DLVGen: A Dual Latent Variable Approach to Personalized Dialogue Generation

Jing Yang Lee[1][a], Kong Aik Lee[2][b] and Woon Seng Gan[1][c]
[1]*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore*
[2]*Institute for Infocomm Research, A*Star, Singapore*
jingyang001@e.ntu.edu.sg, kong_aik_lee@i2r.a-star.edu.sg, ewsgan@ntu.edu.sg

Keywords: Personalized Dialogue, Natural Language Generation, Conversational AI, Latent Variables

Abstract: The generation of personalized dialogue is vital to natural and human-like conversation. Typically, personalized dialogue generation models involve conditioning the generated response on the dialogue history and a representation of the persona/personality of the interlocutor. As it is impractical to obtain the persona/personality representations for every interlocutor, recent works have explored the possibility of generating personalized dialogue by finetuning the model with dialogue examples corresponding to a given persona instead. However, in real-world implementations, a sufficient number of corresponding dialogue examples are also rarely available. Hence, in this paper, we propose a Dual Latent Variable Generator (DLVGen) capable of generating personalized dialogue in the absence of any persona/personality information or any corresponding dialogue examples. Unlike prior work, DLVGen models the latent distribution over potential responses as well as the latent distribution over the agent's potential persona. During inference, latent variables are sampled from both distributions and fed into the decoder. Empirical results show that DLVGen is capable of generating diverse responses which accurately incorporate the agent's persona.

1 INTRODUCTION

Personalized dialogue generation refers to the task of generating coherent, fluent dialogue consistent with a specific persona or personality. The generation of personalized dialogue is key to achieving natural, human-like dialogue. Conventionally, this task requires a representation of the interlocutor's persona or personality. This representation can either be explicitly provided in the form of a textual persona description, which consists of multiple persona statements, or via personality related metadata such as gender, age etc. Recent work have also explored inferring the persona statement directly from the dialogue context. The persona/personality representation, along with the dialogue context, is then used to condition the decoder.

Due to the difficulty of crafting persona descriptions and personality representations in a practical setting, the Persona Agnostic Meta-Learning (PAML) framework (Madotto et al., 2019) and the Multi-Task Meta-Learning (MTML) framework (Lee et al., 2021) were proposed. Both PAML and MTML involved pretraining the model via meta-learning and finetuning with dialogue examples corresponding to the dialogue agent's persona. Similarly, (Zheng et al., 2019b) also introduced a pretraining and finetuning approach for persona-sparse dialogue which features an attention routing structure and a learned persona attribute embedding. However, even though these methods do not require the explicit provision of any persona/personality information, multiple dialogue examples corresponding to the dialogue agent's persona is still needed to finetune the model. This is impractical as corresponding dialogue examples are rarely available in real-world implementations. Hence, in this paper, we propose a novel Dual Latent Variable Generator (DLVGen) for personalized dialogue generation, which does not rely on any persona/personality information or on any corresponding dialogue examples. Instead, DLVGen generates the response given only the dialogue context.

A common issue with regard to personalized dialogue agents is the low response diversity. Prior work address this issue by exploring the application of latent variable models, specifically the Condi-

[a] https://orcid.org/0000-0002-3611-3512
[b] https://orcid.org/0000-0001-9133-3000
[c] https://orcid.org/0000-0002-7143-1823

tional Variational Auto Encoder (CVAE) (Sohn et al., 2015). These approaches involved modelling the potential dialogue responses as a latent Gaussian distribution, which improves diversity due to the introduced stochasticity. Typically, the persona or personality information and dialogue history are used to generate the latent Gaussian distribution. During inference, the persona or personality vector, the dialogue context, along with a latent variable sampled from the latent distribution, are passed to the decoder for response generation. The Persona-CVAE framework (Song et al., 2019) aims to generate responses which incorporate information from the provided textual persona description, which is incorporated into the decoding process via a multi-hop attention mechanism. On the other hand, the Persona-Aware Variational Response Generator (PAGenerator) (Wu et al., 2020) relies on user embeddings trained concurrently with the CVAE. In addition to the CVAE, the basic Variational Auto Encoder (VAE) and the Wasserstein Auto Encoder (WAE) have also been applied to personalized dialogue generation in a similar manner (Chan et al., 2019). The Common Sense and Persona Aligned Chatbot (COMPAC) (Majumder et al., 2020), on the other hand, relies on a latent variable to select the appropriate persona information from a set of expanded persona descriptions.

Unlike the aforementioned models, DLVGen involves generating two latent Gaussian distributions: the latent distribution over the agent's potential persona (defined by multiple persona statements in the persona description), and the latent distribution over potential dialogue responses. While it has been established that modelling the latent distribution over responses would increase response diversity, we hypothesize that modelling latent distribution over the agent's potential persona would result in responses which incorporate a range of potential persona information inferred from the dialogue context. We utilize the persona description only during training of a neural network that is tasked with generating the latent distribution over the agent's potential persona. To the best of our knowledge, this is the first attempt at modelling the latent distribution over the agent's *potential* persona. For this paper, our contributions are threefold:

1. We propose the DLVGen framework which leverages persona descriptions only during training. Unlike prior frameworks, the proposed framework models both the latent distribution over potential dialogue responses as well as the latent distribution over the agent's potential persona.
2. We introduce a variance regularization technique which involves maximizing or minimizing the variance of the distribution over the agent's potential persona and the distribution over potential responses respectively. In particular, evaluation results reveal that minimizing the variance of the distribution over potential responses improves the persona consistency of the generated responses.
3. We present a selection framework based on lexical diversity to select the final response from a pool of generated responses. Generally, we find that the response selected via lexical diversity selection demonstrate greater persona consistency and diversity.

2 METHODOLOGY

2.1 Task Definition

In this paper, we will tackle the task of generating personalized dialogue responses in the absence of any persona/personality information or corresponding dialogue examples. Hence, during inference, the model is expected to generate personalized dialogue given only the dialogue context, which consists of all previous utterances in the conversation. In other words, the model is expected to infer the agent's persona from only the dialogue context. For this task, the generated response is expected to be diverse as well as persona consistent i.e., accurately reflect the agent's persona. However, in cases where no persona information can be inferred from the dialogue context, the responses generated should aim to be persona neutral. This means that, as far as possible, responses should not contain any personal information which could potentially contradict the agent's persona.

For our experiments, we utilize the ConvAI2 dialogue corpus. We chose the ConvAI2 corpus as the textual persona descriptions, which consists of multiple persona statements, can be used during training. Also, the available persona descriptions allows us to evaluate the amount of persona information that is accurately reflected in the generated response relatively easily.

2.2 Dual Latent Variable Generator

Essentially, the proposed Dual Latent Variable Generator (DLVGen) framework requires learning two distinct networks to model the latent distribution over potential dialogue responses as well as the latent distribution over the agent's potential persona. During inference, latent variables are sampled from both distributions and fed to the decoder, which consists of a

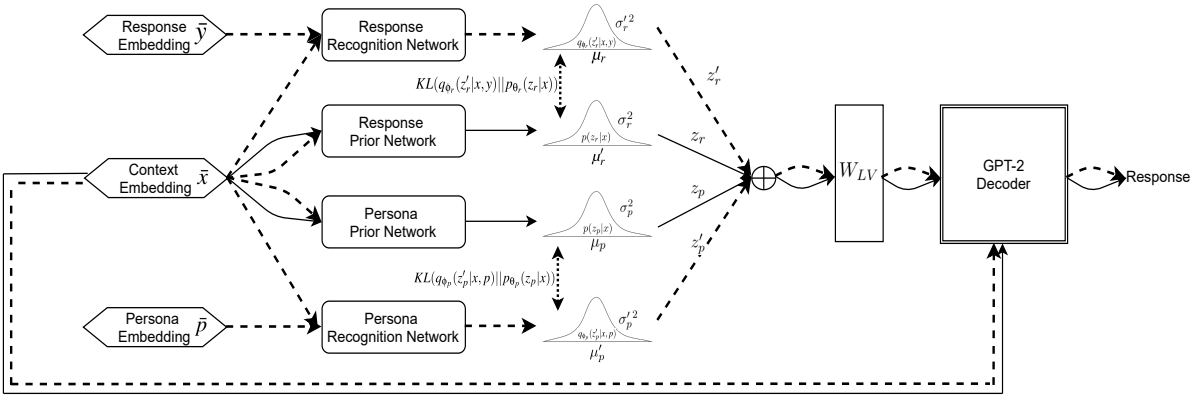

Figure 1: Flowchart depicting the architecture of the DLVGen model. The dashed lines and solid lines represents the connections that occur only during training and inference respectively. The dotted bidirectional connections indicate the KL divergence computation. $\oplus$ represents the concatenate operation.

GPT-2 pretrained language model (Radford and Wu, 2019). Randomly sampling the latent distribution over responses would improve response diversity due to the stochasticity introduced by the latent variable. Similarly, randomly sampling the latent distribution over the agent's persona would incorporate a wide range of potential persona information (inferred from the dialogue context) in the generated response. The induced stochasticity would also reduce the likelihood of the model incorporating the same persona information in multiple responses, further contributing to the overall response diversity. In our approach, the persona statements describing each interlocutor is only utilized during training. An overview of the model is provided in Figure 1.

In our discussion, x represents the dialogue context which consists of all the prior utterances in the dialogue history. p refers to all statements in the provided textual persona description. y denotes the dialogue response. During inference, the dialogue response is generated based on the dialogue context x, the latent variable sampled from the distribution over the agent's potential persona z_p, and the latent variable sampled from the distribution over potential dialogue responses z_r. Since z_p and z_r each represent different aspects of the generated response (z_p encompasses the persona information and z_r captures information regarding the flow of the dialogue), we will regard z_p and z_r as independent i.e., $z_p \perp\!\!\!\perp z_r$. The persona information and dialogue flow information, which are both derived from the dialogue context, are disentangled via the prior and recognition networks introduced in the remainder of this section. Hence, by assuming z_p and z_r are independent random variables, the generation process can be expressed via the following conditional distribution $p(y,z_p,z_r|x) = p(y|x,z_p,z_r)p(z_p|x)p(z_r|x)$. We also assume that z_p and z_r can be modelled by multivariate isotropic Gaussian distributions:

$$p(z_p|x) = N(\mu_p, \sigma_p^2 \mathbf{I}) \tag{1}$$

$$p(z_r|x) = N(\mu_r, \sigma_r^2 \mathbf{I}) \tag{2}$$

where μ_p, μ_r and σ_p^2, σ_r^2 refer to the mean and variance of the distribution over the potential persona and the distribution over potential responses respectively. Hence, to approximate $p(z_p|x)$ and $p(z_r|x)$, we define 2 prior networks $p_{\theta_p}(z_p|x)$ and $p_{\theta_r}(z_r|x)$, which are single-layer Multi-Layer Perceptrons (MLPs) parameterized by θ_p and θ_r. $p_{\theta_p}(z_p|x)$ and $p_{\theta_r}(z_r|x)$ are termed the persona prior network and the response prior network respectively. Before being fed to the prior networks, a GPT-2 pretrained model (Radford and Wu, 2019) is used to obtain $\bar{x}$, $\bar{y}$ and $\bar{p}$, the sequence embeddings of the dialogue context x, response label y and persona descriptions p respectively. Utilizing the persona and response prior networks, the following means μ_p, μ_r and variances σ_p^2, σ_r^2 are derived as follows:

$$\begin{bmatrix} \mu_p \\ log(\sigma_p^2) \end{bmatrix} = W_p \left[\bar{x}\right] + b_p \tag{3}$$

$$\begin{bmatrix} \mu_r \\ log(\sigma_r^2) \end{bmatrix} = W_r \left[\bar{x}\right] + b_r \tag{4}$$

where W_p, W_r refer to the weight vectors and b_p, b_r refer to the bias vectors corresponding to the persona and response prior networks.

Additionally, to approximate the posterior of $p(z_p|x)$ and $p(z_r|x)$, we introduce latent variables z'_p and z'_r as well as latent distributions $q(z'_p|x,p)$ and $q(z'_r|x,y)$, which are generated by 2 recognition networks $q_{\phi_p}(z'_p|x,p)$ and $q_{\phi_r}(z'_r|x,y)$ (defined as single-layer MLPs parametrized by ϕ_p and ϕ_r) respectively.

$q_{\phi_p}(z'_p|x,p)$ is termed the persona recognition network and $q_{\phi_r}(z'_r|x,y)$ is termed the response recognition network. Similarly, the sequence embeddings $\bar{x}$, $\bar{y}$ and $\bar{p}$ are fed to the recognition networks. The persona and response recognition networks define the approximate posterior $q(z'_p|x,p)$ and $q(z'_r|x,y)$ via generating the respective means μ'_p, μ'_r and variances ${\sigma'_p}^2, {\sigma'_r}^2$:

$$\begin{bmatrix} \mu'_p \\ log({\sigma'_p}^2) \end{bmatrix} = W'_p \begin{bmatrix} \bar{x} \\ \bar{p} \end{bmatrix} + b'_p \tag{5}$$

$$\begin{bmatrix} \mu'_r \\ log({\sigma'_r}^2) \end{bmatrix} = W'_r \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} + b'_r \tag{6}$$

where W'_p, W'_r refer to the weight vectors and b'_p, b'_r refer to the bias vectors corresponding to the persona and response recognition networks. The KL divergence between the prior and recognition networks will be included in the final loss function. It should be noted that for the persona prior network and the persona recognition network, before being fed into the GPT-2 pretrained model to obtain the embedding, the utterances corresponding to the opposing interlocutor (the user) are masked. This prevents the generated Gaussian distribution from modelling the persona of the opposing interlocutor.

To ensure end-to-end differentiability, latent variables are sampled using the reparameterization trick (Kingma and Welling, 2014). During training, the latent variables z'_p and z'_r are sampled from the persona and response recognition networks. However, during inference, the latent variables z_p and z_r are sampled from the persona and response prior networks. After sampling, the latent variables z'_p and z'_r (training) or z_p and z_r (inference) are concatenated and fed to a linear layer (W_{LV}) before being passed to the decoder, which constitutes a GPT-2 pretrained language model. The resultant representation is added to the GPT-2 input, which consists of the token embedding and the positional encoding, at every decoding step. The persona description p is only used during training to learn the latent distribution over the the agent's potential persona. A simple figure depicting this process in provided in Figure 2.

Since maximizing the conditional log likelihood during training requires an intractable marginalization over z_p and z_r, DLVGen is trained via the Stochastic Gradient Variational Bayes (SGVB) framework (Kingma and Welling, 2014) by maximizing the variational lower bound. Also, the Bag-of-Words (BoW) loss (Zhao et al., 2017) is incorporated into the loss function to address the vanishing latent variable problem. Essentially, SGVB involves maximizing the

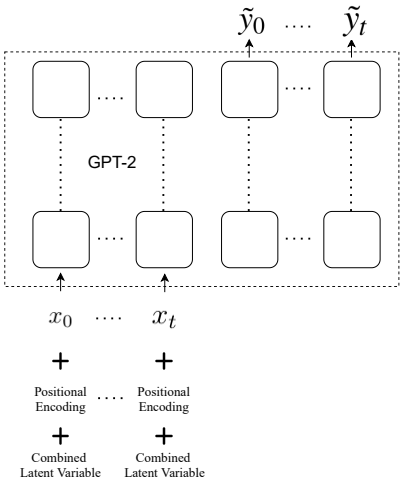

Figure 2: Diagram depicting the GPT-2 decoder during inference. After the latent variables are concatenated and fed into a linear layer W_{LV}, the resultant combined latent variable is added to the positional encoding and token embedding at every decoding step. x_0 and x_t represent the first and last token embedding of the dialogue context. y_0 and y_t represent the first and last token of the generated response.

variational lower bound of $p(y|x)$:

$$\begin{aligned} &L(\theta_p, \theta_r, \phi_p, \phi_r, x, y, p, z_p, z_r) = \\ &E_{q_{\phi_p}(z'_p|x,p); q_{\phi_r}(z'_r|x,y)}[log\, p(y|x, z_p, z_r)] \\ &- KL(q_{\phi_p}(z'_p|x,p)||p_{\theta_p}(z_p|x)) \\ &- KL(q_{\phi_r}(z'_r|x,y)||p_{\theta_r}(z_r|x)) \\ &+ E_{q_{\phi_p}(z'_p|x,p); q_{\phi_r}(z'_r|x,y)}[log\, p(y_{bow}|x, z_p, z_r)] \end{aligned} \tag{7}$$

where $KL(\cdot)$ refers to the KL divergence and y_{bow} represents the response bag-of-words.

2.3 Variance Regularization

Additionally, we introduce a novel regularization technique based on the variances of distributions generated by the prior networks i.e., $p(z_p|x)$ and $p(z_r|x)$. Essentially, we either maximize the variance of the latent distribution over the agent's potential persona σ_p^2, or minimize the variance of the latent distribution over potential responses σ_r^2. Specifically, this is realized by either adding the euclidean norm of the log-variance of $p(z_r|x)$, $log(\sigma_r^2)$, or the inverse of the euclidean norm of the log-variance of $p(z_p|x)$, $log(\sigma_p^2)$. The inverse of the euclidean norm of the log-variance can also be interpreted as the precision of the distribution. The variance regularization terms can be expressed as:

$$R_r = \lambda_r \, \| log(\sigma_r^2) \| \tag{8}$$

$$R_p = \lambda_p \| \frac{1}{log(\sigma_p^2)} \| \tag{9}$$

where R_r refers to the loss term which minimizes the variance of the distribution over potential responses, and R_p refers to the loss term which maximizes the variance of the distribution over the agent's potential persona. λ_r and λ_p are penalty terms to be tuned. Then, the 2 variance regularization terms can be individually added to the loss function. Alternatively, both terms can be added to the loss function to form the final loss function $\bar{L}$:

$$\begin{aligned}\bar{L}(\theta_p,\theta_r,\phi_p,\phi_r,x,y,p,z_p,z_r) = \\ L(\theta_p,\theta_r,\phi_p,\phi_r,x,y,p,z_p,z_r) \\ +R_r(\sigma_r^2)+R_p(\sigma_p^2)\end{aligned} \tag{10}$$

By maximizing the variance of the latent distribution over the agent's potential persona, we hope to encourage the persona prior network to model a wider range of potential persona information, which would further improve response diversity. On the other hand, by minimizing the variance of the latent distribution over potential responses, we aim to constrain the variability of the sampled random variables, improving the contextual coherence of the generated responses.

2.4 Lexical Diversity Selection

Lexical diversity is a measure of how many different lexical words appear in a given text. Unlike filler words such as 'and' or 'the', lexical words are words that convey information or meaning ('cat','computer','sky' etc.) . By sampling the Gaussian distributions N times, N different responses can be generated. Hence, we introduce a method of selecting the best response from the pool of N generated responses. Intuitively, we assume that lexically diverse sentences make for more informative responses. Hence, to quantify the lexical diversity of the generated responses, we utilize the Measure of Textual Lexical Diversity (MTLD) and the Moving Average Type Token Ratio (MATTR) scores (Fergadiotis et al., 2015). Obtaining the MTLD involves computing the Token-Type Ratio (TTR) for sequentially larger segments of the sentence. For each segment, when the TTR drops below a predefined threshold h, a count is incremented by 1. Then, the total number of tokens in the generated response is divided by the final count. The final MTLD is obtained by averaging the forward and backward MTLD scores. The MATTR, on the other hand is based on averaging the TTR of successive segments of the generated response with a fixed window size w. After the computing the MATTR and MTLD, we select the response with the highest combined score. This is summarized in the following expression:

$$\underset{r\in R}{\operatorname{argmax}}(0.1M_1^h(r)+M_2^w(r)) \tag{11}$$

where M_1^h and M_2^w refer to the functions for calculating the MTLD score (with threshold h) and MATTR (with window size w) score respectively. R refers to the pool of generated responses. The MTLD score was penalized by a factor of 0.1 to prevent it from overwhelming the MATTR score. The proposed lexical diversity selection method is relatively straightforward to implement and can be applied on top of any latent variable dialogue model.

3 EXPERIMENT

3.1 Corpus

We evaluate our approach on the ConvAI2 personalized dialogue corpus (Dinan et al., 2019). Since the ConvAI2 corpus is based on the PersonaChat corpus (Zhang et al., 2018), both corpora are structurally similar. Both corpora provide personalized dialogues as well as persona descriptions in the form of several sentences. However, the ConvAI2 corpus features numerous crowd-sourced rewrites and rephrases, which increases the overall task difficulty. In ConvAI2, the training set $\mathbf{D}_{train}$ contains 17,878 dialogues from 1155 unique personas, while the test set (validation set is used as the test set in our experiments as actual test set is hidden) $\mathbf{D}_{test}$ contains 1000 dialogues from 100 unique personas.

3.2 Implementation

For our experiments, we utilize the PyTorch, ParlAI, and HuggingFace libraries. During training, the Adam optimizer (learning rate = 0.0001) is used with a batch size of 16. The size of latent variables z_p, z_r, z_p', z_r' are fixed at 256. We utilize the GPT-2 pretrained language model (Radford and Wu, 2019) from HuggingFace to obtain the context, persona and response embeddings. The decoder also consists of a GPT-2 model (12 layers, 768 dimensional hidden state, 12 heads, 117 million parameters).

For variance regularization, the values of λ_r and λ_p was set to 0.5 and 1.0 respectively. For the MTLD and MATTR computation during lexical diversity selection, h and w were fixed at 0.72 and 4 respectively. Responses are generated via beam search with a beam size of 3. N is also set to 3.

3.3 Baselines

To evaluate our proposed model on the task described in section 2.1, we implement the following baselines:

1. **PAML** Following Madotto et al. (2019), we pretrain a standard transformer with PAML. However, using the ConvAI2 corpus, the model was applied to the task defined in section 2.1 instead (PAML was originally evaluated on the PersonaChat corpus), which involved directly evaluating the pretrained transformer without further finetuning.
2. **MTML** Following Lee et al. (2021), we pretrain a standard transformer with MTML ($\alpha = 0.8$). Similarly, using the ConvAI2 corpus, the model was applied to the task defined in section 2.1 (MTML was also originally evaluated on the PersonaChat corpus), which involved directly evaluating the pretrained transformer without further finetuning.
3. **GPT-2** Similar to TransferTransfo (Wolf et al., 2019). However, instead of GPT, GPT-2 was finetuned for dialogue generation. Since the model was applied to the task defined in section 2.1, the decoder input consists of only the dialogue context. The persona description was not utilized.
4. **CVAE** Similar to the CVAE-based dialogue model proposed by Zhao et al. (2017), which involves only generating and sampling from the latent distribution over responses. However, our implementation utilizes the GPT-2 pretrained model instead of the bidirectional GRU during decoding as well as to generate the sequence embeddings. Additionally, since the model was applied to the task defined in section 2.1, the persona description was not used. Instead, the prior network generates the latent Gaussian distribution based on only the dialogue context, and the decoder input consists of only the sampled latent variable and the dialogue context.
5. **DLVGen** A model which consists of our proposed DLVGen architecture with a GPT-2 decoder. We also implement additional variations of this model which include each of the variance regularization terms R_r and R_p, as well as the combination of both terms. Additionally, we implement the lexical diversity selection (LS) for all variants.

3.4 Evaluation

In our experiments, N was set to 3 i.e., the Gaussian distributions were sampled 3 times to generate 3 responses. For the latent variable models (CVAE and DLVGen), the final automatic and human evaluation scores are obtained by averaging the individual scores from the N responses generated from each dialogue example in $\mathbf{D}_{test}$.

3.4.1 Automatic Metrics

We evaluate the generated responses with the following automatic metrics:

1. **Distinct 1 & 2** Distinct-1 and 2 scores quantify the diversity of a generated response. The Distinct-1 and Distinct-2 scores accounts for the number of distinct 1-grams and 2-grams in the generated response respectively. A higher Distinct-1 or 2 score would indicate greater response diversity.
2. **C-score** The C-score (Madotto et al., 2019) reveals the extent to which the persona is accurately reflected in the generated response. Essentially, it is generated via a BERT model, which is finetuned to indicate if the generated response entails, contradicts or is independent to each statement in the persona description. A higher C-score would indicate a larger amount of accurately incorporated persona information.

3.4.2 Human Evaluation

We engaged 5 graduates to evaluate the various generated responses based on 3 criteria which are similar to the criteria used during human evaluation in prior work(Lee et al., 2021; Madotto et al., 2019):

1. **Persona Consistency** The individuals were told to rate the responses in terms of the amount of persona information present in the generated response on a scale from -1 to 1. A score of -1 would indicate a contradiction with the corresponding persona description and a score of 1 would indicate an accurate incorporation of persona information.
2. **Naturalness** The individuals were told to rate the generated responses according to human-likeness on a scale from -1 to 1. This criteria accounts for both the diversity and general fluency of the generated response. A score of 1 would be assigned to a perfectly fluent response that is indiscernible from a human response. A score of -1 would indicate an awkwardly phrased response with multiple grammatical and lexical errors.
3. **Engagingness** Individuals were told to rate the generated responses in terms of how engaging the generated response was, or how well the response attempts to continue the conversation, on a scale from -1 to 1. A score of 1 would indicate a response which follows the logical flow of the dialogue and aims to continue the conversation. A score of -1 would be assigned to a laconic response that is contextually incoherent.

Table 1: Results on the ConvAI2 dialogue corpus. R_r, R_p and *LS* refer to the variance regularization terms presented in Equation 5 and 6, and the lexical diversity selection respectively.

	C-score	Distinct-1	Distinct-2	Persona Consistency	Naturalness	Engagingness
PAML	-0.029	0.089	0.182	0.041	0.141	0.064
MTML	-0.055	0.118	0.268	0.057	0.116	0.083
GPT-2	-0.011	0.097	0.177	-0.079	**0.378**	0.114
CVAE	-0.046	0.152	0.422	0.012	0.227	0.091
DLVGen	0.048	0.151	0.426	0.105	0.276	0.128
+LS	0.049	0.367	0.757	0.123	0.356	0.117
+R_r	0.075	0.171	0.474	0.226	0.325	0.104
+R_r+LS	**0.081**	**0.390**	**0.809**	**0.269**	0.367	0.151
+R_p	0.024	0.161	0.448	0.093	0.181	0.134
+R_p+LS	0.017	0.370	0.775	0.142	0.262	0.148
+R_r+R_p	0.054	0.171	0.474	0.173	0.279	**0.174**
+R_r+R_p+LS	0.062	0.386	0.801	0.214	0.369	0.145

4 RESULTS & DISCUSSION

4.1 Results

The automatic and human evaluation results attained by the models described in section 3.4 are displayed in Table 1. An example of dialogue responses generated by the best performing DLVGen variant (DLVGen+R_r+LS) is provided in Table 2.

4.2 Discussion

Based on the C-scores and Persona Consistency scores attained, it can be easily observed that our proposed model incorporates the corresponding persona information to a larger extent compared to PAML, MTML, GPT-2 and CVAE. All DLVGen variants achieved higher C-scores and Persona Consistency scores. However, while the addition of the variance regularization term R_r generally improves the C-score/Persona Consistency score, the introduction of R_p leads to a slight drop in the same metrics. We suspect that this is because a large variance would increase the chances of modelling incorrect or contradictory persona traits, which would negatively impact the persona consistency of the responses. DLVGen variant that achieved the highest C-score and Persona Consistency score was DLVGen+R_r+LS. Also, it can be observed that the DLVGen model variants with lexical diversity selection achieved better C-scores/ Persona Consistency scores. This indicates that lexically diverse responses tend to contain more persona information.

It should be noted that the C-scores and Persona Consistency scores obtained are dependant on the length of the dialogue context. A longer dialogue context (greater number of utterances) would usually imply greater persona consistency and larger C-score. This is expected as a short dialogue context would not have sufficient persona information to be inferred by the persona prior network. For these cases, we observe that there is a tendency for the model to incorporate a randomly generated persona, which could be due to the failure of the persona prior network to model any persona information from the context. As a result, the generated response might contradict the agent's persona and negatively impact the C-score as well as Persona Consistency score. Ideally, in such cases, the model should be able to recognize that there is insufficient persona information in the dialogue context, and the generated response should be largely persona neutral (section 2.1) while remaining contextual and fluent. Addressing this issue could be a direction for future research. A potential approach could involve estimating the uncertainty, which could be represented by the variance, of the persona prior network, and incorporating the uncertainty estimate in the decoding process.

Upon closer inspection of the responses generated by each of model, we notice that out of the responses generated by GPT-2 and CVAE, few attempt to incorporate the persona information. Instead, a significant fraction of responses were relatively short and generic. On the other hand, responses generated by DLVGen are more likely to a attempt some form of persona incorporation. It should be highlighted that even though it is less likely for generic responses to achieve high C-scores, it is also less likely for them to be assigned negative C-scores. A potential avenue for future work could involve designing a metric which

Table 2: Dialogue responses generated by DLVGen+R_r+LS. Ref indicates the response label. Since $N = 3$, r_1, r_2, r_3 refer to the 3 responses generated after sampling the latent distributions 3 times. The response selected by the lexical diversity selector is in bold.

Persona	
i have been trying all types of food everywhere i go.	
hey there i'm 23 and i love food.	
i've been traveling the world for a years.	
i also like to cook but i'm not very good at it.	
i own a yacht and i rent it out when i'm not using it.	
Dialogue Context	
User:	hello. how are you today ?
Agent:	good today . just cooking some mexican food. i cooking but am not very good.
User:	oh i like mexican food , but my favorite food are cheeseburgers.
Agent:	i inherited some money and bought a yacht to travel, i try different foods traveling
User:	i help out at a soup kitchen since i grew up poor
Responses	
Ref :	cheeseburgers are great , i try all kinds of foods everywhere i go , gotta love food.
r_1 :	i love to cook , i am a chef.
r_2 :	i love to cook.
r_3 :	**i love to travel , i have been to many countries.**

penalizes generic responses in addition to inaccurate responses.

From our experimental results, the GPT-2 base model attained noticeably lower Distinct 1 and 2 scores compared to the latent variable models. This follows previous work which reported an increase in response diversity due to the stochasticity introduced by the latent variable (Zhao et al., 2017; Song et al., 2019). Furthermore, the DLVGen model variants with lexical diversity selection achieved better Distinct-1 and 2 scores compared to all other models. This is not surprising since lexical diversity would generally imply general diversity. The improvements on both diversity and persona consistency demonstrate the efficacy of our lexical diversity selection approach.

When it comes to the Naturalness, we can observe that the usage of the GPT-2 pretrained language model resulted in higher scores. Both PAML and MTML, which utilize standard Transformers, attained noticeably lower Naturalness scores compared to all other models. This is to be expected as pretrained language models such as GPT-2 have demonstrated high language understanding and generation capability even without any finetuning. We can also observe that most of the latent variable models attained comparable Naturalness scores. The only exception is the inclusion of the variance regularization term R_p, which results in a lower Naturalness score. The GPT-2 base model, however, maintains an edge over the latent variable models in terms on Naturalness score. This could be attributed to the short, generic responses generated by the base model, which are relatively more fluent and human-like. On the other hand, when it comes to the Engagingness score, all implemented models achieved relatively poor performance. Relatively few responses attempt to continue the conversation. Instead, most responses are largely informative.

5 RELATED WORK

5.1 Latent Variable Models

Latent variable models are a category of models that involve inferring latent random variables from a group of observable variables (Verbeke and Molenberghs, 2017). Latent variable models such as the Variational Auto Encoder (VAE) (Kingma and Welling, 2014) and the Conditional Variational Auto Encoder (CVAE) (Sohn et al., 2015) have been applied to the task of open-domain dialogue generation, where the potential dialogue responses are modelled as a latent Gaussian distribution (Li et al., 2020; Shen et al., 2018; Zhao et al., 2017; Serban et al., 2017). In addition to personalized dialogue generation (examples provided in the introduction), CVAEs have been applied to conditional dialogue generation tasks such as emotional dialogue generation (Liu et al., 2021; As-

ghar et al., 2020; Zhou and Wang, 2018) as well as topical dialogue generation (Wang et al., 2020).

5.2 Personalized Dialogue

Over the past few years, there has been numerous publications exploring various approaches to the task of personalized dialogue generation. As mentioned in the introduction, there are numerous approaches to the task of personalized dialogue generation. A popular approach typically involves conditioning dialogue responses on the dialogue context in addition to textual persona descriptions (Lee et al., 2021; Na et al., 2021; Liu et al., 2020; Majumder et al., 2020; Madotto et al., 2019; Wolf et al., 2019; Zheng et al., 2020; Chan et al., 2019; Song et al., 2019). This approach focuses on generating responses which incorporate the provided persona information. Another approach involves incorporating personality or persona related metadata into the decoding process (Qian et al., 2018). Some approaches also involve implicitly learning personality user embeddings (Zheng et al., 2019a; Wu et al., 2020; Al-Rfou et al., 2016; Li et al., 2016). Another approach entails inferring the dialogue agent's personality directly from the dialogue history (Zheng et al., 2019b; Su et al., 2019). Typically, for this approach, the primary aim is to train the agent to mimic the dialogue style of the interlocutor.

6 CONCLUSION

In this paper, we introduced DLVGen, a dual latent variable model which models the potential responses and the agent's potential persona as latent Gaussian distributions. Through our experiments, we find that responses generated by DLVGen effectively incorporates persona information inferred from the dialogue context. We also introduced a variance regularization technique and lexical diversity selection method which improves the quality of the generated responses in terms of both persona consistency and human-likeness. However, an area for improvement is the relatively poor engagingness of the dialogue. Encouraging the generation of persona consistent, diverse yet engaging open-domain dialogue is a potential avenue for future research. A possible approach involves designing an objetcive function which explicitly accounts for the engagingness of the generated response. The dialogue model could then be trained on both objective functions via a multi-task learning framework.

REFERENCES

Al-Rfou, R., Pickett, M., Snaider, J., Sung, Y., Strope, B., and Kurzweil, R. (2016). Conversational contextual cues: The case of personalization and history for response ranking. *CoRR*, abs/1606.00372.

Asghar, N., Kobyzev, I., Hoey, J., Poupart, P., and Sheikh, M. B. (2020). Generating emotionally aligned responses in dialogues using affect control theory.

Chan, Z., Li, J., Yang, X., Chen, X., Hu, W., Zhao, D., and Yan, R. (2019). Modeling personalization in continuous space for response generation via augmented Wasserstein autoencoders. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1931–1940, Hong Kong, China. Association for Computational Linguistics.

Dinan, E., Logacheva, V., Malykh, V., Miller, A., Shuster, K., Urbanek, J., Kiela, D., Szlam, A., Serban, I., Lowe, R., Prabhumoye, S., Black, A. W., Rudnicky, A., Williams, J., Pineau, J., Burtsev, M., and Weston, J. (2019). The second conversational intelligence challenge (convai2).

Fergadiotis, G., Wright, H. H., and Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research*, 58(3):840–852.

Kingma, D. P. and Welling, M. (2014). Auto-encoding variational bayes.

Lee, J. Y., Lee, K. A., and Gan, W. S. (2021). Generating personalized dialogue via multi-task meta-learning. In *Proceedings of the 25th Workshop on the Semantics and Pragmatics of Dialogue - Full Papers*, Potsdam, Germany. SEMDIAL.

Li, C., Gao, X., Li, Y., Peng, B., Li, X., Zhang, Y., and Gao, J. (2020). Optimus: Organizing sentences via pre-trained modeling of a latent space. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4678–4699, Online. Association for Computational Linguistics.

Li, J., Galley, M., Brockett, C., Spithourakis, G., Gao, J., and Dolan, B. (2016). 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.

Liu, M., Bao, X., Liu, J., Zhao, P., and Shen, Y. (2021). Generating emotional response by conditional variational auto-encoder in open-domain dialogue system. *Neurocomputing*, 460:106–116.

Liu, Q., Chen, Y., Chen, B., Lou, J.-G., Chen, Z., Zhou, B., and Zhang, D. (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.

Madotto, A., Lin, Z., Wu, C.-S., and Fung, P. (2019). Personalizing dialogue agents via meta-learning. In *Pro-*

ceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 5454–5459, Florence, Italy. Association for Computational Linguistics.

Majumder, B. P., Jhamtani, H., Berg-Kirkpatrick, T., and McAuley, J. (2020). Like hiking? you probably enjoy nature: Persona-grounded dialog with commonsense expansions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9194–9206, Online. Association for Computational Linguistics.

Na, Y., Park, J., and Sohn, K.-A. (2021). Is your chatbot perplexing?: Confident personalized conversational agent for consistent chit-chat dialogue. In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART,*, pages 1226–1232. INSTICC, SciTePress.

Qian, Q., Huang, M., Zhao, H., Xu, J., and Zhu, X. (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.

Radford, A. and Wu, J. (2019). Rewon child, david luan, dario amodei, and ilya sutskever. 2019. *Language models are unsupervised multitask learners. OpenAI Blog*, 1(8):9.

Serban, I. V., Sordoni, A., Lowe, R., Charlin, L., Pineau, J., Courville, A., and Bengio, Y. (2017). A hierarchical latent variable encoder-decoder model for generating dialogues. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, AAAI'17, page 3295–3301. AAAI Press.

Shen, X., Su, H., Niu, S., and Demberg, V. (2018). Improving variational encoder-decoders in dialogue generation. In *AAAI*.

Sohn, K., Lee, H., and Yan, X. (2015). Learning structured output representation using deep conditional generative models. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

Song, H., Zhang, W., Cui, Y., Wang, D., and Liu, T. (2019). Exploiting persona information for diverse generation of conversational responses. In *IJCAI*.

Su, F.-G., Hsu, A. R., Tuan, Y.-L., and Lee, H.-Y. (2019). Personalized Dialogue Response Generation Learned from Monologues. In *Proc. Interspeech 2019*, pages 4160–4164.

Verbeke, G. and Molenberghs, G. (2017). Modeling through latent variables. *Annual Review of Statistics and Its Application*, 4(1):267–282.

Wang, Y., Si, P., Lei, Z., and Yang, Y. (2020). Topic enhanced controllable cvae for dialogue generation (student abstract). *Proceedings of the AAAI Conference on Artificial Intelligence*, 34:13955–13956.

Wolf, T., Sanh, V., Chaumond, J., and Delangue, C. (2019). Transfertransfo: A transfer learning approach for neural network based conversational agents. *CoRR*, abs/1901.08149.

Wu, B., Li, M., Wang, Z., Chen, Y., Wong, D. F., Feng, Q., Huang, J., and Wang, B. (2020). Guiding variational response generator to exploit persona. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 53–65, Online. Association for Computational Linguistics.

Zhang, S., Dinan, E., Urbanek, J., Szlam, A., Kiela, D., and Weston, J. (2018). Personalizing dialogue agents: I have a dog, do you have pets too? In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia. Association for Computational Linguistics.

Zhao, T., Zhao, R., and Eskenazi, M. (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.

Zheng, Y., Chen, G., Huang, M., Liu, S., and Zhu, X. (2019a). Personalized dialogue generation with diversified traits. *CoRR*, abs/1901.09672.

Zheng, Y., Zhang, R., Huang, M., and Mao, X. (2020). A pre-training based personalized dialogue generation model with persona-sparse data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):9693–9700.

Zheng, Y., Zhang, R., Mao, X., and Huang, M. (2019b). A pre-training based personalized dialogue generation model with persona-sparse data. *CoRR*, abs/1911.04700.

Zhou, X. and Wang, W. Y. (2018). Mojitalk: Generating emotional responses at scale. *ArXiv*, abs/1711.04090.