  $s_0$  is used as the initial hidden state of the decoder.

<sup>3</sup>In our experiments,  $k = 3$  performs better than  $k = 1$  or 2, but there is no significant improvement when  $k = 4$  or 5.

<sup>4</sup>Gensim is an open-source library for unsupervised topic modeling and natural language processing. It is available at <https://radimrehurek.com/gensim/>.

In the decoding process, the decoder generates the response word by word sequentially by incorporating the optimal persona  $p^*$ . The hidden state of the decoder at time  $t$  is:

$$c_t^{p^*} = \text{Attention}(s_{t-1}, \mathbf{h}^{(p^*)}) \quad (15)$$

$$\tilde{c}_t^i = \text{Attention}(s_{t-1}, \mathbf{h}^{(\tilde{X}_i)}) \quad (16)$$

$$\tilde{c}_t = \frac{1}{K} \sum_{j=1}^K \tilde{c}_t^j \quad (17)$$

$$s_t = \text{GRU}([e(y_{t-1}); \tilde{c}_t; c_t^{p^*}], s_{t-1}) \quad (18)$$

where  $\tilde{c}_t$  is computed by a series of sub-attentions, and each sub-attention is responsible for attending the details of those congeneric contexts.

### 3.4 Multi-Task Training

As mentioned in Luong et al. (2016), multi-task learning for Seq2Seq can improve the performance of a task using other related tasks. In our work, PDNet and CDNet are tied together in the training phase through a multi-task training approach. We use two independent tasks to train these two networks alternately, where the encoder-decoder backbone can be benefited from both:

- The Context-Dominated task: We expose CDNet with  $\{X, Y, P\}$  training examples.
- The Persona-Dominated task: We expose PDNet with  $\{\{\tilde{X}_i\}_{i=1}^K, Y, p^*\}$  training examples.

Specifically, in each batch, all training data is sampled from one task only. For each batch, we randomly select a task from these two tasks, where the probability of the Persona-Dominated task is  $\gamma$  and the probability of the Context-Dominated task is  $1 - \gamma$ .

The loss function is the cross-entropy error between the predicted token distribution and the gold distribution in the training corpus. As aforementioned in Section 3.3.3, the optimal persona text is labeled in advance, so an additional cross-entropy loss is also added to ensure the accuracy of persona selection for CDNet.

## 4 Experiments

### 4.1 Dataset

We conduct our experiments on a recently created dataset, namely the Persona-chat dataset (Zhang et al., 2018). In Persona-chat, each conversation was constructed from crowd-workers who were randomly paired. To produce personalized conversations, each worker was asked to chat with each other conditioned on a given persona profile. This dataset contains 164,356 utterances over 10,981 dialogs and has 1155 persona profiles, each consisting of a set of persona texts. We adapt this dataset into a single-turn conversation paradigm and each turn corresponds to a context-response pair and a persona profile. We divide this dataset into 65,490/7800/7512 turns for training/validation/testing.

### 4.2 Models for Comparison

We compare the proposed PEDNet with the following baseline models:

- Seq2Seq: a vanilla Seq2Seq with attention, which does not incorporate external persona profiles. (Vinyals and Le, 2015).
- Generative Profile Memory Network (GPMN): a persona-based dialogue model, where the persona is stored as memory representation in a Memory Network and fed into the decoder with attention. (Zhang et al., 2018).
- Posterior Knowledge Selection (PostKS): a knowledge-grounded model<sup>5</sup> where both prior and posterior distributions over persona are used to facilitate persona selection. (Lian et al., 2019).

<sup>5</sup>In PostKS, the persona is served as knowledge.

Among these baselines, Seq2Seq is compared for demonstrating the effect of integrating persona profiles in response generation while other models, which incorporate persona profiles, are compared to verify that the effectiveness of our PEDNet.

### 4.3 Implementation Details

We use PyTorch to implement the proposed model. We apply 2-layer GRU structures with 800 hidden states for both the encoder and decoder with a dropout rate of 0.2. We set the size of word embedding to 300 and initialize it by using GloVe (Pennington et al., 2014). The input contexts, generated responses and persona texts share the same word embedding layer. The vocabulary size is limited to 20,000. The Adam optimizer is used to update the gradient and the gradient is clipped in 5.0. We train the model with a minibatch size of 128 and an initial learning rate of 0.0005. We choose the model parameters that yield the highest score to each metric on our validation set.

### 4.4 Automatic Evaluation

Model	Distinct-1/2	P-R/P/F1	P-Cover
Seq2Seq	0.018/0.064	0.048/0.030/0.037	0.016
GPMN	0.028/0.160	0.082/0.048/0.060	0.034
PostKS	0.031/0.174	0.111/0.068/0.084	0.037
PEDNet ( $K = 1$ )	0.031/0.165	0.128/0.077/0.096	0.040
PEDNet ( $K = 5$ )	0.031/0.175	0.136/0.082/0.102	0.044
PEDNet ( $K = 10$ )	<b>0.032/0.177</b>	0.141/0.086/0.107	0.044
PEDNet ( $K = 15$ )	0.031/0.173	<b>0.143/0.088/0.109</b>	<b>0.045</b>
PEDNet ( $K = 20$ )	0.030/0.164	0.133/0.082/0.101	0.041

Table 2: The automatic evaluation on the Persona-chat dataset.  $K$  is the number of the contexts in the congeneric contexts set.

**Metrics:** We adopt several automatic metrics to evaluate the models, and the results are summarized in Table 2. Since the goal of the proposed model is to generate personalized responses, we use Distinct-1/2 (Li et al., 2016a) to evaluate the word-level diversity of generated responses. Besides, we also use Persona Recall/ Precision/F1 and Persona Coverage, adapted from Dinan et al. (2018) and Song et al. (2019) respectively, to evaluate how well the selected persona is expressed. Specifically, given the set of non-stop words in generated response  $Y$  and each persona text  $\{p_i\}_{i=1}^M$  defined as  $W_Y$  and  $\{W_{p_i}\}_{i=1}^M$ , the Persona F1 (P-F1) is computed by  $\text{Persona Recall} = \frac{\max_{i \in [1, M]} |W_Y \cap p_i|}{|W_{p_i}|}$  and  $\text{Persona Precision} = \frac{\max_{i \in [1, M]} |W_Y \cap p_i|}{|W_Y|}$ .

$$\text{Persona } F_1 = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (19)$$

where  $|W_Y|$  and  $|W_{p_i}|$  denote the amount of words in  $W_Y$  and  $W_{p_i}$  respectively, and  $W_Y \cap p_i$  is the words shared by  $W_Y$  and  $W_{p_i}$ . The Persona Coverage (P-Cover) is computed by:

$$\text{Persona Coverage} = \max_{i \in [1, M]} \frac{\sum_{w_j \in W_Y \cap p_i} (idf_j)}{|W_Y \cap p_i|} \quad (20)$$

where  $idf_j$  is the inverse document frequency for  $w_j \in W_Y \cap p_i$ :  $idf_j = 1/(1 + \log(1 + tf_j))$ , and  $tf_j$  is from the GloVe index via Zipf’s law (Zhang et al., 2018).

**Results:** We test several variants of PEDNet to explore the effect of different values of  $K$  on model performance. As shown in Table 2, even if  $K = 1$ , our model still outperforms all baselines in persona expression performance as well as performs well in diversity. When  $K = 15$ , our model obtains the best performance in persona expression compared with all the baseline models. For instance, comparing PEDNet against PostKS, Persona F1 is increased from 0.084 (PostKS) to 0.109 (PEDNet) and Persona

Coverage also achieves a good improvement from 0.037 (PostKS) to 0.045 (PEDNet), which demonstrate our model’s ability in enriching persona information in responses. In our experiments, when the value of  $K$  exceeds 20, the performance of the model begins to decline. This is probably because there are many contexts without enough similarity to the current context in the congeneric contexts set, leading to an information scattering of the general representation. In order to better evaluate our model, the following experiments are all based on  $K = 15$ .

#### 4.5 Human Evaluation

Model	Rel.	Per-Exp.
Seq2Seq	0.873	0.652
GPMN	1.028	0.792
PostKS	1.047	0.867
PEDNet	<b>1.085</b>	<b>0.983</b>

Table 3: The human evaluation on *Relevance* (Rel.) and *Persona Expression* (Per-Exp.).

In order to further evaluate the performance of our method, we also perform a human evaluation study. Given a context and a persona profile, we generate responses from all the baseline models and our model. These responses are presented to three human annotators along with the context and the persona profile.

**Metrics:** Three annotators have been applied to score the quality of the responses in terms of *Relevance* and *Persona Expression*. The rating ranges from 0 to 2, where 0 means not good, 1 means acceptable, and 2 means excellent. *Relevance* is defined to measure whether the response is fluency and relevant to its context. *Persona Expression* is defined to measure whether the response displays sufficient persona information.

**Annotation Statistics:** We randomly sample 200 responses by each model, resulting in 800 responses in total for human annotation. We calculate the Fleiss’ kappa (Fleiss, 1971) to measure inter-rater consistency. Fleiss’ kappa for *Relevance* and *Persona Expression* is 0.496 and 0.648, indicating “Moderate agreement” and “Substantial agreement” respectively.

**Results:** The results are shown in Table 3. Our model outperforms all the baselines in both metrics. Specifically, the performance of our model on *Persona Expression* is improved from 0.867 (PostKS) to 0.983 (PEDNet), indicating that PEDNet can further enrich the persona information of generated responses. Besides, the performance on *Relevance* of PEDNet is also better than other baselines, which shows that the PEDNet can also ensure that the generated response is semantically relevant to the current context.

#### 4.6 Ablation Test

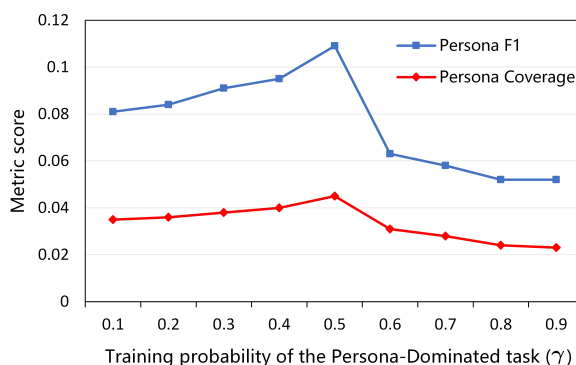


Figure 2: The ablation results for the training probability of the Persona-Dominated task.

To investigate the influence of PDNet for persona expression, we also conduct an ablation test where the training probability of the Persona-Dominated task is set from 0.1 to 0.9 (refer to  $\gamma$  in Section 3.4).



The results are shown in Figure 2. As we can see, when increasing the training probability of the Persona-Dominated task, the metric scores of Persona F1 and Persona Coverage achieve the best values with the probability equal to 0.5. And as this training probability continues to increase, it begins to drop dramatically. We analyze the reasons as follows: If the training probability of the Persona-Dominated task is very small, the model is mainly trained by CDNet, without the assistance of PDNet, the model performs comparatively worse in persona expression, which proves the effectiveness of PDNet. On the contrary, if the training probability of the Persona-Dominated task reaches a higher level, the model is mainly trained by PDNet, which results in the model being unable to see the complete context during the training process, so the model cannot learn to choose the appropriate persona and generate the appropriate responses.

#### 4.7 Case Study

Context	Hi, do you have sons? I have two.	Wow. That is super cool. I prefer movies to books though.	
Persona	<p><math>p_1</math>: I have a father and a brother.</p> <p><math>p_2</math>: My favorite food is sushi.</p> <p><math>p_3</math>: I listen to rap music.</p> <p><math>p_4</math>: I like to work out.</p>	<p><math>p_1</math>: I own my very successful dieting business.</p> <p><math>p_2</math>: I spend my free time reading plays.</p> <p><math>p_3</math>: I also write diet cookbooks.</p> <p><math>p_4</math>: I enjoy exercising for long hours.</p>	
Response	Seq2Seq	I do not have any kids.	I also like to watch movies.
	GPMN	I do. I am busy with work outside.	I am a personal trainer. What do you do?
	PostKS	I have two sons and a wife.	I like to read books mostly.
	PEDNet	I do not have a child but I have a brother, do you?	I read a lot of plays. I am more of a reader.

Table 4: Examples of the Generated Responses on the Persona-chat dataset.

We further conduct a case study to analyze the quality of the generated responses. Some sampled cases are shown in Table 4. As we can see, the responses predicted by PEDNet contain richer persona information than other baselines. In the first example, the context is asking whether the chatbot has sons. There is no persona directly related to “sons” in the given persona profile, except that the first persona “*I have a father and a brother*” is regarded as an indirectly related persona of the same type of conversation topic. Compared with other baselines, our PEDNet generates a response that is not only semantically appropriate to its context “*I do not have a child*”, but also display adequate persona content (i.e., “*but I have a brother*”). The second example also proves the superiority of PEDNet in persona expression.

## 5 Conclusion and Future Work

In this paper, we focus on the task of personalized response generation and propose the Persona Enhanced Dual Alternating Learning Network to improve the explicit expression of personality. Two networks that learn through a multi-task training approach are designed to select the appropriate persona according to the conversation context, as well as enhance the utilization of persona information in response generation. Experimental results on both automatic and human evaluations verify the effectiveness and usefulness of our model. For future work, we plan to consider the long-distance conversation history to improve the persona selection and extend our persona integration method in multi-turn conversations.

## Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under grant 62072169 and 61702176, and CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences under grant CASNDST202001.

## References

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *Proceedings of the 3rd International Conference on Learning Representations, 2015*, San Diego, CA, USA.
- Kyunghyun Cho, Bart van Merriënboer, Çağlar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *Proceedings of the 2014 conference on Empirical Methods in Natural Language Processing 2014*, Doha, Qatar.
- Junyoung Chung, Çağlar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *Computing Research Repository*, arXiv:1412.3555.
- Kenneth Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1):22–29.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics 2019*, pages 4171–4186, Seattle, Washington.
- Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2018. Wizard of wikipedia: Knowledge-powered conversational agents. *Computing Research Repository*, arXiv:1811.01241.
- Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378.
- Satwik Kottur, Xiaoyu Wang, and Vítor Carvalho. 2017. Exploring personalized neural conversational models. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence 2017*, pages 3728–3734, Melbourne, Australia.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016a. A diversity-promoting objective function for neural conversation models. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics 2016*, pages 110–119, San Diego, California.
- Jiwei Li, Michel Galley, Chris Brockett, Georgios Spithourakis, Jianfeng Gao, and Bill Dolan. 2016b. A persona-based neural conversation model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics 2016*, pages 994–1003, Berlin, Germany.
- Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to select knowledge for response generation in dialog systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence 2019*, pages 5081–5087, Macao, China.
- Minh-Thang Luong, Quoc V. Le, Ilya Sutskever, Oriol Vinyals, and Lukasz Kaiser. 2016. Multi-task sequence to sequence learning. In *Proceedings of the 4th International Conference on Learning Representations 2016*, San Juan, Puerto Rico.
- Lili Mou, Yiping Song, Rui Yan, Ge Li, Lu Zhang, and Zhi Jin. 2016. Sequence to backward and forward sequences: A content-introducing approach to generative short-text conversation. In *Proceedings of the 26th International Conference on Computational Linguistics 2016*, pages 3349–3358, Osaka, Japan.
- Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on Empirical Methods in Natural Language Processing 2014*, pages 1532–1543, Doha, Qatar.
- Qiao Qian, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018. Assigning personality/profile to a chatting machine for coherent conversation generation. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence 2018*, Stockholm, Sweden.
- Lifeng Shang, Zhengdong Lu, and Hang Li. 2015. Neural responding machine for short-text conversation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing 2015*, pages 1577–1586, Beijing, China.
- Haoyu Song, Weinan Zhang, Yiming Cui, Dong Wang, and Ting Liu. 2019. Exploiting persona information for diverse generation of conversational responses. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence 2019*, Macao, China.

- Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus. 2015. End-to-end memory networks. In *Proceedings of the Advances in Neural Information Processing Systems 2015*, pages 2440–2448, Montreal, Canada.
- Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Proceedings of the Advances in Neural Information Processing Systems 2014*, pages 3104–3112, Montreal, Canada.
- Oriol Vinyals and Quoc Le. 2015. A neural conversational model. *Computing Research Repository*, arXiv:1506.05869.
- Semih Yavuz, Abhinav Rastogi, Guan-Lin Chao, and Dilek Hakkani-Tur. 2019. Deepcopy: Grounded response generation with hierarchical pointer networks. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue 2019*, pages 122–132, Stockholm, Sweden.
- Weinan Zhang, Ting Liu, Yifa Wang, and Qingfu Zhu. 2017. Neural personalized response generation as domain adaptation. *Computing Research Repository*, arXiv:1701.02073.
- 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 2018*, pages 2204–2213, Melbourne, Australia.