An Adversarial Approach to High-Quality, Sentiment-Controlled Neural Dialogue Generation

Xiang Kong, Bohan Li, Graham Neubig, Eduard Hovy, Yiming Yang
Language Technologies Institute
Carnegie Mellon University
{xiangk, bohanl1, gneubig, hovy, yiming}@cs.cmu.edu

Abstract

In this work, we propose a method for neural dialogue response generation that allows not only generating semantically reasonable responses according to the dialogue history, but also explicitly controlling the *sentiment* of the response via sentiment labels. Our proposed model is based on the paradigm of conditional adversarial learning; the training of a sentiment-controlled dialogue generator is assisted by an adversarial *discriminator* which assesses the fluency and feasibility of the response generating from the dialogue history and a given sentiment label. Because of the flexibility of our framework, the generator could be a standard sequence-to-sequence (SEQ2SEQ) model or a more complicated one such as a conditional variational autoencoder-based SEQ2SEQ model. Experimental results using automatic and human evaluation both demonstrate that our proposed framework is able to generate both semantically reasonable and sentiment-controlled dialogue responses.

Introduction

Sentiment is a fundamental part of the human communication, and reflecting sentiment in human-computer interfaces is a key to making them engaging and interesting to use. This is certainly true for dialog systems and there has been a large body of literature that attempts to equip dialogue systems with the ability to understand and express the sentiment (Polzin and Waibel 2000; Skowron et al. 2011; Partala and Surakka 2004; Prendinger and Ishizuka 2005; Hasegawa et al. 2013; Zhou et al. 2017; Shen et al. 2017). However, these methods are either based on templates or rules which require extensive hand engineering.

End-to-end neural dialogue generation (Shang, Lu, and Li 2015; Vinyals and Le 2015; Serban et al. 2016) is now a popular research topic because of the ease and flexibility of creating systems within this paradigm. While early works (Shang, Lu, and Li 2015; Vinyals and Le 2015) just employ simple sequence-to-sequence (SEQ2SEQ) models similar to those used in machine translation (Cho et al. 2014; Sutskever, Vinyals, and Le 2014), a number of papers have aimed to further improve the quality and diversity of dialogue responses in manners specific to dialog systems (Li et al. 2015; Li et al. 2016; Gu et al. 2016; Xing et al. 2016; Li et al. 2017; Serban et al. 2017; Zhao, Zhao, and Eskenazi 2017).

Furthermore, There have also been a few attempts (Skowron et al. 2011; Hasegawa et al. 2013; Shen et al. 2017; Zhou et al. 2017) to incorporate sentiment information into data-driven end-to-end dialog systems, but each has their own shortcomings. For example, Hasegawa et al. (2013) propose a method to train individual systems for each kind of emotion, which will cause the system to suffer from data sparsity and high computational cost. In addition, Shen et al. (2017) incorporate latent variables expressing emotion into the dialogue system but do not provide an explicit way to control the sentiment of these responses.

Recently, Zhou and Wang (2018) collect a large corpus of tweets from Twitter with emojis in the response, and assume that these emojis could reflect the sentiment of the response. Furthermore, they train a conditional variational autoencoder (CVAE)-based neural dialogue system which is capable of controlling the sentiment of the generated response explicitly. In this work, we investigate the application of another powerful model, i.e., generative adversarial networks (GANs) to this problem.

In this paper, we propose a conditional generative adversarial network (CGAN)-based framework for sentiment-controlled dialogue generation. In this framework, the desiderata of fluency and controlability are explicitly enforced by creating a model with two subcomponents: a generator and a discriminator. The generator is in charge of generating sentimental responses given a dialogue history and a sentiment label, while the adversarial discriminator enforces sentimental response quality by trying to determine whether the item (*dialogue history, sentiment label, dialogue response*) comes from the real data distribution. By training the generator to fool the discriminator, our system can simultaneously improve the quality of dialogue responses and generate responses with different sentiments depending on the sentiment label.

Background

Problem Setting

Our task is to train a dialogue system which is able to generate high-quality and sentiment-controlled responses. Given the dialogue history $\mathbf{W}_h = \{\mathbf{w}_1, \ldots, \mathbf{w}_i, \ldots, \mathbf{w}_{h_N}\}$, where $\mathbf{w}_i$ is the i-th token and h_N is the length of the sequence, and a sentiment label y, the task is to generate a response

$\mathbf{W}_r = \{\mathbf{w}_1, \ldots, \mathbf{w}_i, \ldots, \mathbf{w}_{r_N}\}$ where r_N is the length of the generated response. We would like this response to be consistent with the sentiment label y and semantically appropriate for the dialogue history $\mathbf{W}_h$.

Encoder-Decoder for Dialog Generation

Most neural models for dialogue generation are based on the encoder-decoder structure, a.k.a, sequence-to-sequence (SEQ2SEQ) framework (Cho et al. 2014; Sutskever, Vinyals, and Le 2014), in which an encoder reads in the previous dialog history/context and encodes it a continuous vector representation c, which the decoder then uses to output the next dialogue utterance.

Specially, given the dialogue history $\mathbf{W}_h$, the hidden state of the encoder at time t, h_t^{enc}, is computed according to:

$$h_t^{enc} = \mathbf{RNN}_{enc}(h_{t-1}^{enc}, \mathbf{w}_t),\ t = 1, ..., h_N, \quad (1)$$

where $\mathbf{RNN}_{enc}$ is the encoder RNN. Finally, we obtain the vector representation of $\mathbf{W}_h$, i.e., $c = h_{h_N}^{enc}$.

The decoder is another RNN which is capable of generating a response $\mathbf{W}_r$ given the context vector c. The hidden state of the decoder at the time step t is calculated by:

$$h_t^{dec} = \mathbf{RNN}_{dec}(h_{t-1}^{dec}, c, \mathbf{w}_{t-1}) \quad (2)$$

where $\mathbf{RNN}_{dec}$ is the decoder RNN.

The probability p_t^{dec} over the whole vocabulary at t-th time step is then calculated by a softmax function conditioned on the hidden state h_t^{dec}.

By multiplying all probabilities of the gold word tokens at each time step, we can calculate the probability $p(\mathbf{W}_r|\mathbf{W}_h)$ of the response sequence $\mathbf{W}_r$ given the dialogue history sequence $\mathbf{W}_h$.

Generative Adversarial Nets

The second important technology contributing to the proposed method is Generative Adversarial Networks (GANs; Goodfellow et al. (2014)).

In the original GAN framework, there are two models: a generative model G, which is in charge of generating outputs (one example being the SEQ2SEQ model in the previous section), and a discriminative model D that attempts to discriminate whether its input samples are real or generated outputs. By training G to create outputs that are able to fool D into thinking that they are real, it is possible to generate samples that seem highly realistic, improving the quality of generation of images (Salimans et al. 2016) and text (Li et al. 2017; Yu et al. 2017).

However, one major problem in the GAN framework is that there is no mechanism to control attributes of generated items. Therefore, Mirza and Osindero (2014) propose a condition adversarial nets (CGANs) in which both the generator and discriminator are conditioned on some extra information so that the generator can control the types of items being generated according to this extra information.

Back to the dialogue scenario, Li et al. (2017) propose an adversarial dialogue generation model, in which the generative model G is a standard SEQ2SEQ model (Sutskever, Vinyals, and Le 2014) which could generate a response $\mathbf{W}_r$

given the dialogue history $\mathbf{W}_h$ according to Eq.1 and Eq.2. The discriminative model D is a binary classifier which takes the dialogue history $\mathbf{W}_h$ and a dialogue response $\mathbf{W}_r$ as an input and outputs a label $D(\mathbf{W}_h, \mathbf{W}_r)$ indicating whether the dialogue response $\mathbf{W}_r$ is generated from machines or human beings.

In more detail, its objective is to maximize the expected reward, i.e., $D(\mathbf{W}_h, \mathbf{W}_r)$ of generated responses :

$$J = \mathbb{E}_{\mathbf{W}_r \sim G_\theta(\cdot|\mathbf{W}_h)}[D(\mathbf{W}_h, \mathbf{W}_r)] \quad (3)$$

Variational Autoencoders

Another popular deep generative model recently is the framework of variational autoencoders (VAEs). VAEs have been successfully applied to many text generation tasks (Bowman et al. 2015; Serban et al. 2017; Zhou and Neubig 2017; Hu et al. 2017). Specifically, in Serban et al. (2017), the dialogue generation model has been augmented by introducing a latent variable z at the decoder. Shen et al. (2017) and Zhou and Wang (2018) present a CVAEs-based framework for dialogue generation in which the response $\mathbf{W}_r$ is generated from a stochastic latent variable z and and the context vector c. Mathematically, a CVAE-based dialogue generation system maximizes a variational lower bound on the conditional likelihood of $\mathbf{W}_r$ given the latent variable z and the context vector c.

Approaches

In this section, we build upon the SEQ2SEQ-based dialogue generation model with the CVAE and CGAN techniques introduced in the previous section.

Sentiment-Context SEQ2SEQ

In order to explicitly control the sentiment of the generated response, we slightly change the structure of the standard SEQ2SEQ model (Sutskever, Vinyals, and Le 2014). Specifically, after obtaining the context vector c of the dialogue history $\mathbf{W}_h$ from the encoder by Eq.1, the concatenation of a sentiment vector s and c is fed into the decoder to generate a response and this vector is called "sentiment context", s_c.

To generate the sentiment vector s, similarly to word embedding, we first map the sentiment label y to a vector y_e, and this vector will be fed into a fully-connected neural network to output the sentiment vector s.

The computation graph of the Sentiment-Context SEQ2SEQ model is shown in Figure 1.

Conditional Variational Autoencoders (CVAEs) SEQ2SEQ

We follow the model structure described in Sohn, Lee, and Yan (2015) and Zhou and Wang (2018) to build the CVAE-SEQ2SEQ model.

Mathematically, the objective of CVAE-SEQ2SEQ is to maximize the lower bound probability of the response given the sentiment context vector, i.e.,

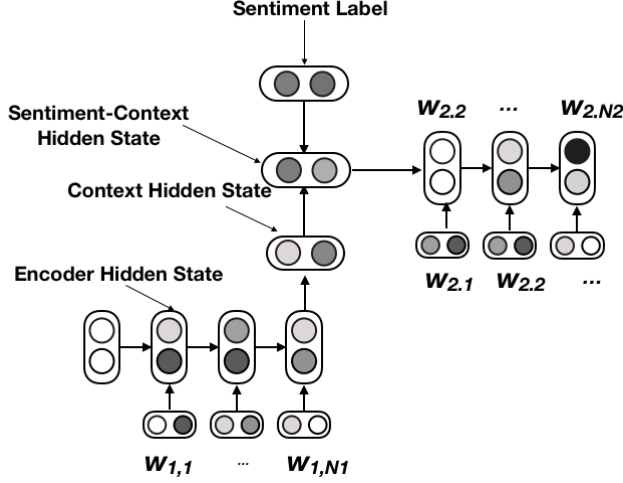$$p(\mathbf{W}_r|s_c) = \int p(\mathbf{W}_r|z, s_c)p(z|s_c)dz \quad (4)$$

Figure 1: The computational graph of the sentiment-context SEQ2SEQ architecture. The dialogue history is encoded into a dense vector c via an encoder RNN, then the concatenation of the context vector c and the sentiment vector s computed from the sentiment label y is fed into a decoder RNN to generate tokens in the response.

where z is the latent variable, s_c is the sentiment context vector mentioned before. Based on the assumption that the latent variable follows a multivariate Gaussian distribution with a diagonal covariance matrix, the lower bound of $\log p(\mathbf{W}_r|s_c)$ is:

$$\mathcal{L}_{\text{CVAE}} = \mathbb{E}_{q(z|\mathbf{W}_r, s_c)}(\log p(\mathbf{W}_r|z, s_c) \\ - \text{KL}(q(z|\mathbf{W}_r, s_c)||p(z|s_c)) \quad (5)$$

where $p(\mathbf{W}_r|z, s_c)$ is modeled by another decoder which is different from the decoder in the seq2seq model, $q(z|\mathbf{W}_r, s_c)$ is described by a recognition network and $p(z|s_c)$ is modeled by a prior network, both of which are MLP-based neural networks..

In more detail, for the encoder RNN, CVAE-SEQ2SEQ uses the same setting as the SEQ2SEQ model to encode the dialogue history $\mathbf{W}_h$ into a context vector. The decoder RNN, however, is different because it now takes the concatenation of the sentiment context vector and the sampled stochastic latent variable as input to generate a response. At training time, the latent variable sample z is drawn from an approximate posterior network and used for optimizing the variational lower-bound given by Eq. 5. At test time, the latent variable sample z is drawn from a prior network for decoding, which has no knowledge of the ground-truth response. Furthermore, the bag-of-word loss (Zhao, Zhao, and Eskenazi 2017) has been added in the above objective function. Therefore, the final objective function for the CVAE-SEQ2SEQ is:

$$\mathcal{L} = \mathcal{L}_{\text{CVAE}} + \mathcal{L}_{bow} \quad (6)$$

Conditional Generative Adversarial Net SEQ2SEQ

Adversarial training methods have been successfully applied to neural dialogue generation (Li et al. 2017) to improve
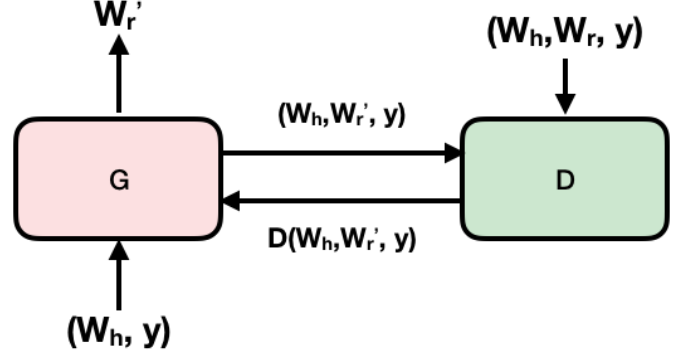

Figure 2: The computational graph of the conditional generative adversarial networks-based SEQ2SEQ architecture. "G" denotes the generator and "D" refers to the discriminator. $W_r^{'}$ is the response being generated from the generator.

the quality of generated responses. However, in their model, the property of the response such as the sentiment, could not be controlled explicitly. Therefore, we propose a conditional generative adversarial network-based dialogue system named CGAN-SEQ2SEQ which is able to improve the quality of the response and control its sentiment at the same time. The model is shown in Figure 2.

Our proposed CGAN-SEQ2SEQ consists of two components, i.e., a conditional generator G and a conditional discriminator D.

Generator G. We adopt the original sentiment-context SEQ2SEQ as the G which could generate a response $\mathbf{W}_r$ given the dialogue history $\mathbf{W}_h$ and a sentiment label y. The goal of the G is to produce high-quality and sentiment-controlled responses as similar as those being generated from human beings so as to fool the discriminator D.

Discriminator D. The discriminator G in our framework is to identify whether the input response is generated from human beings or machines given the dialogue history $\mathbf{W}_h$ and the sentiment label y. Specifically, the discriminator consists of two encoders. The first encoder is similar to that in the SEQ2SEQ model which is able to encode the input dialogue history to a representation vector which will be concatenated with the sentiment vector to compose the sentiment context vector. For the second encoder, the initialize state will be set as the sentiment context vector allowing the decoder to condition on the $\mathbf{W}_h$ and y, then it encodes the response sequence to a representation vector. Finally, the concatenation of this vector and sentiment context vector will be fed into a fully-connected neural network-based binary classifier to compute the final result. The reason why we utilize the sentiment context vector again is to let the D pay more attention to the sentiment information. The computation graph of the D is illustrated in the Figure 3.

A Game with Two Players. Following the training process mentioned in the Li et al. (2017), we first pre-train a generator without the discriminator, then freeze the parameters of the pre-trained generator to pre-train the discriminator. During pre-training of the discriminator, responses
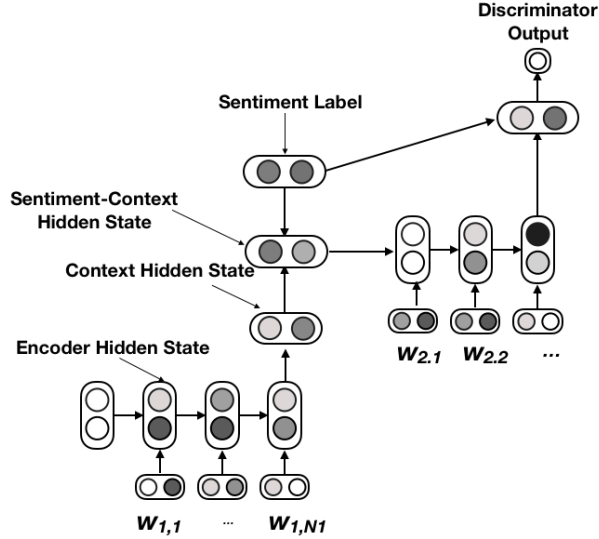
Figure 3: The discriminator structure in the CGAN-SEQ2SEQ model. After obtaining the final vector from the second encoder, the concatenation of this vector and the sentiment context vector will be fed into an MLP layer for the final output.

generated from the pre-trained generator and human beings are regarded as negative and positive samples respectively. Finally, the generator G and the discriminator D play a two-player game. In this game, the generator first generates a response $\mathbf{W}_r^{'}$ given the dialogue history $\mathbf{W}_h$ and the sentiment label y, then the discriminator provides $D(\mathbf{W}_r^{'}, \mathbf{W}_h, y)$ back to the generator and use triples ($\mathbf{W}_h$, $\mathbf{W}_r^{'}$, y) and ($\mathbf{W}_h$, $\mathbf{W}_r$, y) to train itself. The generator will be optimized according to the $D(\mathbf{W}_r^{'}, \mathbf{W}_h, y)$ obtained from discriminator.

Policy Gradient Training. Similarly to Li et al. (2017), the generator in our proposed framework is a probabilistic transformation from the dialogue history to the dialogue response, both in discrete space. Therefore, we also employ the REINFORCE algorithm (Williams 1992) to optimize it.

The objective of the generator is to maximize the expected reward of generated responses:

$$J = \mathbb{E}_{\mathbf{W}_r^{'} \sim G_\theta(\cdot|\mathbf{W}_h, y)}[D(\mathbf{W}_h, \mathbf{W}_r^{'}, y)] \qquad (7)$$

.

Note that $D(\mathbf{W}_h, \mathbf{W}_r^{'}, y)$ can be regarded as the probability of the response $\mathbf{W}_r^{'}$ being generated from human beings given the $\mathbf{W}_h$ and y. The gradient with respect to the θ in Eq.7 could be approximately computed by the likelihood ratio trick (Williams 1992):

$$\nabla_\theta \approx [D(\mathbf{W}_h, \mathbf{W}_r^{'}, y) - b(\mathbf{W}_h, \mathbf{W}_r^{'}, y)]$$
$$\nabla_\theta \sum_t \log p(w_t|\mathbf{W}_h, y, W_{1:t-1}) \qquad (8)$$

where $b(\mathbf{W}_h, \mathbf{W}_r^{'}, y)]$ is the baseline value of the expected reward which could reduce the variance of the estimate while keeping it unbiased (Ng, Harada, and Russell 1999).

Intuitively, when the generated response is more likely to fool the D, the larger reward the G will get, and thus parameters will be updated with a larger step.

One advantage of the CGAN-SEQ2SEQ over CVAE-SEQ2SEQ is that it will not change the structure of the SEQ2SEQ model. During response generation, the discriminator could be removed and the SEQ2SEQ model remains the same.

We found training to be unstable if we just use Eq.8 to optimize the G. Therefore, we also employ the teacher forcing procedure (Li et al. 2017) to assist the training so that the generator has access to the golden response.

CGAN-CVAE SEQ2SEQ

From the Figure 2, it is easy to find that the generator could also be the CVAE-SEQ2SEQ. Therefore, we propose a CGAN-CVAE SEQ2SEQ model, in which the generator is the CVAE-SEQ2SEQ model and the discriminator stays the same as that in the CGAN SEQ2SEQ model. Intuitively, from the reinforcement learning perspective, the discriminator is regarded as the reward provider and for a high-quality generated response, it will assign a higher reward back to the generator.

In order to stabilize the training process, besides adding the teaching forcing method mentioned in the previous section, we also add the original CVAE objective to the Eq. 7 to create a hybrid objective function (Zhou and Wang 2018), i.e.,

$$\mathcal{L}^{'} = \mathbb{E}_{\mathbf{W}_r^{'} \sim G_\theta(\cdot|\mathbf{W}_h, y)}[D(\mathbf{W}_h, \mathbf{W}_r^{'}, y)] + \mathcal{L} \qquad (9)$$

Experimental Results

Dataset

To evaluate our proposed framework, we use the large corpus of tweets with emojis collected and used in Zhou and Wang (2018). To simplify the task, we classify all emojis into two clusters, i.e., positive and negative. As a result, there are approximately $374\mathbf{K}$, $21\mathbf{K}$ and $21\mathbf{K}$ tweets in train, dev and test sets respectively. The ratio of the ratio of positive to negative samples is around 3:1.

Evaluation Metrics

Perplexity: Perplexity is a common metric used in many natural language tasks and connected to the likelihood of the gold response given a dialogue generation model. Although the diversity of responses generated by the dialogue system is very important, the system should nonetheless assign a relatively high likelihood to the ground truth response.

Sentiment Accuracy: Because the goal of our task is to control the sentiment of the response given a sentiment label and a dialogue history, whether the generated response correctly reflects the sentiment is very important. Therefore, we build a sentiment classifier on the training set and evaluate the generated responses by how often the classifier-predicted label represents the specified sentiment y.

Human Evaluation: Because automatic metrics are suboptimal for evaluating performance of dialogue generation systems (Liu et al. 2016), we ask three judges to evaluate

30 random items, each of which consists of a dialogue history, a gold response, and a generated response. Judges are expected to evaluate in two settings.

- In one setting, the goal is to evaluate the quality of dialogue responses from different models. We use a 1-5 scale where 5 means that the response and the dialogue history is highly relevant semantically and syntactically, and 1 means they are irrelevant.
- In the other setting, judges are asked to label the sentiment of the given responses as positive or negative. In this case only the generated responses are provided to the judges.

Note that these two experiments are conducted separately and the items are different in order to avoid bias.

Implementation Details

Sentiment Classifier Our sentiment classifier is 1-layer bidirectional RNN-GRU encoder with 128 hidden units in each direction. This is fed into an MLP classifier to predict the final sentiment class.

We employ a standard SEQ2SEQ (Sutskever, Vinyals, and Le 2014) model with attention (Luong, Pham, and Manning 2015) to build the sentiment-context SEQ2SEQ model. The encoder is a 1-layer bidirectional GRU with hidden size 128 in each direction, and the dimension of the sentiment vector is 12. The decoder is a 1-layer GRU of size $128 * 2 + 12 = 268$. The Adam optimizer with a 1e-3 learning rate and gradients clipped to 5 is employed to train this model.

Following the experiment settings in Zhou and Wang (2018), we incorporate a response encoder, a recognition network and a prior network into the above SEQ2SEQ model to build a CVAE-SEQ2SEQ model. The response encoder is another 1-layer bidirectional GRU of size 128 in each direction. The mean and log variance of latent variable z is obtained from the recognition and the prior network, both of which are two fully-connected networks, then latent variables are sampled via the reparameterization trick (Kingma and Welling 2013). During generation without golden responses, the latent variable sampled from the prior network will be directly fed into the decoder.

Main Results

Capacity for Sentiment Control: The sentiment control capacity of each model is evaluated by the sentiment accuracy metric. As shown in Table 1, the CGAN-CVAE SEQ2SEQ model outperforms all the other models in sentiment accuracy, indicating that, combining CGANs and CVAEs together, the generator could control the sentiment of the response more effectively than the respective baselines. Although the sentiment accuracy of the CGAN-SEQ2SEQ is better than the SEQ2SEQ model, it can not control the sentiment of the response as well as the CVAE-SEQ2SEQ model. We suspect that this is because during REINFORCE training the generator can only access the generated sentences, which will be noise to deteriorate the generator if they are of low quality. We have found that the responses from the pre-trained generator are indeed generic and do not control the

sentiment information well. However, the CVAE-SEQ2SEQ model can utilize the golden response at every training step.
Response Quality: We employ Perplexity (PPL), which is shown in Table 1, as a proxy to evaluate the response quality. Compared with other models, the CGAN-CVAE SEQ2SEQ model achieves the lowest PPL score, which means that its likelihood of generating the golden response is highest. Similarly to the sentiment accuracy, the PPL of the CGAN-SEQ2SEQ is higher than that of the CVAE-SEQ2SEQ and we attribute this to the same reason mentioned above.

Human Evaluation

The human evaluation result is shown in Table 2. With respect to both content quality and sentiment accuracy, the CGAN-CVAE has better accuracy than other models. This demonstrates that our proposed CGAN-CVAE could not only generate high-quality dialogue response but effectively control the sentiment of dialogue responses as well, which is consistent with the automatic evaluation results. The overall performance of the CVAE model is also better than that of the CGAN model.

Case Study

In order to show the differences between the performances of these models more concretely, we show some examples in Table 3. We can clearly see that the responses generated from the CGAN-CVAE SEQ2SEQ and CVAE-SEQ2SEQ models are more distinctive given different sentiment labels and topics related to the dialogue context. For CGAN-SEQ2SEQ, the sentiment of the response is relatively consistent with the sentiment label but compared with CVAE-SEQ2SEQ, the diversity of responses is relatively low. As for the SEQ2SEQ model, it seems that it only remembers some sentimental words and the responses are quite dull and generic.

	Perplexity	Sentiment Acc (%)
SEQ2SEQ	157.5	55.6
CVAE	81.83	75.6
CGAN	120.3	64.4
CGAN-CVAE	**69.54**	**78.8**

Table 1: Evaluation of various dialogue systems with perplexity and sentiment accuracy.

	Quality	Sen-Acc(%)
SEQ2SEQ	2.1	54.4
CVAE	3.6	73.3
CGAN	2.9	66.7
CGAN-CVAE	**3.9**	**78.9**

Table 2: Dialogue response quality and sentiment accuracy (Sen-Acc) of different dialogue systems based on human evaluation.

Context	goldlink is dope live one of my favorite shows i 've been to	
Sentiment	Positive	Negative
SEQ2SEQ	i 'm so happy for you	i 'm so sad
CGAN	i 'm gonna be there	i 'm not sure i 'm gonna be able to find it
CVAE	i like the song	i feel like i was gonna cry
CVAE-CGAN	omg i love it	that 's the worst
Context	and i never got $\langle dgt \rangle$ lol	
Sentiment	Positive	Negative
SEQ2SEQ	i 'm so excited	i 'm so sorry
CGAN	he 's so cute	i 'm not sure i 'm not going to be a fan of $\langle dgt \rangle$
CVAE	i love to hear that ! i 'm so happy to hear this	well , didn 't realize you had to get the wrong name
CVAE-CGAN	lmao i m looking for it	i 'm sorry for you
Context	always got ya my dude no matter what ! lets bowl	
Sentiment	Positive	Negative
SEQ2SEQ	i'm so mad	i 'm not sure if it 's a good idea
CGAN	i 'm glad you 're enjoying it	i 'm not sure if you 're joking
CVAE	we are doing a great job !	wow . i hate you
CVAE-CGAN	yes i love you guys , but it is a good time	i mean i hate my bestfriend

Table 3: Response samples from different dialogue models given different sentiment labels.

Related Work

The sentiment is crucial to human-human communication, and thus for machines to communicate smoothly with humans, it is necessary for machines to generate utterances with sentiment.

Skowron et al. (2011) propose that affective profiles in a dialogue system are strongly correlated with the emotional changes experienced by participants. Partala and Surakka (2004) show that positive affective intervention could be especially useful to enhance users' problem solving performance. Prendinger and Ishizuka (2005) indicate that a computer agent with empathic feedback could support people preparing for an interview. Polzin and Waibel (2000) describe how a dialogue system adjusts its interaction according to the emotion of users.

While these works are valuable proofs-of-concept, they generally either focus on small-scale corpora or using rule-based templates to generate responses, which make them difficult to extend to large-scale open-domain dialogue generation.

Some recent works have tried to incorporate sentiment in large-scale conversation generation. Hasegawa et al. (2013) study how utterances affect the emotion of other speakers, and try to predict the emotion of the user to generate a response which reflects the specific emotion using the statistical machine translation framework (Ritter, Cherry, and Dolan 2011). Within the framework of neural dialogue systems, pioneering work by Shen et al. (2017) incorporate sentiment into the variational hierarchical encoder-decoder model (Serban et al. 2017). However, their model is trying to improve the quality of the response instead of controlling the sentiment of the response. In parallel, Zhou et al. (2017) tried to model sentiment in a dialogue conversation system through three mechanisms: emotion category embedding, internal emotion memory, and external memory. The internal memory models the change of the internal emotion state of the decoder, and therefore encodes how much an emotion has already been expressed. The external memory decides whether to choose an emotional or generic (non-emotional) word at a given step during decoding.

Conclusion

In this paper, we propose an intuitive training objective for neural dialogue generation, which is to control the sentiment of the generated response explicitly. This objective is implemented through a conditional adversarial training paradigm, in which the generator is trained to generate sentiment-controlled responses via sentiment labels assisted by a discriminator. Furthermore, the generator in our system could be the standard SEQ2SEQ or CVAEs-based SEQ2SEQ models. Our system adopts a policy gradient algorithm to deal with the optimization challenge posed by discrete generator outputs. Experiments clearly demonstrate the effectiveness of such an adversarial training objective in successfully controlling the sentiment of the response and improving the content quality.

Future directions include validating our approach on more fine-grained sentiment-based data and improved combination with other advanced techniques in reinforcement learning and adversarial learning such as reward shaping, etc.

References

[2015] Bowman, S. R.; Vilnis, L.; Vinyals, O.; Dai, A. M.; Jozefowicz, R.; and Bengio, S. 2015. Generating sentences from a continuous space. *arXiv preprint arXiv:1511.06349*.

[2014] Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

[2014] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *NIPS*, 2672–2680.

[2016] Gu, J.; Lu, Z.; Li, H.; and Li, V. O. 2016. Incorporating copying mechanism in sequence-to-sequence learning. *arXiv preprint arXiv:1603.06393*.

[2013] Hasegawa, T.; Kaji, N.; Yoshinaga, N.; and Toyoda, M. 2013. Predicting and eliciting addressee's emotion in online dialogue. In *Proceedings of the 51st ACL*, 964–972.

[2017] Hu, Z.; Yang, Z.; Liang, X.; Salakhutdinov, R.; and Xing, E. P. 2017. Toward controlled generation of text. In *International Conference on Machine Learning*, 1587–1596.

[2013] Kingma, D. P., and Welling, M. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

[2015] Li, J.; Galley, M.; Brockett, C.; Gao, J.; and Dolan, B. 2015. A diversity-promoting objective function for neural conversation models. *arXiv preprint arXiv:1510.03055*.

[2016] Li, J.; Monroe, W.; Ritter, A.; Galley, M.; Gao, J.; and Jurafsky, D. 2016. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*.

[2017] Li, J.; Monroe, W.; Shi, T.; Ritter, A.; and Jurafsky, D. 2017. Adversarial learning for neural dialogue generation. *arXiv preprint arXiv:1701.06547*.

[2016] Liu, C.-W.; Lowe, R.; Serban, I. V.; Noseworthy, M.; Charlin, L.; and Pineau, J. 2016. How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. *arXiv preprint arXiv:1603.08023*.

[2015] Luong, M.-T.; Pham, H.; and Manning, C. D. 2015. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*.

[2014] Mirza, M., and Osindero, S. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

[1999] Ng, A. Y.; Harada, D.; and Russell, S. 1999. Policy invariance under reward transformations: Theory and application to reward shaping. In *ICML*, volume 99, 278–287.

[2004] Partala, T., and Surakka, V. 2004. The effects of affective interventions in human–computer interaction. *IWC*.

[2000] Polzin, T. S., and Waibel, A. 2000. Emotion-sensitive human-computer interfaces. In *ITRW*.

[2005] Prendinger, H., and Ishizuka, M. 2005. The empathic companion: A character-based interface that addresses users'affective states. *Applied Artificial Intelligence* 19(3-4):267–285.

[2011] Ritter, A.; Cherry, C.; and Dolan, W. B. 2011. Data-driven response generation in social media. In *Proceedings of the conference on EMNLP*, 583–593. ACL.

[2016] Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; and Chen, X. 2016. Improved techniques for training gans. In *NIPS*, 2234–2242.

[2016] Serban, I. V.; Sordoni, A.; Bengio, Y.; Courville, A. C.; and Pineau, J. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *AAAI*, 3776–3784.

[2017] Serban, I. V.; Sordoni, A.; Lowe, R.; Charlin, L.; Pineau, J.; Courville, A. C.; and Bengio, Y. 2017. A hierarchical latent variable encoder-decoder model for generating dialogues. In *AAAI*, 3295–3301.

[2015] Shang, L.; Lu, Z.; and Li, H. 2015. Neural responding machine for short-text conversation. *arXiv preprint arXiv:1503.02364*.

[2017] Shen, X.; Su, H.; Li, Y.; Li, W.; Niu, S.; Zhao, Y.; Aizawa, A.; and Long, G. 2017. A conditional variational framework for dialog generation. *arXiv preprint arXiv:1705.00316*.

[2011] Skowron, M.; Rank, S.; Theunis, M.; and Sienkiewicz, J. 2011. The good, the bad and the neutral: affective profile in dialog system-user communication. In *ACII*, 337–346. Springer.

[2015] Sohn, K.; Lee, H.; and Yan, X. 2015. Learning structured output representation using deep conditional generative models. In *NIPS*, 3483–3491.

[2014] Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to sequence learning with neural networks. In *NIPS*, 3104–3112.

[2015] Vinyals, O., and Le, Q. 2015. A neural conversational model. *arXiv preprint arXiv:1506.05869*.

[1992] Williams, R. J. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8(3-4):229–256.

[2016] Xing, C.; Wu, W.; Wu, Y.; Liu, J.; Huang, Y.; Zhou, M.; and Ma, W.-Y. 2016. Topic augmented neural response generation with a joint attention mechanism. *arXiv preprint arXiv:1606.08340*.

[2017] Yu, L.; Zhang, W.; Wang, J.; and Yu, Y. 2017. Seqgan: Sequence generative adversarial nets with policy gradient. In *AAAI*, 2852–2858.

[2017] Zhao, T.; Zhao, R.; and Eskenazi, M. 2017. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders. *arXiv preprint arXiv:1703.10960*.

[2017] Zhou, C., and Neubig, G. 2017. Multi-space variational encoder-decoders for semi-supervised labeled sequence transduction. *arXiv preprint arXiv:1704.01691*.

[2018] Zhou, X., and Wang, W. Y. 2018. Mojitalk: Generating emotional responses at scale. In *Proceedings of the 56th ACL*. Melbourne, Victoria, Australia: ACL.

[2017] Zhou, H.; Huang, M.; Zhang, T.; Zhu, X.; and Liu, B. 2017. Emotional chatting machine: Emotional conver-

sation generation with internal and external memory. *arXiv preprint arXiv:1704.01074*.