Enhancing Personalized Dialogue Generation with Contrastive Latent Variables: Combining Sparse and Dense Persona

Yihong Tang[1], **Bo Wang**[2,*], **Miao Fang**[4],
Dongming Zhao[3], **Kun Huang**[3], **Ruifang He**[2], **Yuexian Hou**[2]
[1]School of New Media and Communication, Tianjin University, Tianjin, China
[2]College of Intelligence and Computing, Tianjin University, Tianjin, China
[3]AI Lab, China Mobile Communication Group Tianjin Co., Ltd.
[4]School of Computer and Communication Engineering,
Northeastern University at Qinhuangdao, Qinghuangdao, China
{toyhom, bo_wang}@tju.edu.cn

Abstract

The personalized dialogue explores the consistent relationship between dialogue generation and personality. Existing personalized dialogue agents model persona profiles from three resources: sparse or dense persona descriptions and dialogue histories. However, sparse structured persona attributes are explicit but uninformative, dense persona texts contain rich persona descriptions with much noise, and dialogue history query is both noisy and uninformative for persona modeling. In this work, we combine the advantages of the three resources to obtain a richer and more accurate persona. We design a **C**ontrastive **L**atent **V**ariable-based model (**CLV**) that clusters the dense persona descriptions into sparse categories, which are combined with the history query to generate personalized responses. Experimental results on Chinese and English datasets demonstrate our model's superiority in personalization.

1 Introduction

In order to develop personalized dialogue agents, current approaches enhance the personality of generated responses mainly utilizing three kinds of resources: (1) Defined sparse persona attributes (Zhang et al., 2018a; Song et al., 2019; Wolf et al., 2019; Liu et al., 2020; Song et al., 2021); (2) Dense persona description texts (Qian et al., 2018; Zheng et al., 2020; Song et al., 2021); (3) Historical queries of current dialogue (Li et al., 2016b; Ma et al., 2021). Each of the three resources has its advantages and disadvantages.

Sparse persona attributes (e.g., gender, age) are highly interpretable and have high information utilization, but the information is limited and cannot express complex persona features. Dense persona description text contains rich and flexible persona information but suffers from noisy expressions. Modeling personality directly from dialogue histories is free of additional persona information, but the persona information in history queries is both noisy and uninformative.

To address these issues, in this paper, we improve personalized dialogue generation by combining the advantages of the three resources. We design a contrastive latent variable (CLV)-based model that clusters the dense persona descriptions into sparse categories, which are combined with the history query to generate personalized responses. Specifically, first, the dialog's latest query and response together with dense persona description texts are encoded. Then the recognition distribution of query and response is jointly modeled with a pre-designed dual conditional variational autoencoder (CVAE (Sohn et al., 2015)). Simultaneously, the persona information is automatically separated into multiple parts to participate in the above process in parallel. These partitioned persona pieces of information are considered to hide different angles of portrayal. This process is also reinforced by contrastive learning. Next, a decider decides which category of persona information is used for persona modeling. Finally, a personalized generator combines the history query and additional persona information for response generation. Without explicit supervised signals, we design a pseudo-labeling and joint training method to train the decider.

Our contributions are summarized as follows:

(1) We design a framework named CLV based on contrastive latent variables to combine the advantages of three persona resources for personalized dialogue generation. The framework contains a self-separation algorithm and a decider, which are jointly trained to work in conjunction with each other. In this way, our work can both extract information more efficiently from the cluttered persona description text and not require persona informa-

*Corresponding author.

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics
Volume 1: Long Papers, pages 5456–5468

tion in the inference phase.

(2) Under the designed CLV-based framework, we propose a self-separation algorithm to mine and categorize dense persona description text into sparse persona profiles. Furthermore, a decider is proposed to decide whether the dialogue should involve persona information and choose appropriate persona profiles among the persona profiles generated by the self-separation algorithm. This process helps to improve the consistency of personalized dialogue generation.

(3) We conduct extensive experiments on the Chinese and English personalized dialogue datasets to demonstrate our model's superiority. We also propose a refined evaluation framework for personalized dialogue generation, which considers the consistency, coherence, and diversity of dialogue generation at the same time.

2 Related Work

Personalized Dialogue Generation Open-domain dialogue has been studied in depth for a long time (Koehn et al., 2003; Ni et al., 2021), and under the influence of the psychological theory, personality has been incorporated into the requirements for dialogue generation. Personalized dialogue generation has three typical approaches: (1) Using well-defined sparse persona attributes (e.g., gender, age), the model can utilize different attributes efficiently and interpretably, and knowledge-enhanced dialogue generation approaches can be borrowed (Zhang et al., 2018a; Song et al., 2019; Wolf et al., 2019; Liu et al., 2020; Bao et al., 2020; Song et al., 2021). However, sparse attributes can only provide little persona information without complex semantics. (2) Mining information from dense textual persona descriptions, which contain rich and deep persona information but are very noisy (Qian et al., 2018; Song et al., 2020; Zheng et al., 2020; Song et al., 2021). (3) Implicitly modeling persona profiles from historical dialogue query (Li et al., 2016b; Ma et al., 2021; Zhong et al., 2022). This approach does not rely on additional persona information, but it is difficult to acquire personality implicitly from dialogue history without reference objects.

Dialogue generation based on CVAE Besides personalization, another essential goal of personalized dialogue generation is the diversity of dialog expression. To this end, existing works have explored hidden variable models that model the variables in the dialogue process as Gaussian distributions, which can enhance the diversity of dialogue generation by introducing randomness (Zhao et al., 2017; Song et al., 2019; Hu et al., 2022). In this direction, one typical approach is to include persona information as a condition in regular Seq2Seq constructs and to model responses and queries as recognition distributions in CVAE (Li et al., 2018); another approach is to combine persona information or other external conditions and responses as generation targets before modeling joint distributions together with queries (Lee et al., 2021). In addition, many CVAE text generation models focus on other tasks, and they modify model details as well as probability maps for different tasks, which are not considered in this paper.

3 Methodology

3.1 Overview

Given multi-turn dialogue of two users u_i, u_j. The dialogue context of u_i is $U^i = \{(Q_1^i, R_1^i), \cdots, (Q_t^i, R_t^i)\}$. Q^i is the *query* initiated by u_j to u_i. The goal of the personalized dialogue is to generate a personalized *response* R_i using the corresponding *personal* information P_i in text form.

The overview of our model is shown in Figure 1. The overall model is composed of four modules: encoder, self-separation module, decider, and generator (marked in Figure 1 with orange borders). Specifically, the encoder module encodes dialogue queries, persona information, and responses respectively. The self-separation module separates the persona information in the hidden sentence vector space to form the grouping of persona information with implicit categories. We use multiple CVAEs to process the grouping persona information and get the grouping latent variables. The decider then automatically selects the latent variable to use from the group and feeds it into the generator along with the query. Finally, the generator autoregressively generates personalized responses based on the query and latent variables.

3.2 Encoder

we use a pre-trained GPT-2 (Radford et al., 2019) to encode the personal information text P_i, dialog query Q_i, and dialog response R_i. We take the hidden vector of the last time step in the last layer of GPT-2 as the representation of the whole

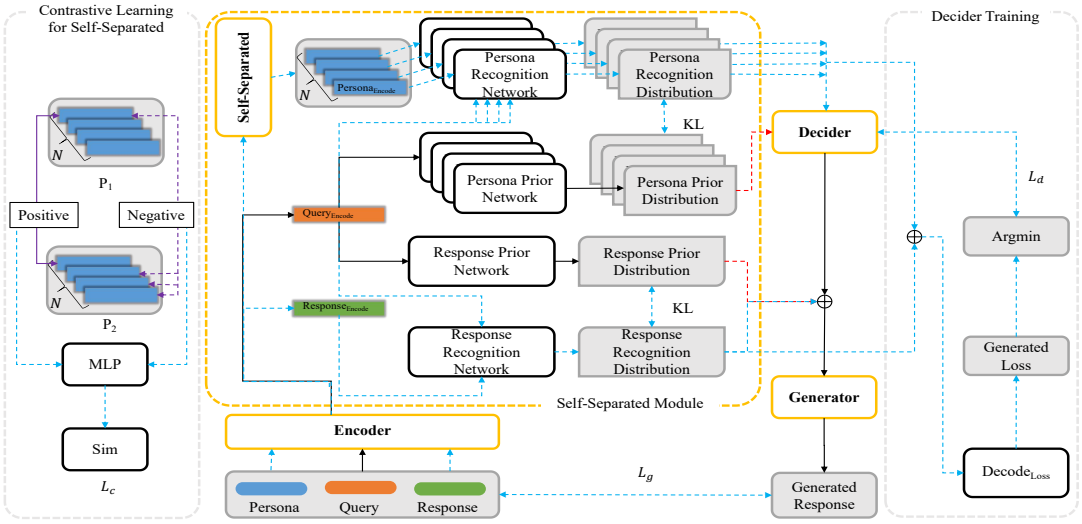

Figure 1: The overview structure of the proposed model. Connections with dashed blue lines only appear during training, connections with dashed red lines only appear during inference, and connections with solid black lines indicate that they appear during both training and inference phases. The purple lines represent positive and negative example constructions in contrastive learning.

paragraph:

$$p_i = \text{GPT-2}_{\text{Hidden}}(P_i), \quad (1)$$
$$q_i = \text{GPT-2}_{\text{Hidden}}(Q_i), \quad (2)$$
$$r_i = \text{GPT-2}_{\text{Hidden}}(R_i), \quad (3)$$

where $p_i, q_i, r_i \in \mathbb{R}^d$, and d is the dimension of the hidden state.

Algorithm 1: Persona Self-Separation

Input: $p \in \mathbb{R}^{1\times d}$: the vector representation of original sentence;
N : hyper-parameter, the self-separation coefficient;
d : the dimension of the hidden state;

Output: $P_g \in \mathbb{R}^{N\times d}$: vector representations of persona information after processing, in this context, it is the form of a set;

Initialize P_g;
Set $s \leftarrow$ the integer of d/N;
for $i = 1$ to N **do**
  Initialize augment vector $c_i \leftarrow (0, 0, \ldots, 0)_{1\times d}$;
  Set $c_i\,[(i-1)\times s + 1 : i \times s] \leftarrow (1, 1, \ldots, 1)_{1\times s}$;
  $P_g\,[i, :] \leftarrow \text{MLP}(p + c_i; c_i)$;
end for
return P_g

3.3 Self-Separated Module

After obtaining the hidden state representation of P, Q and R, their representation vectors are further processed. As mentioned above, sparse personal information is more explicit and interpretable, while dense information text contains rich information but needs to be more organized. Therefore, referring to the research of Sun et al. (2021), we propose a self-separation method of persona information, which implicitly divides dense text persona information into N categories:

$$P_g = \text{P-Sepa}(p), \quad (4)$$

where $P_g = \{p_1, p_2, \cdots, p_N\}$, and P_g represents the persona information after grouping, which is composed of multiple parallel persona information. For the algorithm of P-Sepa, see Algorithm 1.

In order to let the model automatically classify the grouped persona information, we use contrastive learning on the data in the same batch to let the model learn the similarities between the grouped persona information. Specifically, for two data points, P_g^i and P_g^j, we use a contrastive loss to help the model better represent group persona information.Following simcse, we denote $h_k^i = f_\theta(p_k^i)$ where $p_k^i \in P_g^i$. Then we get the training objective:

$$L_c = -log\frac{e^{sim(h_k^i, h_k^j)/\tau}}{\sum_{n=1}^{N} e^{sim(h_k^i, h_n^j)/\tau}}, \quad (5)$$

where τ is a temperature hyperparameter and $sim(h_k^i, h_k^j)$ is the cosine similarity.

The model samples the persona latent variable z_p from the persona distribution and the response latent variable z_r from the potential response distribution. Since z_p and z_r respectively represent different aspects of the generated responses (z_p contains the persona, and z_r captures the specific query-response association), we assume that z_p and z_r are independent of each other, namely $z_p \perp z_r$. So, the response generation process can be said to use the following conditional distribution $p(r, z_p, z_r|q) = p(r|q, z_p, z_r)p(z_p|q)p(z_r|q)$. Our goal is to use the deep learning method to approximate $p(r|q, z_p, z_r)$, $p(z_p|q)$ and $p(z_r|q)$, in which, according to Zhao et al. (2017) and Song et al. (2019), we refer to $p(r|q, z_p, z_r)$ as a response generator and $p_\theta(z_p|q)$, $p_\theta(z_r|q)$ as a *prior network*. In order to approximate the posterior distribution of the true, we refer to $q_\varphi(z_p|q, p)$ and $q_\varphi(z_r|q, r)$ as recognition networks.

We train this CVAE using Stochastic Gradient Variational Bayes(SGVB) (Kingma and Welling, 2013) by maximizing the *variational lower bound* of conditional log-likelihood. Following Zhao et al. (2017) and Song et al. (2019), we assume that potential variables z_p and z_r follows a multivariate Gaussian distribution with the diagonal covariance matrix. The lower bound of the variation of CLV-CVAE can be written as:

$$\begin{aligned} L_g = & E_{q_{\varphi_r}(z_r|q,r);q_{\varphi_r}(z_p|q,p)}(\log p(r|q,z)) \\ & - KL(p_{\theta_q}(z_p|q)||q_{\varphi_r}(z_p|q,p)) \\ & - KL(p_{\theta_r}(z_r|q)||q_{\varphi_r}(z_r|q,r)), \end{aligned} \tag{6}$$

Because we assume that the underlying variables z_p and z_r follow isotropic multivariate gaussian distribution, both recognition networks $q_{\varphi_p}(z_p|q, p) \sim \mathcal{N}(\mu_p, \sigma_p^2\mathbf{I})$ and $q_{\varphi_r}(z_r|q, r) \sim \mathcal{N}(\mu_r, \sigma_r^2\mathbf{I})$, both prior networks $p_{\theta_p}(z_p|q) \sim \mathcal{N}(\mu'_p, \sigma'^2_p\mathbf{I})$ and $p_{\theta_r}(z_r|q) \sim \mathcal{N}(\mu'_r, \sigma'^2_r\mathbf{I})$. In order to sample z_p and z_r from the prior network and recognition network in training and to make the sampling operation differentiable, using the *reparameterization* technique (Kingma and Welling, 2013), we have:

$$\begin{bmatrix} \mu_p \\ \sigma_p^2 \end{bmatrix} = W_q^{recog} \begin{bmatrix} q \\ p \end{bmatrix} + b_q^{recog}, \tag{7}$$

$$\begin{bmatrix} \mu_r \\ \sigma_r^2 \end{bmatrix} = W_r^{recog} \begin{bmatrix} q \\ r \end{bmatrix} + b_r^{recog}, \tag{8}$$

$$\begin{bmatrix} \mu'_p \\ \sigma'^2_p \end{bmatrix} = W_q^{prior} q + b_q^{prior}, \tag{9}$$

$$\begin{bmatrix} \mu'_r \\ \sigma'^2_r \end{bmatrix} = W_r^{prior} r + b_r^{prior}, \tag{10}$$

where p, r, q are the representation vectors obtained in Section 3.2.

Finally, z is fed into the generator to generate r together with the dialogue query q, where: $z = z_p + z_r$. How to get the final z_p is explained in detail in Section 3.4.

3.4 Decider

In fact, in order to make the model can find the appropriate persona information, we do not let CLV choose from the grouped persona information directly, but first, use the recognition network or prior network to obtain the grouped persona information latent variables $Z_p^g = \{z_p^1, z_p^2, \cdots, z_p^N\}$, which is obtained by sampling a set of distributions constructed separately for each vector in P_g. Then, the *Decider* is trained to choose between them. We call it the Decider because it also includes the decision not to use personal information.

Specifically, the decider is a classification neural network composed of multi-layer sensing units which use a soft decision method to make a selection. The decider-matrix is composed of classification probability, and the classification probability is multiplied by the grouping persona information latent variable to get the final persona information latent variable z_p. For grouped persona information latent variable Z_p^g:

$$W_d = \text{Softmax}(\text{MLP}([Z_p^g; q])), \tag{11}$$

$$z_p = W_d \cdot Z_p^g, \tag{12}$$

where $Z_p^g \in \mathbb{R}^{N\times d}$, $W_d \in \mathbb{R}^{1\times N}$ and $z_p \in \mathbb{R}^d$.

It is difficult to directly let the decider learn how to choose from the latent variables of grouping persona information generated by sampling the persona distribution of implicit clustering. Therefore, we introduce the pseudo-label to guide the learning of the decider. The more intuitive idea is that if a latent variable in the group of persona information latent variables can achieve a minor decoding loss in the generator, then it may be a better latent variable. Based on this idea, we designed the decision loss to train the decider:

$$y = \text{Argmin}(\text{GPT-2}_{\text{Loss}}(Z_p^g)), \tag{13}$$

$$L_d = -y \log(W_d), \tag{14}$$

where y is the index corresponding to z_p input into the generator to obtain the minimum decoding loss.

Dataset	# Train	# Valid	# Test
ConvAI2	43,410	4,213	2,138
Baidu PersonaChat	376,016	19,923	4,456

Table 1: Statistics of persona dialogue datasets.

3.5 Generator

We use a pre-trained GPT-2 as the generator, which uses the dialogue query as input and adds cross-attention to the latent variable z:

$$\hat{R} = \text{GPT-2}_{\text{Generator}}(Pre(z), q), \quad (15)$$

where $Pre(z)$ is the pre-cross attention object added before the standard GPT-2, which autoregressively generates a personalized response $\hat{R}$.

3.6 Training and Optimizer

In our practice, we find that there are some challenges in training the decider, which is probably the reason for the mutual influence between loss functions. Firstly, there will be conflicts between the KL divergence and the decoding loss of the generator. Secondly, the loss of the decider depends on the dummy label monitoring signal set by us. Finally, for the purpose of implicit clustering of persona information, the contrastive enhancement loss is largely independent of the above losses.

In order to promote gradient learning involving the above loss functions, a joint training process is designed to train CVAE and decider alternately. Specifically, in each training iteration, we first sample query Q, response R, and persona information P of two data points from batch data D, conduct contrastive training on encoders encoding persona information according to the self-separation algorithm 1, and then generate latent variables after self-separation respectively according to the method described in Section 3.4. The generator's loss value creates a dummy label y (Eq. 13), which is used to train the decider by optimizing the loss L_d(Eq. 14).

Further, we traverse D, generate a personalized response R, and update the generator and CVAE MLP by optimizing loss L_g (Eq. 6).

4 Experiments

4.1 Datasets

ConvAI2 (Dinan et al., 2019) is an English dataset containing rich personal information, and the dialogues in this dataset are based on the personal facts corresponding to the characters. It is derived from PersonaChat (Zhang et al., 2018b) and obtained after filtering and refinement. It is a crowdsourced dataset covering rich persona features, and we have processed it to remove some noise.

Baidu PersonaChat[1], which is a personalization dataset collected and open-sourced by Baidu, is similar to ConvAI2, although it's Chinese.

We summarize the key statistics of the two personalized dialogue datasets in Table 1. As mentioned earlier, we only use the persona information of the two datasets during training.

4.2 Baselines

We compare the proposed model with 6 baselines, which can be classified into 3 categories.

Non-Personalized Approaches **Seq2Seq** with Attention (Sutskever et al., 2014) is a sequence-to-sequence model with an attention mechanism (Luong et al., 2015). The pre-trained **GPT-2** (Radford et al., 2019) performs well in various text generation tasks and is used as a dialogue generation model after training on a dialogue corpus.

Approaches based on Dense Persona Information These methods use persona information to construct knowledge enhancement models, and for better model comparison, we tested these methods using the dialogue history as an approximation of the persona information. **PerCVAE** (Zhao et al., 2017) encodes the persona information text as a conditional representation and uses CVAE to generate personalized responses. **BoB** (Song et al., 2021) uses the Bert model for personalized dialogue generation and integrates the consistency generation task with the consistency inference task jointly to provide insight into the evaluation mechanism of personalized dialogue generation.

The Dialogue History-based Approach **DHAP** (Ma et al., 2021) uses historical memory to store and construct dynamic query-aware user profiles from dialogue histories and then uses a personalized decoder to generate responses. **MSP** (Zhong et al., 2022) enhances personalized dialogue generation by retrieving similar conversations from similar users via User Refiner and Topic Refiner and uses a Token Refiner to find the relevant tokens to be used during training, which is the best overall performance model for persona-free information personalized dialogue generation.

Implementation Details are in Appendix A.1.

[1] https://www.luge.ai/#/luge/dataDetail?id=38

		Coherence			Diversity				Consistency
		BLEU-1	ROUGE-L	Coh.Score	C-Dist-1	C-Dist-2	S-Dist-1	S-Dist-2	Coh-Con.Score
ConvAI2	Seq2Seq	3.45†	5.45†	34.85†	1.23†	3.84†	34.21†	61.59†	10.85†
	GPT-2	6.77†	10.96†	56.71†	7.35†	28.13†	68.22†	88.81†	13.29†
	PerCVAE	6.89†	10.54†	53.26†	**12.57**†	**39.54**†	67.48†	89.46†	12.95†
	BoB	7.85†	12.46†	62.47†	7.24†	26.41†	63.85†	85.02†	15.97†
	DHAP	7.21†	9.90†	64.27†	9.24†	30.98†	69.86†	90.23†	16.04†
	MSP	8.19†	11.67†	65.81†	10.49†	29.96†	65.79†	89.43†	15.45†
	CLV (Ours)	**11.85**	**15.10**	**71.72**	5.63	26.91	**71.24**	**92.89**	**23.01**
Baidu PersonaChat	Seq2Seq	7.14†	8.66†	40.39†	0.97†	5.19†	29.61†	76.65†	8.96†
	GPT-2	10.53†	11.29†	49.37†	5.64†	24.98†	51.93†	84.06†	12.14†
	PerCVAE	10.86†	10.44†	51.19†	**10.39**†	27.86†	58.24†	87.37†	11.33†
	BoB	14.26†	13.30†	58.13†	5.36†	27.45†	52.91†	82.93†	16.33†
	DHAP	12.96†	12.54†	55.21†	6.23†	25.37†	57.09†	85.44†	12.30†
	MSP	15.84†	14.06†	**61.52**†	5.37†	**28.41**†	54.06†	86.24†	14.37†
	CLV (Ours)	**24.77**	**22.33**	60.74	2.42	22.96	**60.27**	**88.15**	**18.15**

Table 2: Automatic evaluation on two datasets. The best results are in **bold**. "†" indicates that our model passed the t-test with p-value < 0.05.

4.3 Evaluations

In order to obtain accurate performance comparisons, we use both automatic and human evaluations.

Automatic Evaluation We divide the automatic evaluation methods into three categories in order to evaluate and model the diversity, consistency, and coherence of the generated dialogues.

(1) **Diversity** Distinct-1/2 (Li et al., 2016a) considers the number of single or double frames in the generated responses and is usually used to evaluate diversity. Most experiments do not specify the object of evaluation for Distinct-1/2, whether it is the whole corpus or multiple sentences, so we propose C-Dist-1/2(Corpus-Distinct-1/2) and S-Dist-1/2(Sentence-Distinct-1/2) according to the different objects of evaluation, the former evaluating the dialogue responses generated by the model on the whole test set, and the latter evaluating multiple responses (set to generate five responses in this paper). S-Dist-1/2 provides a better evaluation of whether the model can generate interesting responses in the same situation.

(2) **Consistency** The personalized dialogue generation task requires consistency between the generated responses and the persona information, and we propose Con.Score (Consistency Score) based on C.score (Madotto et al., 2019), which is obtained based on the referee model and can be defined as:

$$\text{Con.Score}(P,Q,R) = \begin{cases} 1, & \text{if NLI}(P,Q,R) = 1 \textit{ or } 2, \\ 0, & \text{if NLI}(P,Q,R) = 0. \end{cases} \tag{16}$$

where the NLI model is a triple classification model and can be found in Appendix A.

(3) **Coherence** BLEU-1 (Papineni et al., 2002) and ROUGE-L (Lin and Och, 2004) are classical words overlap-based metrics for measuring the similarity between generated responses and factual responses, which we believe can indirectly measure the coherence of dialogues. The reason we didn't look at BLEU-2/3/4 because we think that too much rigid coverage doesn't reflect the coherence of the model. And similar to the Con.Score, we propose the Coh-Con.Score (Coherence-Consistency Score), which is also obtained based on the NLI model:

$$\text{Coh-Con.Score}(P,Q,R) = \begin{cases} 0, & \text{if NLI}(P,Q,R) = 0, \\ 1, & \text{if NLI}(P,Q,R) = 2. \end{cases} \tag{17}$$

Human Evaluation Taking into account the uncertainty of the criteria when evaluating, we perform human evaluations of all models, and we convert the scoring method to a ranking method. Specifically, we extract 100 data points(queries, responses, and persona information) and hire three well-educated annotators to score the responses generated by the different models in a ranking style and to normalize them into specific scores on a scale of $[0, 1]$ at the end. We focus on four aspects: readability, diversity, consistency, and coherence, and ask the evaluators to rank eight options for the seven model-generated responses and the factual responses.

4.4 Experimental Results

Automatic Evaluation Table 2 shows the performance of all models on different automatic metrics for both Chinese and English datasets, and

Model	Readability	Diversity	Consistency	Coherence
Seq2Seq	0.57†	0.69†	0.11†	0.34†
GPT-2	0.73†	0.72†	0.43†	0.69†
PerCVAE	0.71†	0.82†	0.41†	0.65†
BoB	0.72†	0.80†	0.57†	0.73†
DHAP	0.77†	0.85	0.49†	0.69†
MSP	0.75†	0.83†	0.51†	0.72†
CLV (N=4)	**0.79**	**0.85**	**0.61**	**0.81**
Ground-Truth	0.80	0.91	0.86	0.97

Table 3: The result of human evaluation on ConvAI2 dataset. "†" indicates that our model passed the t-test with p-value < 0.05.

it can be clearly observed that our CLV model improves on key metrics and these improvements are statistically significant (t-test with p-value $<$ 0.05). Specifically, we can observe that: (1) **Diversity**. CLV shows different results on the two diversity evaluation dimensions. For S-Dist-1/2, CLV leads the other models, which indicates that our model is able to make more diverse and flexible responses compared to other models when facing the same situation. However, C-Dist-1/2 is lower than most models, which indicates that our model makes some sacrifices to improve consistency and coherence, and we will analyze this reason further in Section 5. (2) **Consistency**. The lead of the consistency personalization metric Con.Score implies that our approach can integrate persona information into the generation, especially when this integration is done without information generation, which is more indicative of the superiority of CLV. (3) **Coherence**. The performance of our model in coherence is also outstanding, whether it is the coverage index BLEU-1, Rouge-L, or the learning index Coh-Con.Score, which also shows that it is feasible to use the coverage index as a kind of evaluation basis for dialogue coherence. Our task diversity, coherence, and consistency can be used as three key bases for evaluating personalized dialogue generation, and the findings in the experiments suggest that our model is able to produce more personalized responses than all baselines.

Human Evaluation Human evaluation results on ConvAI2 are shown in Table 3. We calculated the Fleiss Kappa among the three annotators and obtained a Kappa of 0.67, which implies that the three annotators are in *substantial agreement* (Landis and Koch, 1977). In general, the results of human annotations are consistent with the results of automatic evaluations. They both demonstrate the advantages of our model in terms of personalized dialogue generation and basic readability.

5 Further Analysis

We further describe our model through a series of analyses. All analyses are based on the ConvAI2 dataset, and similar phenomena can be observed on Baidu PersonaChat.

Ablation Study To investigate the effects of different modules in CLV, we conducted an ablation study by removing modules. The results of the ablation study are shown in Table 5. We **first** investigated the impact of the core mechanism of the model, the self-separation algorithm. After removing the complete self-separation mechanism, the model degenerates to the most basic GPT-2 model, and it can be observed that the performance is on par with GPT-2. If we just remove the contrastive learning in the self-separation algorithm and keep the CVAE, we can see that the performance of the model also has a large decline, but the model's C-Dist-1/2 has an improvement, which is due to the global diversity due to the randomness of the sampled hidden variables in CVAE, which also indicates that CLV does sacrifice global diversity for other performance. **Then**, for the *decider*, we eliminate persona information by directly computing the mean of the grouped persona information latent variables, and we can find that the *decider* also plays an important role in CLV, especially when many dialogues are generated without considering persona, which shows that our decider can make decisions automatically. **Finally**, we conducted an experiment to validate our proposed joint training, and its performance degradation shows that it is difficult for the decider to learn how to make decisions without additional supervised signals.

Effect of Self-Separation Coefficients In CLV, the self-separation mechanism categorizes the persona information in an approximate implicit clustering way, and the self-separation coefficient N corresponds to the number of categories in the clusters. Intuitively, the self-separation factor will affect the model's performance, and we report this effect in Figure 2. The self-separation mechanism cannot do much good when the N is small. When N is set too large, the decider is also unable to make good decisions, which is due to the increased noise caused by too many categories, making the persona information too scattered, which is also consistent with the fact that the descriptive texts are always

Persona	I have a happy family. I sleep about six hours a day. I usually like to play the guitar. I work in a factory. I like listening to music.	I am a school teacher. I love to taste good food. I teach grade five in the Affiliated Primary School of Renmin University.	I'm afraid of heights. I don't believe in religion. I'm a lawyer.
Query	Yes, and what are your hobbies?	So you're from Beijing. I'm from Beijing, too.	A case? Are you a policeman?
Golden-Response	My life is relatively simple, in addition to making music, to help the factory. What do you like to do in your spare time?	Really? Are you from Beijing, too?	No, I'm a lawyer. I specialize in property disputes.
GPT-2	I like running and my body feels good.	Nice. I own a bakery.	No, I'm not. I'm a fireman.
MSP	I usually like to dance, I will dance with other anime anime anime anime voices.	Haha, I went to school in Beijing. I want to be a doctor in the future.	No, just talking and playing.
BoB	I love studying and I like music very much.	No, I grew up in Beijing.	No, I'm an accountant in this company.
CLV(N=4)	I like music very much and I like dancing. How about you?	Yes, I'm from Beijing. I'm a primary school teacher.	Yes, I am a special policeman, and often travel, because my work is very tiring, I am suitable for legal affairs.

Table 4: A case study. Keywords are marked in red.

Models	BLEU-1	C-Dist-1/2	Coh-Con.Score
MSP (N=4)	**11.85**	**5.63/26.91**	**23.01**
w/o Self-Separation	6.74	7.31/28.02	13.17
w/o Contrastive Learning	9.36	7.13/27.92	16.17
w/o Decider	10.01	4.99/24.89	17.59
w/o Joint Training	9.69	5.09/24.71	18.16

Table 5: Ablation experiments results on ConvAI2.

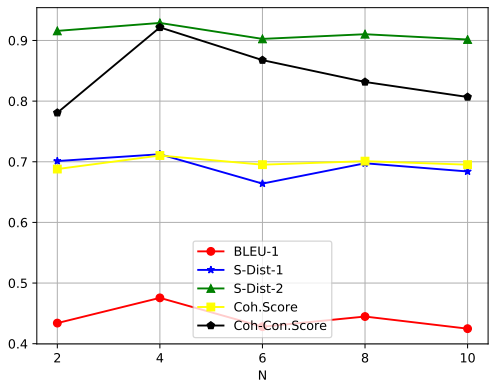

Figure 2: Experiments with the different N on the ConvAI2 dataset. For ease of viewing, BLEU-1 and Coh-Con.Score are multiplied by a factor of 4.

confined to several fixed perspectives.

To demonstrate the model's effectiveness more concretely, we conduct case studies. The results are shown in Table 4, which show that CLV can extract personal information, reconstruct persona profiles from queries alone, extract personal information, and generate fluent, personalized responses.

In Case 1, both CLV and BoB accurately answered "music" when asked about their hobbies, while CLV also used "How about you? " to keep the conversation going. In Case 2, CLV not only answered the address accurately but also flexibly used "school teacher" and "Affiliated Primary School of Renmin University" in the persona information to generate the response. In Case 3, all four models failed to accurately answer the question consistent with personality, but CLV still connected "lawyer" and "legal affairs".

By observing Cases 1 and 2, we can see that CLV can balance consistency and coherence, and its generation is consistent with persona and maintains context coherence. GPT-2 can only achieve basic sentence fluency. BoB and MSP can also generate good answers due to the help of context in reasoning. In Case 3, CLV creates a slightly fit answer, which is also better than the other models.

6 Conclusion

In this work, we propose a CLV model for personalized dialogue generation. Unlike existing works, we integrate the advantages of sparse and dense persona information. We use a *self-separation* mechanism to implicitly cluster the persona information in the dense persona information text so that the *decider* can consider different sparse categories of persona information during dialogue and enhance the personalization of dialogue generation. We also propose a more effective evaluation metric framework for personalized dialogue generation. The experimental results confirm the effectiveness of the model in generating personalized responses.

Limitations

First, our model is a method of approximating clustering by contrastive learning, but due to the limitations of the model structure, we cannot directly explore the performance of past clustering algorithms on this task. Secondly, due to the large scale of the experiment, our dialogue generator only considers GPT-2. Although the ablation study proves the effectiveness of our model, it is a limitation. Finally, this paper proposes a complete evaluation framework for personalized dialogue generation. It is very effective, but the specific indicators in it still need to be discussed and further studied. In addition, the model assumes that response and persona are independent Gaussian distributions in CVAE. Although it performs well in the experiment, it does not conform to realistic cognition.

Ethics Statement

From a general moral point of view, the generation of personalized dialogue in a broad sense may indeed cause problems such as identity forgery and the spread of false information. However, in this study, personalized corpus and responses are limited to the scope of experiments, which are not enough to threaten the real conversation.

Furthermore, all models in this paper are trained on public corpus. The used datasets do not contain unethical language. We also ensure the anonymization of the human evaluation.

Acknowledgements

This work was supported by National Natural Science Foundation of China(62272340, 61876128, 61876129, 62276187, 61976154, 61402323), State Key Laboratory of Communication Content Cognition(Grant No.A32003).

References

Siqi Bao, Huang He, Fan Wang, Hua Wu, and Haifeng Wang. 2020. PLATO: Pre-trained dialogue generation model with discrete latent variable. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 85–96, Online. Association for Computational Linguistics.

Emily Dinan, Varvara Logacheva, Valentin Malykh, Alexander Miller, Kurt Shuster, Jack Urbanek, Douwe Kiela, Arthur Szlam, Iulian Serban, Ryan Lowe, et al. 2019. The second conversational intelligence challenge (convai2). *arXiv preprint arXiv:1902.00098*.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2020. The curious case of neural text degeneration. In *International Conference on Learning Representations*.

Jinyi Hu, Xiaoyuan Yi, Wenhao Li, Maosong Sun, and Xing Xie. 2022. Fuse it more deeply! a variational transformer with layer-wise latent variable inference for text generation. *arXiv preprint arXiv:2207.06130*.

Diederik Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *International Conference on Learning Representations*.

Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Philipp Koehn, Franz J. Och, and Daniel Marcu. 2003. Statistical phrase-based translation. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pages 127–133.

J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174.

Jing Yang Lee, Kong Aik Lee, and Woon Seng Gan. 2021. Dlvgen: A dual latent variable approach to personalized dialogue generation. *arXiv preprint arXiv:2111.11363*.

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: Human Language Technologies*, pages 110–119, San Diego, California. Association for Computational Linguistics.

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 (Volume 1: Long Papers)*, pages 994–1003, Berlin, Germany. Association for Computational Linguistics.

Juntao Li, Yan Song, Haisong Zhang, Dongmin Chen, Shuming Shi, Dongyan Zhao, and Rui Yan. 2018. Generating classical Chinese poems via conditional variational autoencoder and adversarial training. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3890–3900, Brussels, Belgium. Association for Computational Linguistics.

Chin-Yew Lin and Franz Josef Och. 2004. Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, ACL '04, page 605–es, USA. 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.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Ilya Loshchilov and Frank Hutter. 2019. Decoupled weight decay regularization. In *International Conference on Learning Representations*.

Zhengyi Ma, Zhicheng Dou, Yutao Zhu, Hanxun Zhong, and Ji-Rong Wen. 2021. One chatbot per person: Creating personalized chatbots based on implicit user profiles. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '21, page 555–564, New York, NY, USA. Association for Computing Machinery.

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.

Jinjie Ni, Tom Young, Vlad Pandelea, Fuzhao Xue, Vinay Vishnumurthy Adiga, and E. Cambria. 2021. Recent advances in deep learning based dialogue systems: A systematic survey. *ArXiv*, abs/2105.04387.

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.

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 Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 4279–4285. International Joint Conferences on Artificial Intelligence Organization.

Yushan Qian, Bo Wang, Shangzhao Ma, Wu Bin, Shuo Zhang, Dongming Zhao, Kun Huang, and Yuexian Hou. 2023. Think twice: A human-like two-stage conversational agent for emotional response generation. *arXiv preprint arXiv:2301.04907*.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Kihyuk Sohn, Honglak Lee, and Xinchen Yan. 2015. Learning structured output representation using deep conditional generative models. *Advances in neural information processing systems*, 28.

Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang, and Ting Liu. 2021. BoB: BERT over BERT for training persona-based dialogue models from limited personalized data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 167–177, Online. Association for Computational Linguistics.

Haoyu Song, Yan Wang, Wei-Nan Zhang, Zhengyu Zhao, Ting Liu, and Xiaojiang Liu. 2020. Profile consistency identification for open-domain dialogue agents. In *EMNLP (1)*, pages 6651–6662.

Haoyu Song, Weinan Zhang, Yiming Cui, Dong Wang, and Ting Liu. 2019. Exploiting persona information for diverse generation of conversational responses. In *IJCAI*.

Bin Sun, Shaoxiong Feng, Yiwei Li, Jiamou Liu, and Kan Li. 2021. Generating relevant and coherent dialogue responses using self-separated conditional variational autoencoders. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5624–5637.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc.

Sean Welleck, Jason Weston, Arthur Szlam, and Kyunghyun Cho. 2019. Dialogue natural language inference. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3731–3741, Florence, Italy. Association for Computational Linguistics.

Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019. Transfertransfo: A transfer learning approach for neural network based conversational agents.

Zhitong Yang, Bo Wang, Jinfeng Zhou, Yue Tan, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. 2022. Topkg: Target-oriented dialog via global planning on knowledge graph. In *International Conference on Computational Linguistics*.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018a. 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.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018b. 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.

Tiancheng Zhao, Ran Zhao, and Maxine Eskenazi. 2017. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 654–664, Vancouver, Canada. Association for Computational Linguistics.

Yinhe Zheng, Rongsheng Zhang, Minlie Huang, and Mao Xiaoxi. 2020. A pre-training based personalized dialogue generation model with persona-sparse data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34:9693–9700.

Hanxun Zhong, Zhicheng Dou, Yutao Zhu, Hongjin Qian, and Ji-Rong Wen. 2022. Less is more: Learning to refine dialogue history for personalized dialogue generation. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5808–5820, Seattle, United States. Association for Computational Linguistics.

Jinfeng Zhou, Bo Wang, Ruifang He, and Yuexian Hou. 2021. Crfr: Improving conversational recommender systems via flexible fragments reasoning on knowledge graphs. In *Conference on Empirical Methods in Natural Language Processing*.

Jinfeng Zhou, Bo Wang, Minlie Huang, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. 2022a. Aligning recommendation and conversation via dual imitation. In *Conference on Empirical Methods in Natural Language Processing*.

Jinfeng Zhou, Bo Wang, Zhitong Yang, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. 2022b. Cr-gis: Improving conversational recommendation via goal-aware interest sequence modeling. In *International Conference on Computational Linguistics*.

A Appendix

A.1 Default Parameter Settings

Our experiments are done based on pre-trained GPT-2, and we tried various model structures and hyperparameters, and the final hyperparameters are as follows: the size of GPT-2 embedding and GPT-2 hidden vector is 768. All word embedding dimensions are set to 768, and we use word2vec to initialize word embedding. The number of layers of Transformer is 12. The self-separation coefficient N is set from 2 to 16(default is 4), the MLP input dimension and output dimension in the model are kept the same as the hidden vector, and the number of batches was set to 16. The maximum learning rate is 1e-4. The training of the proposed model was done on an Nvidia Telsa V100 16G GPU. The total training time takes approximately 10 hours. The temperature hyperparameter τ is 0.5. The pre-trained models used in these experiments of this paper include gpt2[2], gpt2-chinese-cluecorpussmall[3], xlm-roberta-base[4], and chinese-roberta-wwm-ext[5].

We use kernel sampling (Holtzman et al., 2020) as our decoding strategy, use the Adam (Kingma and Ba, 2014) optimizer to train the model and use AdamW (Loshchilov and Hutter, 2019) to warm up the generator. Please refer to the published project for additional details, which is publicly available[6].

A.2 NLI Model

NLI model is a triple classification model and can be design as:

$$\begin{aligned}&\text{NLI}(P, Q, R)\\&= \begin{cases} 2,\text{if } P \textit{ is consistent with } R \\ \quad \textit{and } Q \textit{ is coherent with } R\,, \\ 1,\text{if } P \textit{ is consistent with } R \\ \quad \textit{but } Q \textit{ is not coherent with } R \\ 0,\text{otherwise,} \end{cases}\end{aligned} \tag{18}$$

Here NLI (Welleck et al., 2019) is a pre-trained RoBERTa model (Liu et al., 2019), fine-tuned using a dataset constructed based on ConvAI2 and Baidu PersonaChat, and the test set accuracy of NLI model on Chinese and English is 83.2% and 83.1%, respectively.

[2] https://huggingface.co/gpt2

[3] https://huggingface.co/uer/gpt2-chinese-cluecorpussmall

[4] https://huggingface.co/xlm-roberta-base

[5] https://huggingface.co/hfl/chinese-roberta-wwm-ext

[6] https://github.com/Toyhom/CLV

ACL 2023 Responsible NLP Checklist

A For every submission:

☑ A1. Did you describe the limitations of your work?
limitations, 4.1 Datasets and 4.4 Evaluations.

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

☑ A3. Do the abstract and introduction summarize the paper's main claims?
Abstract, 1 Introduction.

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

B ☑ Did you use or create scientific artifacts?

4.1 Datasets, 4.4 Evaluations, A.1.

☑ B1. Did you cite the creators of artifacts you used?
4.1 Datasets, 4.4 Evaluations, A.1.

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

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

☒ B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?
In the original text of the dataset, the relevant data description has been included.

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

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

C ☑ Did you run computational experiments?

A.1 Default Parameter Settings, 4.5 Experimental Results.

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
A.1 Default Parameter Settings.

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

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
A.1 Default Parameter Settings.

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

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?
A.1 Default Parameter Settings.

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

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

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

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

☒ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
It will be mentioned later in the acknowledgments.

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