A Pre-training Based Personalized Dialogue Generation Model with Persona-sparse Data

Yinhe Zheng [1,3*]**, Rongsheng Zhang**[2*]**, Xiaoxi Mao**[2]**, Minlie Huang**[1†]

[1] Institute for Artifical Intelligence, State Key Lab of Intelligent Technology and Systems.
Beijing National Research Center for Information Science and Technology.
Department of Computer Science and Technology, Tsinghua University, Beijing, China.
[2] Fuxi AI Lab, NetEase Inc., Hangzhou, China.
[3] Samsung Research China - Beijing (SRC-B), Beijing, China.
yh.zheng@samsung.com, zhangrongsheng@corp.netease.com, maoxiaoxi@corp.netease.com,
aihuang@tsinghua.edu.cn

Abstract

Endowing dialogue systems with personas is essential to deliver more human-like conversations. However, this problem is still far from well explored due to the difficulties of both embodying personalities in natural languages and the persona sparsity issue observed in most dialogue corpora. This paper proposes a pre-training based personalized dialogue model that can generate coherent responses using persona-sparse dialogue data. In this method, a pre-trained language model is used to initialize an encoder and decoder, and personal attribute embeddings are devised to model richer dialogue contexts by encoding speakers' personas together with dialogue histories. Further, to incorporate the target persona in the decoding process and to balance its contribution, an *attention routing* structure is devised in the decoder to merge features extracted from the target persona and dialogue contexts using dynamically predicted weights. Our model can utilize persona-sparse dialogues in a unified manner during the training process, and can also control the amount of persona-related features to exhibit during the inference process. Both automatic and manual evaluation demonstrates that the proposed model outperforms state-of-the-art methods for generating more coherent and persona consistent responses with persona-sparse data.

Introduction

Building a "human-like" conversation system has been an important topic in artificial intelligence, where one of the major challenges is to present a consistent persona so that the system can interact with users in a more natural way to gain users' confidence and trust. The user engagement of a dialogue agent increases when the agent is conditioned on various persona settings, including age, gender, language, location, or even a proper accent (**?**; **?**; **?**; **?**). Various approaches have been explored to personalize a dialogue system (**?**; **?**; **?**).

Recent advances in pre-training methods have led to state-of-the-art results in a range of natural language processing tasks (**?**; **?**; **?**; **?**). Promising results are also obtained by
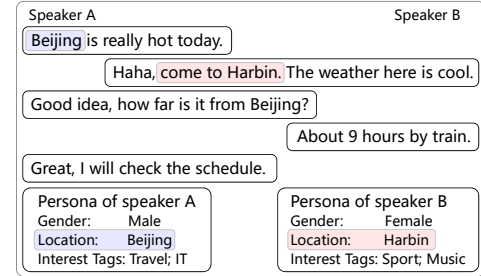
Figure 1: An example dialogue session and the personal profile of each speaker. Words in responses are in the same color with the corresponding personal attributes.

applying these approaches in personalized dialogue generation models, such as fine-tuning a pre-trained model on a small set of persona-related dialogues (e.g. PERSONA-CHAT (**?**)) (**?**; **?**; **?**). However, the dialogue data used in the fine-tuning stage of these methods are usually crowd-sourced, where speakers are required to exchange their personas within limited turns of conversation. This data collection scheme is guaranteed to yield dialogues that cover rich persona related features (i.e., "**persona-dense**") and thus facilitate fine-tuning directly. However, such a setting is expensive and can only produce a limited amount of dialogues. Further, models fine-tuned using these data tend to over-fit to the routine that persona-related features should be exhibited in every response. This does not meet the common practice observed in our daily communication.

As a matter of fact, most speakers in our daily conversations are not aiming to exhibit their personas within limited turns of interactions, namely, real-world dialogues are not always persona-related. For example, as shown in the dialogue session of Figure 1, speaker A and B only reveal their personas in the first turn of the conversation, while the rest part of this dialogue is not persona-related. Therefore, we argue that data collected from real-world conversations only contain a limited amount of dialogues that relate to speakers' persona. In other words, real dialogue data are "**persona-sparse**". Directly fine-tuning on these real-world conversations may mislead the model to focus on dialogues

that are not persona-related, since these dialogues are in the majority. Further, speakers' personas may be regarded as the noises and tend to be ignored by the dialogue model since they are not related to most responses.

To address the above issues, we propose a pre-training based method that can utilize persona-sparse data to build a personalized dialogue agent. Specifically, the dialogue data we use come from the daily conversations on social media, where speakers are not asked to reveal their personas *intentionally*. This differs from previous pre-training based approaches that utilize crowdsourced dialog data such as PERSONA-CHAT (**?**), which is persona-dense and thus a direct fine-tuning process will suffice for the pre-trained dialogue model to capture persona related features (**?**; **?**).

In this paper, we adopt the encoder-decoder framework and use a pre-trained language model to initialize an encoder and decoder. Attribute embeddings are added in the encoder to capture rich persona related features when modeling dialogue histories, and an *attention routing* mechanism is proposed in the decoder to incorporate the target persona in the decoding process. Three attention routes are used in this study and each route models a certain source of features, i.e., features extracted from the target persona, dialogue histories, and previously decoded tokens. A dynamic weight predictor is built to weigh the output of each route, so that the contribution of the target persona in the final output can be balanced. In this manner, we can leverage persona-sparse dialogue data in the training stage and control the amount of persona information to exhibit in the inference stage. Automatic and manual evaluation indicates that our method can produce dialogue responses that are more coherent and contain richer persona features.

Our main contributions can be summarized as follows:

1. We propose a pre-training based method that can utilize persona-sparse data to build personalized dialogue models. Our method can take full advantage of the pre-trained model for generating diverse and coherent dialogues, while effectively leveraging real-world data that are persona-sparse to produce persona-related responses.

2. We propose an attention routing mechanism to weigh persona features dynamically in the decoder. It allows us to utilize persona-sparse dialogue data in a unified manner during the training process and to control the amount of persona-related features to exhibit in the decoded responses.

3. Both automatic and manual evaluation shows that our method outperforms previous methods in producing more coherent and persona-related responses.

Related Work

Personalized Dialogue Models: Traditional studies to build personalized dialogue agents focused on psychology inspired approaches, such as modeling "Big Five" of speakers (**?**). However, such psychological metrics are too subjective to model and the corresponding dialogue data are extremely difficult to collect. This limits the application of these approaches in building large-scale personalized dialogue systems.

Recent studies try to tackle the personalized dialogue generation problem in a data-driven manner, i.e., learning persona related features directly from large-scale dialogue datasets. Early attempts in this direction focused on modeling characters in movie dialogues (**?**), while recent studies took advantages of the sequence to sequence learning framework (**?**; **?**) to model a speaker's persona by utilizing social media data (**?**). Specifically, persona in these studies was either implicitly modeled using a persona embedding (**?**; **?**; **?**) which requires sufficient data from each speaker, or explicitly given as the input (**?**; **?**). Some models were also proposed to personalize task-oriented dialogue systems (**?**). However, these models were all trained from scratch without a pre-training process, where the benefits of using language models that are pre-trained with large corpora are yet to be explored.

Pre-training Methods: Recent advances in the pre-training methods have led to state-of-the-art results in many tasks (**?**; **?**; **?**; **?**). Various pre-training approaches have also been proposed in the dialogue modeling task (**?**). Specifically, **?** (**?**) proposed two pre-training objectives to boost dialogue tasks; **?** (**?**) adapted the pretrained GPT2 model (**?**) to multi-domain task-oriented dialogues. As for personalized dialogue modeling, **?** (**?**) and **?** (**?**) showed that the pre-trained GPT model (**?**) can significantly improve the quality of the generated dialogues by fine-tuning on a small persona-dense dialogue dataset.

Compared to ours, the most relevant prior work was done by **?** (**?**). However, their method requires a direct fine-tuning process on a persona-dense dataset, while our study can make use of the persona-sparse dialogues with the proposed dynamic weighting scheme. Further, we also add explicit attribute embeddings to model structured personas when encoding dialogue contexts, whereas their approaches do not consider speakers' personas when modeling dialogue contexts.

Model

Our study aims at generating a fluent response Y that is coherent with a given dialogue context C and an explicitly represented target persona T of the responder:

$$Y = \arg\max_{Y'} P(Y'|C, T) \qquad (1)$$

Specifically, the persona T can be regarded as a set of attributes (such as gender, location, or personal interest) $T=\{t_1, t_2, ..., t_N\}$ and each attribute is represented as a key-value pair $t_i=\langle k_i, v_i \rangle$. The dialogue context $C=\{(U_1, T_1), ..., (U_M, T_M)\}$ contains several turns of conversations (i.e., utterances U_i) and the persona T_i of the associated speaker.

Figure 2 shows an overview of our model. The encoder and decoder used in our study follow the Transformer framework (**?**), and share the same set of weights. The encoder is used to encode the dialogue context C into a context encoding E_C and the target persona T into a persona encoding E_T, independently. Attribute embeddings are added when producing E_C. The decoder takes as input E_C and E_T and decodes the output in an auto-regressive way. An
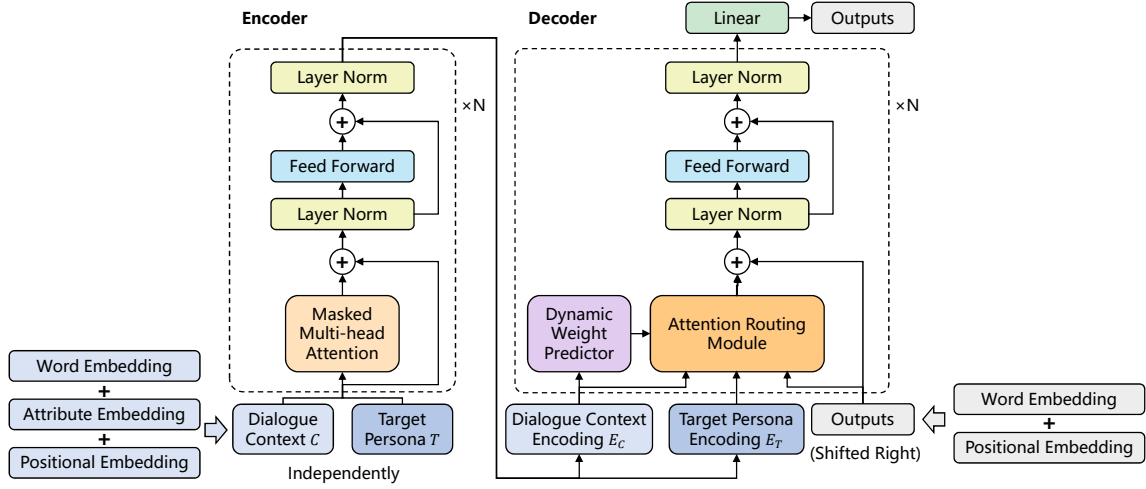
Figure 2: The framework of the proposed personalized dialogue generation model. The encoder and decoder share the same set of parameters. The dialogue context and the target persona are encoded independently using the encoder and their encodings are fed into the attention routing module in each decoder block. A dynamic weight predictor is trained to weigh the contribution of each route.

attention routing mechanism is proposed by extending the original multi-head attention module and introducing a dynamic weight predictor. Outputs of these attention routes are merged using the dynamically predicted weight.

Encoding with Personas

We introduce attribute embeddings in the encoder to model each persona $T_i, (i = 1, 2, ..., n)$ that is involved in the dialogue context C. Specifically, we first concatenate all the utterances in C with a special token "_SPE" and map each attribute t_j in T_i to an embedding representation. The input embedding for the Transformer encoder at each time step is constructed by adding the word embedding, positional embedding and attribute embeddings together (Figure 3). The proposed attribute embeddings enhance the dialogue context encoding E_C since the persona of every speaker is modeled in E_C. This differs from the previous work of **?** (**?**), where only word embeddings are used.

More precisely, three attributes are modeled in this study: Gender, Location, and Interest Tags. The embedding of Gender and Location can be obtained simply utilizing look-up tables since these attributes only have one unique value for each speaker, while the embedding of Interest Tags is computed as the average of all the tag embeddings for a speaker since each speaker may have several different interest tags.

Moreover, for the target persona T that the generated response should be coherent to, we pack all the key-value pairs in T into a word sequence and feed the corresponding word embeddings to the encoder to obtain E_T.

Attention Routing

In order to effectively utilize the persona-sparse dialogue data in a unified manner, it is expected to involve little or no persona features in the decoding process when training on
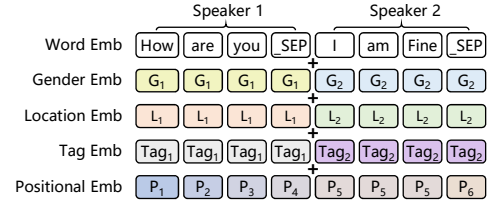

Figure 3: Input representation of the dialogue context. The input embedding for each token is the sum of a word embedding, a positional embedding, and attribute embeddings. Three kinds of attribute embeddings are modeled, i.e., gender embedding, location embedding, and tag embedding. The tag embedding of a speaker is calculated as the average of all the tag representations since each speaker may have several interest tags.

non-persona-related dialogues, whereas to incorporate abundant persona features when modeling persona-related dialogues. In this study, we devise an attention routing mechanism in the decoder to control the contribution of the target persona E_T in the decoding process.

Specifically, the vanilla multi-head attention module in the original Transformer block is extended to model the encodings of the target persona E_T, the dialogue context E_C and previously decoded tokens E_{prev}, respectively. We name each set of the multi-head attention operation as an *attention route* since it routes to different input features. More specifically, The three attention routes in our study use features extracted from previously decoded tokens E_{prev} as the

query, and attend to E_T, E_C and E_{prev}, respectively, i.e.,

$$O_T = \text{MultiHead}(E_{prev}, E_T, E_T) \qquad (2)$$
$$O_C = \text{MultiHead}(E_{prev}, E_C, E_C) \qquad (3)$$
$$O_{prev} = \text{MultiHead}(E_{prev}, E_{prev}, E_{prev}) \qquad (4)$$

Here, we use unmasked bi-directional self-attention in Eq.2 and 3 to facilitate more efficient interactions, and use masked self-attention in Eq. 4 to avoid feeding the "golden truth" token.

The outputs of each attention route O_T, O_C and O_{prev} are averaged using a persona weight $\alpha \in [0, 1]$:

$$O_{merge} = \alpha O_T + (1-\alpha)O_C + O_C + O_{prev} \qquad (5)$$

where a large α value indicates that more persona information will flow to the final outputs, and thus the generated responses are expected to exhibit more persona-related features. Note that Eq. 5 ensures that features extracted from the dialogue context O_C and previous decoded tokens O_{prev} can always be incorporated in the decoder.

Ideally, the value of α should be annotated based on whether the training dialogue is persona-related or not. However, this would be impractical for a large scale dialogue dataset. In this study, we design a dynamic weight predictor to calculate α automatically in the training stage. Specifically, the predictor is modeled as a binary classifier $P_\theta(r|E_c)$ that takes as input the dialogue context encoding E_C and predicts whether the training dialogue is persona related ($r = 1$) or not ($r = 0$). The confidence of this binary classifier is used as the predicted weight:

$$\alpha = P_\theta(r = 1|E_C) \qquad (6)$$

More precisely, we model the weight predictor using a neural network that is parameterized by θ, and develop a heuristic script to produce labels for the training dialogue to optimize θ. This script applies manually crafted rules such as word matching to decide whether a given dialogue is persona-related or not. The weight predictor is jointly trained with the dialogue model by optimizing the following cross entropy loss on these script-generated noisy labels:

$$L_W(\theta) = -\sum_i r_i \log P_\theta(r_i|E_C) + (1-r_i)\log[1-P_\theta(r_i|E_C)]$$

Note that we can also directly use the heuristic script as the weight predictor, i.e., set α to either 1 (the input dialogue is persona-related) or 0 (otherwise) in the training process. However, these hard decisions may bring bias introduced by the script to our model and thus lead to sub-optimal results. On the contrary, our neural-based predictor models a soft approximation of the prior knowledge provided by the heuristic script, and the joint training approach also guide the encoder to capture more persona related features in the context encoding E_C. Our experiments also verify that the soft weights produced by our predictor lead to better results compared to the hard weights produced by the heuristic script.

Pre-training and Fine-tuning

A pre-training process is employed in this study. Specifically, we collect a large set of text data and follow the

Table 1: Statistics of the dialogue dataset used in this study.

Total number of dialogues	5.44 M
Total number of speakers	1.31 M
Total number of utterances	14.40 M
Dialogues with more than 4 utterances	0.81 M
Average utterances per dialogue	2.65
Average tokens per utterance	9.46

scheme introduced by (**?**) to train a language model by optimizing the standard maximum log likelihood loss:

$$L_{LM}(\phi) = -\sum_i \log P_\phi(u_i|u_{i-k}, \ldots, u_{i-1}) \qquad (7)$$

where ϕ is the parameter set of the language model, k is the size of the context window, and $u_{i-k}, \ldots, u_{i-1}, u_i$ is a sequence of tokens that is sampled from the training corpus.

Once pretrained, the parameter set ϕ is used to initialize the encoder and decoder of our model, and a fine-tuning process is employed to adapt ϕ to the dialogue dataset. We optimize the following loss for the dialogue generation task:

$$L_D(\phi) = -\sum_i \log P_\phi(u_i|u_{i-k}, \ldots, u_{i-1}, E_C, E_T) \quad (8)$$

where E_C and E_T is the dialogue context and target persona encoding, respectively, and $u_{i-k}, \ldots, u_{i-1}, u_i$ is a sequence of tokens from the dialogue response.

Further, in order to bridge the gap between the data used in the pre-training and the fine-tuning stage, we also optimize the language model loss (i.e., Eq. 7) that is evaluated using utterances sampled from the dialogue contexts in the fine-tuning process. This is in line with the prior work (**?**), in which performance improvements are observed when incorporating such an auxiliary loss. Specifically, $L_{LM}(\phi)$ is optimized in the pre-training stage and the following loss is optimized in the fine-tuning stage:

$$L(\phi, \theta) = L_D(\phi) + \lambda_1 L_{LM}(\phi) + \lambda_2 L_W(\theta) \qquad (9)$$

where λ_1 and λ_2 are hyper-parameters to balance each loss.

Dataset

The dialogue data used in this study were sampled from the PersonalDialog dataset (**?**), which were collected from a Chinese social media Weibo. Each dialogue in this dataset is composed of a Weibo post and its following replies. A structured personal profile of each speaker was also provided in this dataset, and three persona attributes (i.e., "Gender", "Location" and "Interest Tags") were approached in our study. Figure 1 shows a sampled dialogue and Table 1 shows a basic statistic of the data used in this study. About 0.88M dialogues are labeled as persona-related by our heuristic script.

We randomly sampled 10K sessions of dialogues as the validation set, and constructed two test sets, i.e., a random test set and a biased test set, to test the behavior of our model in different contexts. Specifically, the random test set contained 10K sessions of dialogues that were randomly sampled. Most of these dialogues did not contain persona-related

features since common Weibo users are not required to exhibit their personas intentionally on Weibo. Therefore, the contexts provided by the random test set are persona-sparse.

The biased test set was deliberately chosen to provide us different contexts under which speakers tend to reveal their personas. For example, the dialogue "Are you a boy or a girl?" and "I am a girl" is biased since the speaker reveals her gender in response to the gender-related post. We manually labeled 521 biased dialogues in this study. The contexts provided by the biased test set are persona-dense. It will be interesting to see if our model can produce more persona consistent responses under these biased contexts.

Experiments

Persona Classifier

In order to better evaluate whether the generated dialogue responses carry rich persona-related features, we built a binary classifier that takes as input a dialogue response R and a persona T in the form of a concatenated token sequence, and predicts whether T is exhibited in R. Specifically, we randomly sampled a batch of response-persona pairs, and manually labeled 1,044 positive pairs such that the persona is exhibited in the response. We also labeled the same number of negative pairs such that the response do not carry persona related features. We split these data into the train, validation, and test set with the ratio of 8:1:1 and fine-tuned a classifier based on the BERT-base model (**?**). The accuracy of this classifier on the test set achieved 75.5%. Similar approach was also used by **?** (**?**) and **?** (**?**).

Implementation Details

Our pre-training stage used a dataset that was collected from a set of Chinese novels, which covered a variety of genres (including Comedy, Romance, Mystery). The final pretraining corpus contains about 0.5 billion tokens and we trained a character-level language model with a vocabulary size of 13,084. The encoder and decoder contained 12 Transformer blocks, and 12 attention heads were used. Token embeddings had the size of 768 and the context window was of size 512. The dynamic weight predictor was implemented as a multi-layer perceptron after an average pooling layer on E_C. The value of λ_1 and λ_2 in Eq. 9 was set to 0.2 and 0.5, respectively. The pretraining stage lasted for 70 epochs, and we fine-tuned our model for another 30 epochs.

Baselines

We chose several baselines:

- **Att+PAB**: A RNN based model that encodes the input persona into a representation vector using a persona fusion module, and decodes personalized responses utilizing a persona-aware bias (**?**).

- **Trans.**: A Transformer model (**?**) that takes concatenated dialogue histories as input and produces the corresponding dialogue response without using persona-related features.

- **TTransfo**: The TransferTransfo model as introduced by **?** (**?**). This model fine-tunes the pre-trained model directly on the persona-sparse dialogues.

Table 2: Automatic evaluation on the random test set.

Model	Acc.	BLEU	F1	Dist.	ppl.
Att+PAB	13.99	1.61	8.60	0.130	69.30
Trans.	7.80	3.97	12.51	0.132	43.12
TTransfo	8.80	4.06	12.63	0.169	**32.12**
TTransfo+P	43.05	3.44	11.28	0.158	43.78
LConv	9.45	4.19	12.99	0.157	32.64
LConv+P	48.00	3.56	11.46	0.136	42.00
Ours	32.80	4.18	12.52	**0.171**	35.06
Ours, $\alpha=1$	**84.55**	3.45	10.96	0.154	38.56
Ours, $\alpha=0$	12.90	**4.56**	**13.02**	**0.171**	33.71
w/o PreT	27.10	3.86	11.62	0.146	48.48
w/o AEmb	31.85	4.15	12.56	0.164	35.75
w/o DWP	30.70	4.15	12.34	0.169	34.10
+ HW	32.55	3.50	11.90	0.151	38.52

- **TTransfo + P**: The same as the TransferTransfo model but this model is fine-tuned using only dialogues that are labeled as persona-related by our heuristic script, i.e., noisy persona dense dialogue data.

- **LConv**: The *multi-input* model [1] proposed in (**?**). This model uses a pre-trained encoder and decoder and is fine-tuned directly on the persona-sparse dialogues without using the dynamic weight predictor.

- **LConv + P**: The same as the *multi-input* model but it is fine-tuned using only dialogues that are labeled as persona-related by our heuristic script.

All the baselines that utilize the Transformer architecture used the same set of hyper-parameters, and the same pre-training approach is employed.

Further, we performed several ablation tests with our model to validate the effectiveness of each component. Specifically, the following variants were tested: (1) without the pre-training process (**w/o PreT**); (2) without the attribute embedding in the encoder (**w/o AEmb**); (3) without the dynamic weight predictor (**w/o DWP**), i.e., λ_2 in Eq. 9 was set to 0 and the outputs from all the attention routes were equally averaged. In order to further demonstrate the effectiveness of the proposed dynamic weighting scheme, we also tested the performance of our model using heuristic weights (**+ HW**), i.e., λ_2 in Eq. 9 was set to 0 and the weight α in Eq. 5 was set to either 1 or 0 based on the results of the heuristic script during the training.

Moreover, we also tried to generate different responses by setting the weight α in Eq. 5 to different values in the inference stage. Specifically, we set α to 0 (no persona), 1 (full persona), and the value predicted by the dynamic weight predictor, respectively.

Automatic Evaluation

Metrics We evaluated the models with the following automatic metrics: (1) *Persona Accuracy* (**Acc.**) was computed

---

[1]This model was proposed by the winning team "Lost in Conversation" in the ConvAI2 competition (**?**).

Table 3: Automatic evaluation on the biased test set.

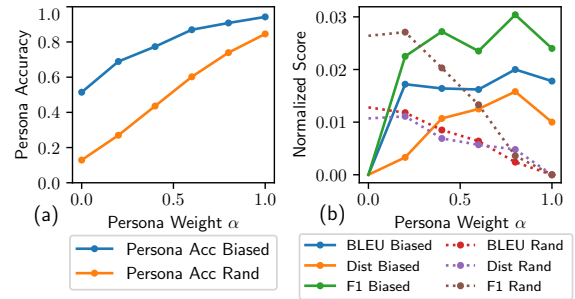| Model	Acc.	BLEU	F1	Dist.	ppl.
Att. + PAB	47.60	3.08	12.50	0.133	94.38
Trans.	34.93	7.06	15.38	0.203	85.80
TTransfo	45.87	8.68	17.39	0.260	**34.83**
TTransfo+P	61.61	9.10	18.41	0.257	38.07
LConv	44.34	8.47	17.08	0.238	37.44
LConv+P	59.88	9.82	18.91	0.231	41.68
Ours	92.13	10.53	19.47	0.256	38.68
Ours, $\alpha=1$	**94.24**	**11.63**	**20.51**	**0.262**	39.74
Ours, $\alpha=0$	51.44	9.00	17.44	0.249	40.89
w/o PreT	71.74	9.36	18.29	0.222	95.00
w/o AEmb	73.51	10.51	19.41	0.247	39.36
w/o DWP	73.90	10.61	19.26	0.256	37.08
+ HW	69.87	9.01	19.81	0.232	36.37

(a)   (b)

Figure 4: Effects for adjusting the persona weight α in the decoding process. Scores shown on the y-axis of (b) are normalized by subtracting the minimum scores.

by feeding the generated responses into the persona classifier together with the target persona, and obtaining the classification accuracy. Higher accuracy values mean the generated responses are more persona consistent. Similar metrics were also used by **?** (**?**) and **?** (**?**). (2) **BLEU** (**?**) was used to evaluate how many n-grams (n=1,2) in the generated responses overlap with those in the reference responses. (3) **F1** (**?**) was calculated based on the character level precision and recall of the generated responses. (4) *Distinct* (**Dist.**) (**?**) was used to measure the proportion of unique n-gram in the generated responses (n=1,2). (5) *Perplexity* (**ppl.**) was used to measure how the model fits the test data.

Results The performance on the random and biased test set is shown in Table 2 and Table 3, respectively. It can be seen that our model outperforms all the baselines on all the metrics except for the perplexity. This indicates that our proposed model can produce diversified dialogue responses carrying rich persona-related features. We can further observe that: **1)** Generating personalized dialogue responses hurts perplexity scores since persona-related words are relative rare and more biased generation of such words will lead to higher perplexity. Though the baseline models fit the test data well (with lower perplexity), they fail to produce more persona-related responses (with lower persona accuracy scores) compared to our model. This observation is in line with the results reported in the ConvAI2 competition (**?**); **2)** Ablation tests show that the pre-training stage significantly boost the model performance, and the proposed attribute embedding and dynamic weight predictor also help to improve the performance of our model; **3)** The weight α in Eq. 5 can be used to control the amount of persona-related features to exhibit in the decoding process. Higher α values lead to more persona-consistent responses; **4)** Larger performance gaps between our model and the baselines are obtained on the biased test set. This shows that our model can generate more persona-consistent responses in biased contexts.

Moreover, the effect of the persona weight α on the generated responses was further evaluated. Specifically, we com-

puted the scores of persona accuracy, BLEU, F1, and distinct corresponding to different α values (Figure 4). It is interesting to observe that: **1)** The persona accuracy increases rapidly with α (Figure 4a). This shows that more persona-related features will be incorporated in the decoded responses when α is larger. **2)** The scores for BLEU, F1 and distinct on the random test set decrease when α increases (dashed lines in Figure 4b). This is because the dialogues in the random test set are persona-sparse and less overlap between model-produced and human-generated responses will be observed if more persona-related features are incorporated. **3)** A clear increasing trend for BLEU, F1 and distinct is observed on the biased test set, but a performance drop is observed when α reaches 1 (solid lines in Figure 4b). This indicates that generating more persona-related responses lead to better performance on the persona-dense contexts, but merely pursuing persona consistency may hurt the performance on other dimensions. This is in line with the manual evaluation results shown in Table 4.

Manual Evaluation

Metrics For a given dialogue context and a target persona, we generated responses using all the transformer-based baselines and our model. Three individual annotators were employed to rate the model-generated responses together with the human-generated responses (Gold Resp) from three aspects: 1) *Utterance Fluency*: whether the responses are fluent and could plausibly have been produced by a human; 2) *Persona Consistency*: whether the responses are consistent with the target persona; 3) *Context Coherency*: whether the responses are coherent with the dialogue context. The rating scale of each measure is of (0, 1, 2), in which rating 0 means worst and 2 best.

Results 200 dialogue sessions were sampled from each of these two test sets, respectively, and 3.2K responses were generated. The inter-rater annotation agreement was measured using the Fleiss's kappa κ (**?**). Particularly, the κ value for *Utterance Fluency*, *Persona Consistency*, and *Context Coherency* was 0.81, 0.70, 0.52, respectively on the random test set, and 0.82, 0.73, 0.49, respectively on the biased test set. This indicates substantial annotation agreement for fluency and persona consistency, and moderate agreement for

Table 4: Manual evaluation on the random and biased test sets.

Model	Utterance Fluency		Persona Consistency		Context Coherency	
	Rand	Biased	Rand	Biased	Rand	Biased
Trans.	1.852	1.810†	0.997†	1.068†	1.428†	1.500
TTransfo	1.832†	1.890	1.015†	1.100†	1.498	1.517
TTransfo+P	1.802†	1.837†	1.125†	1.195†	1.217†	1.483†
LConv	1.863	1.882	1.028†	1.147†	1.490	1.550
LConv+P	1.832†	1.875†	1.093†	1.173†	1.238†	1.478†
Ours	1.837†	**1.912**	1.092†	1.198†	1.487	**1.563**
Ours, α=1	1.835†	1.900	**1.248**	**1.268**	1.303†	1.467†
Ours, α=0	**1.890**	1.880†	0.997†	1.085†	**1.535**	1.463†
Gold Resp	1.928	1.922	1.015	1.423	1.758	1.807

† significant difference with the best result (t-test, p-value<0.05)

context coherency.

Table 4 shows the manual evaluation results. Our model outperforms all the baselines in all the measures. Particularly for persona consistency, our full persona model (i.e., α=1) significantly outperforms all the baselines with a large margin. This indicates that our model can generate more persona-consistent responses that are fluent and context-coherent. Further observations also show that: **1)** Exhibiting too many persona-related features (i.e., obtaining higher persona consistency) hurts response fluency and context coherency. This is in line with the trade-off between the persona accuracy and perplexity as observed in the automatic evaluation results. Moreover, our dynamic weight predictor provides a better balance between the persona-consistency and context coherency, especially on the biased test set; **2)** The persona consistency of our full persona model (i.e., α=1) even surpasses the human-generated response on the random test set. This further proves that our model can incorporate richer persona-related features in the generated responses. **3)** Although directly fine-tuning on the noisy persona dense data (i.e., TTransfo+P and LConv+P) helps to produce more persona-consistent responses, our model still surpasses these baselines significantly. This verifies the effects of the proposed dynamic weighting scheme. This observation is also in line with the automatic evaluation results shown in Table 2 and 3.

Case Study

Figure 5 shows a sampled case, in which our model can generate coherent responses that reveal rich persona features, while responses produced by the baselines either do not exhibit persona-related features or are not grammatically fluent. This case also shows that the persona weight α can be effectively used to control whether to exhibit persona-related features or not. Specifically, our model with the full persona ($\alpha = 1$) reveals the location attribute in the response while our model without persona ($\alpha = 0$) does not exhibit persona related features. See the supplementary file for more cases.

Dialogue Context:	I am free this weekend. It that OK? (我这周末闲，好不) (Gender: Female, Location: Nanjing, Interest Tags: Cate)
Target Persona:	(Gender: Male, Location: **Nanjing**, Interest Tags: Null)
Trans. :	I'm also free, but I don't know what's going on. (我也闲，但是我不知道怎么回事)
TTransfo :	Great, come here and we can have a meal together. (好呀，你来了我们一起吃个饭呗)
TTransfo+P :	I am free this weekend. It that Okay? (我这周末闲，好不好？)
LConv :	It is not OK. (不好)
LConv+P :	Are you in **Nanjing** or **Nanjing**? (你在**南京**还是**南京**啊？)
Ours :	Ok, are you coming to **Nanjing**? (好啊，你来**南京**吗？)
Ours, $\alpha = 1$:	Come to **Nanjing**, I will treat you to dinner and film. (来**南京**，我请你吃饭看电影)
Ours, $\alpha = 0$:	Ok, where are you? I will go back tomorrow. (好啊，你在那呢？我明天就回去了)
Gold Resp :	OK, come to **Nanjing**! (好，来**南京**！)

Figure 5: Sample responses generated by baselines and our model.

Conclusion

In this paper, we present a pre-training based dialogue generation model that can produce coherent persona-consistent responses conditioned on explicitly represented personas. Our method can effectively utilize persona-sparse dialogue data in the fine-tuning stage. We add attribute embeddings in the encoder to model the persona of each speaker involved in the dialogue context and devise a dynamic weighting scheme in the decoder to balance the amount of persona-related features to exhibit in the decoded responses. Automatic and manual evaluation results show that our model can incorporate richer persona-related features in the generated responses compared to state-of-the-art baselines when the dialogues available at the fine-tuning stage are persona-sparse.

Rerum suscipit nisi pariatur repudiandae, exercitationem nihil asperiores amet non quidem deserunt, facere aut voluptate atque dicta nostrum error et saepe quaerat neque, pariatur laudantium quibusdam aspernatur laboriosam suscipit sed tempore, accusamus repudiandae soluta.Vero laboriosam voluptatem molestias iure explicabo, ullam nemo ea obcaecati dicta cumque dolorem pariatur itaque?Dolore harum magnam quibusdam voluptatibus dolorem culpa, commodi expedita obcaecati possimus facilis voluptates quasi unde ut?Modi provident odio laborum, officia consectetur laborum iure sint perferendis autem deleniti, accusantium omnis deleniti minima fugit eos non alias odit, quaerat itaque delectus temporibus accusantium eaque nesciunt possimus facilis cumque?Est perspiciatis corporis provident minus aut iure dolores alias, dolorum perspiciatis incidunt, provident aliquid reiciendis repellendus quam eos velit iste facilis voluptate, error ut voluptates

saepe?Dolorum quia odio velit quo at distinctio ad reprehenderit officia dolore omnis, doloribus hic voluptatum facere ex sequi consequuntur fugiat repudiandae?