Neural Response Generation via GAN with an Approximate Embedding Layer*

Zhen Xu[1], **Bingquan Liu**[1], **Baoxun Wang**[2], **Chengjie Sun**[1], **Xiaolong Wang**[1], **Zhuoran Wang**[2] and **Chao Qi**[2]
[1]School of Computer Science and Technology,
Harbin Institute of Technology, Harbin, China
[2]Tricorn (Beijing) Technology Co., Ltd, Beijing, China
[1]{zxu,bqliu,cjsun,wangxl}@insun.hit.edu.cn
[2]{wangbaoxun, wangzhuoran, qichao}@trio.ai

Abstract

This paper presents a Generative Adversarial Network (GAN) to model single-turn short-text conversations, which trains a sequence-to-sequence (Seq2Seq) network for response generation simultaneously with a discriminative classifier that measures the differences between human-produced responses and machine-generated ones. In addition, the proposed method introduces an approximate embedding layer to solve the non-differentiable problem caused by the sampling-based output decoding procedure in the Seq2Seq generative model. The GAN setup provides an effective way to avoid non-informative responses (a.k.a "safe responses"), which are frequently observed in traditional neural response generators. The experimental results show that the proposed approach significantly outperforms existing neural response generation models in diversity metrics, with slight increases in relevance scores as well, when evaluated on both a Mandarin corpus and an English corpus.

1 Introduction

After achieving remarkable successes in Machine Translation (Sutskever et al., 2014; Cho et al., 2014), neural networks with the encoder-decoder architectures (a.k.a sequence-to-sequence models, Seq2Seq) have been proven to be a functioning method to model short-text conversations (Vinyals and Le, 2015; Shang et al., 2015), where the corresponding task is often called Neural Response Generation. The advantage of applying Seq2Seq models to conversation generation is that the training procedure can be performed end-to-end in an unsupervised manner, based on human-generated conversational utterances (typically query-response pairs mined from social networks). One of the potential applications of such neural response generators is to improve the capability of existing conversational interfaces (informally also known as chatbots) by enabling them to go beyond predefined tasks and chat with human users in an open domain.

However, previous research has indicated that naïve implementations of Seq2Seq based conversation models tend to suffer from the so-called "safe response" problem (Li et al., 2016a), i.e. such models tend to generate non-informative responses that can be associated to most queries, e.g. "I don't know", "I think so", etc. This is due to the fundamental nature of statistical models, which fit sufficiently observed examples better than insufficiently observed ones. Concretely, the space of open-domain conversations is so large that in any sub-sample of it (i.e. a training set), the distribution of most pieces of information are relatively much sparser when compared to safe response patterns. Furthermore, since a safe response can be of relevance to a large amount of diverse queries, a statistical learner will tend to minimize its empirical risk in the response generation process by capturing those safe responses if naïve relevance-oriented loss metrics are employed.

Frequent occurrences of safe responses can dramatically reduce the attractiveness of a chat agent, which therefore should be avoided to the best extent possible when designing the learning algorithms. The pathway to achieve this purpose is to seek a more expressive model with better capacity that can take relevance and diversity (or informativeness) into account simultaneously

*The work was done when the first author was an intern at Tricorn (Beijing) Technology Co., Ltd.

Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 626–635
Copenhagen, Denmark, September 7–11, 2017.

when modelling the underlying distribution of human conversations.

Generative Adversarial Nets (GANs) (Goodfellow et al., 2014; Chen et al., 2016) offers an effective architecture of jointly training a generative model and a discriminative classifier to generate sharp and realistic images. This architecture could also potentially be applied to conversational response generation to relieve the safe response problem, where the generative part can be an Seq2Seq-based model that generates response utterances for given queries, and the discriminative part can evaluate the quality of the generated utterances from diverse dimensions according to human-produced responses. However, unlike the image generation problems, training such a GAN for text generation here is not straightforward. The decoding phase of the Seq2Seq model usually involves sampling discrete words from the predicted distributions, which will be fed into the training of the discriminator. The sampling procedure is non-differentiable, and will therefore break the back-propagation.

To the best of our knowledge, Reinforcement Learning (RL) is first introduced to address the above problem (Li et al., 2017; Yu et al., 2017), where the score predicted by a discriminator was used as the reinforcement to train the generator, yielding a hybrid model of GAN and RL. But to train the RL phrase, Li et al. (2017) introduced two approximations for reward computing at each action (word) selection step, including a Markov Chain Monte Carlo (MCMC) sampling method and a partial utterance scoring approach. It has been stated in their work that the former approach is time-consuming and the latter one will result in lower performance due to the overfitting problem caused by adding a large amount of partial utterances into the training set. Nevertheless, we also want to argue that, besides the time complexity issue of MCMC, RL itself is not an optimal choice either. As shown in our experimental results in Section 5.1, a more elegant design of an end-to-end differentiable GAN can significantly increase the model's performance in this text generation task.

In this paper, we propose a novel variant of GAN for conversational response generation, which introduces an approximate embedding layer to replace the sampling-based decoding phase, such that the entire model is continuous and differentiable. Empirical experiments are conducted based on two datasets, of which the results show that the proposed method significantly outperforms three representative existing approaches in both relevance and diversity oriented automatic metrics. In addition, human evaluations are carried out as well, demonstrating the potential of the proposed model.

2 Related Work

Inspired by recent advances in Neural Machine Translation (NMT), Ritter et al. (2011) and Vinyals and Le (2015) have shown that single-turn short-text conversations can be modelled as a generative process trained using query-response pairs accumulated on social networks. Earlier works focused on paired word sequences only, while Zhou et al. (2016) and Iulian et al. (2017) have demonstrated that the comprehensibility of the generated responses can benefit from multi-view training with respect to words, coarse tokens and utterances. Moreover, Sordoni et al. (2015) proposed a context-aware response generation model that goes beyond single-turn conversations.

In addition, attention mechanisms were introduced to Seq2Seq-based models to capture topic and dialog focus information by Shang et al. (2015) and Chen et al. (2017), which had been proven to be helpful for improving query-response relevance (Wu et al., 2016). Additional features such as persona information (Li et al., 2016b) and latent semantics (Zhou et al., 2017; Serban et al., 2017) have also been proven beneficial within this context.

When compared to previous work, this paper is focused on single-turn conversation modeling, and employs a GAN to yield informative responses.

3 Building a Conversational Response Generator via GAN

3.1 Notations

Let $\mathcal{D} = \{(q_i, r_i)\}_{i=1}^{N}$ be a set of N single-turn human-human conversations, where $q_i = (w_{q_i,1}, \ldots, w_{q_i,t}, \ldots, w_{q_i,m})$ is a query, $r_i = (w_{r_i,1}, \ldots, w_{r_i,t}, \ldots, w_{r_i,n})$ stands for the response to q_i, and $w_{q_i,t}$ and $w_{r_i,t}$ denote the t-th words in q_i and r_i, respectively. This paper aims to learn a generative model $G(r|q)$ based on a discriminator D that can predict informative responses with good diversity for arbitrary input queries.

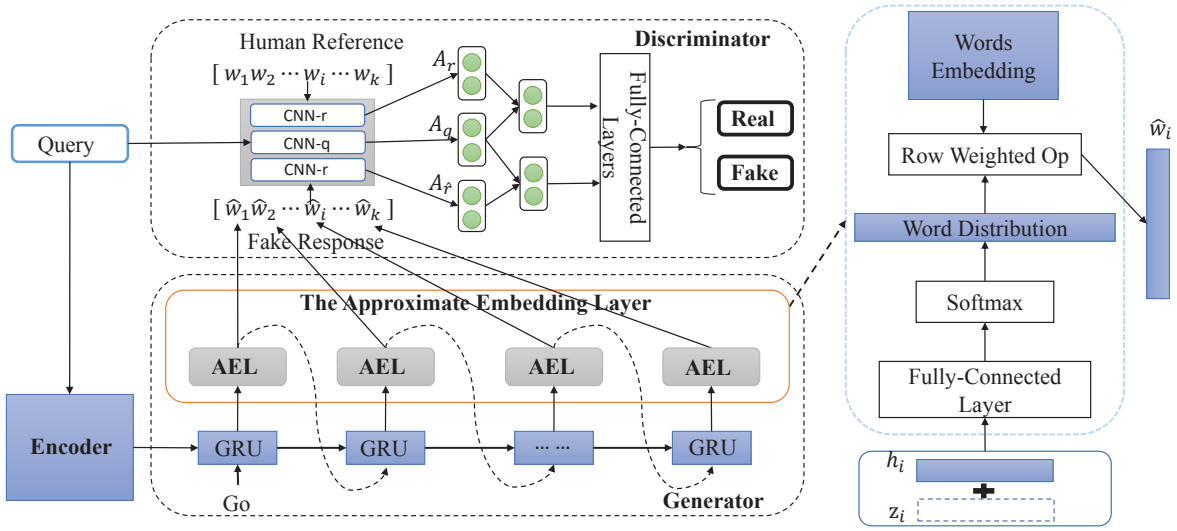

Figure 1: The Framework of GAN for the Response Generator.

3.2 Model Overview

We name the proposed model Generative Adversarial Network with an Approximate Embedding Layer (GAN-AEL), of which Figure 1 illustrates the overall framework. Generally speaking, the whole framework consists of a response generator G, a discriminator D and an embedding approximation layer that connects the G and the D. We explain each of the components in detail as follows. The generator adopts the Gated Recurrent Unit (GRU) (Cho et al., 2014) based encoder-decoder architecture, where the encoder projects the input query (a discrete word sequence) into a real-valued vector, on which the output will be generated conditionally in the decoding process, activated by a starting signal (denoted as "Go" in Figure 1). An approximate embedding layer is designed to guarantee that the response generation procedure is continuous and differentiable, serving as an interface for the discriminator to propagate its loss to the generator. The Convolutional Neural Network (CNN) based discriminator is attached on top of the approximation layer, which aims to distinguish the fake responses output by the approximation layer and the corresponding human-generated references, conditioned on the input query. The judgement of the CNN can be propagated to the Seq2Seq generator through the proposed approximate embedding layer, and forces the generator to be fine-tuned to produce more attractive results.

The proposed GAN framework possesses several advantages over existing conversational response generation models. Firstly, both the generator and the discriminator are conditioned on the input query, which guarantees the relevance of the generated responses. Secondly, the discriminator enforces the generator to produce a response according to the true distribution in better granularity, such that the state of promoting safe responses is leaped out. Thirdly, the approximation layer yields a smooth connection between the generator and the discriminator, avoiding the non-differentiable discrete sampling process.

3.3 Pre-training the Generator by MLE

In our proposed encoder-decoder framework, both the encoder and the generator (i.e. the decoder) G is composed of GRU (Cho et al., 2014) units, which is designed to generate responses $r = \{w_{r,1}, w_{r,2}, \cdots, w_{r,K}\}$ conditioned on an input query $q = \{w_{q,1}, w_{q,2}, \cdots, w_{q,J}\}$. For a given query-response pair (q, r), the target is to maximum the conditional probability $p(r|q)$ in the generative process. Concretely, in this model, q is firstly encoded into a vector representation q_v by the GRU-based encoder as shown in Figure 1, which is actually the last hidden state of the encoder. Then the generator estimates the probability of each word occurring in r conditioned on q_v. Hence $p(r|q)$ can be formulated as follows:

$$p(r|q) = \prod_{t=1}^{K} p(w_{r,t}|q_v, w_{r,1}, \cdots, w_{r,t-1}) \quad (1)$$

Taking the logarithm of the probabilities for

effective computation, the generator is trained by optimising the Maximum Likelihood Estimation (MLE) objective defined as:

$$\frac{1}{|\mathcal{D}|}\sum_{(q,r)\in\mathcal{D}}\sum_{t=1}^{K}\log p(w_{r,t}|q_v, w_{r,1}, \cdots, w_{r,t-1}) \quad (2)$$

Note here, we need to pre-train the generator using Equation 2 as the loss function to guarantee the generator to produce grammatical utterances. Otherwise, the discriminator will tend to learn a rule with ease to distinguish human-produced utterances from those ungrammatical responses generated in the early stages of the training phase, which would cause the failure of the training in satisfying Nash Equilibrium (Goodfellow et al., 2014).

3.4 The Approximate Embedding Layer

In order to smoothly connect the output layer of the generator to the input layer of the discriminator to yield an end-to-end differentiable GAN, one needs to solve the following critical problem. The output of the generator is a sequence of discrete words, which is usually sampled from the distributions predicted by the decoder's RNN units in the Softmax layer, and is non-differentiable.

Since afterward those words will be projected into embedding vectors to feed the CNN-based discriminator, we introduce an embedding approximation layer to merge the generation process of the decoder and the word embedding phrase of the discriminator. This can be done by multiplying the word probabilities in the distributions obtained from the decoder's Softmax layer to the corresponding word vectors, to directly yield an approximately vectorized representation of the generated word sequences for further convolutional computations in the discriminative process. This approximation is based on the assumption that ideally the word distributions should be trained to reasonably approach the one-hot representations of the discrete words.

The structure of the approximation layer is illustrated on the right-hand side of Figure 1. Concretely, the approximation layer takes the output h_i of the generator and a random noise z_i as the input, and reuses the word projection layer (pre-trained in the standard generator) to estimate the probability distribution of w_i. Note that, the noise z_i added to h_i forms a latent feature for the word embedding approximation process to enforce the diversity of the generated responses. The overall word embedding approximation is computed as:

$$\hat{e}_{w_i} = \sum_{j=1}^{V} e_j \cdot \text{Softmax}(W_p(h_i + z_i) + b_p)_j \quad (3)$$

where W_p and b_p are the weight and bias parameters of the word projection layer, respectively, and h_i is the hidden representation of word w_i, from the decoding procedure of the generator G, which is computed as:

$$h_i = g(h_{i-1}, \hat{e}_{w_{i-1}}) \quad (4)$$

where $g(\cdot)$ is the standard GRU inference step in G (Cho et al., 2014).

3.5 Pre-training the CNN-based Discriminator

CNN has been proven to be an appropriate classifier for many NLP tasks, such as sentence classification (Kim, 2014) and matching (Hu et al., 2014). Therefore, in this paper we adopt a CNN-based discriminator as shown in Figure 1.

For the convenience of further discussions, we introduce $\hat{r}$ to denote the underlying (distributional) fake response produced by the decoder. In other words, $\hat{r}$ stands for a sequence of word distributions projected from the hidden layers of the decoder RNN, based on which one would sample the output response utterance in a traditional Seq2Seq generator. The detailed architecture of the discriminator is described as follows. Firstly, the input of the discriminator consist of the word embedding vector sequence V_q for a given query q and the word embedding vector sequence V_r for its human-produced response r, as well as the approximate word embedding vector sequence $V_{\hat{r}}$ produced by the approximate embedding layer for the corresponding fake response $\hat{r}$. All the word embedding vector sequences here are zero-padded or truncated to a same fixed length. After this, two CNNs with shared parameters are employed to encode V_r and $V_{\hat{r}}$ into higher-level abstractions, respectively. In addition, a separate CNN is used to abstract V_q in a similar way. We denote such abstraction layers (i.e. the max-pooling layers before the fully-connected layers) in the above CNNs as A_r, $A_{\hat{r}}$ and A_q, corresponding to r, $\hat{r}$ and q, respectively. Finally, we concatenate A_q

to A_r and $A_{\hat{r}}$ separately, and feed the resulting vectors to their respective fully-connected layers, as illustrated in Figure 1. Here, we make the two fully-connected layers share common parameters and predict probabilities $D(r|q)$ and $D(\hat{r}|q)$, respectively, for r and $\hat{r}$ being true responses of the given q.

In practice, when the Seq2Seq generative network G is pre-trained, we also pre-train the above discriminator D by maximising the following objective function:

$$D_{loss} = \log D(r|q) + \log(1 - D(\hat{r}|q)) \quad (5)$$

with the parameters of G frozen, before the adversarial training procedure described in Section 3.6.

3.6 Adversarial Training of the Generator

After the pre-training of the generator G and the discriminator D as described above, the entire network is trained adversarially. Concretely, we iteratively train G and D, where at each iteration, the parameters of the non-training network will be frozen. The following tricks are utilised in the adversarial training phase to achieve better convergence. Firstly, when training G, we replace the objective function given in Equation 5 with the l_2-loss between A_r and $A_{\hat{r}}$, to maintain a reasonable scale of the gradient. Secondly, we freeze the parameters of the encoder network and the projection layer of the decoder network, but only tune the parameters of decoder's hidden layers. This is based on the assumption that, in principle, after the pre-training, the encoder network is sufficiently effective to represent the entire input utterance, while the projection layer of the decoder is also adequate to decode words from its hidden states. Therefore, the adversarial training here is to adjust the "wording strategy" of the generative model, i.e. the way it organises the semantic contents during the decoding (or in other words, the way it realises the hidden states). Preliminary experiments show that this trick significantly improves the grammaticalness of the generated responses.

The gradient of the generator can be computed as:

$$\begin{aligned} \nabla_{g_{D,G}(\theta_G)} &= \frac{\partial G_{loss}}{\partial V_{\hat{r}}} \frac{\partial V_{\hat{r}}}{\partial \theta_G} \\ &= \frac{\partial G_{loss}}{\partial V_{\hat{r}}} \frac{\partial V_{\hat{r}}}{\partial G} \frac{\partial G}{\partial \theta_G} \end{aligned} \quad (6)$$

where θ_G denotes the active parameters of the generator G, $G_{loss} = \|A_r - A_{\hat{r}}\|$ and $g_{D,G}(\cdot)$ stands for the inference step of the entire GAN. It can be seen that the feedback signals from D can be propagated to G effectively through the approximate embedding layer, which connects G and D smoothly, and avoids the discrete sampling procedure.

4 Experiment Setup

4.1 Datasets

We test our model on two datasets: **Baidu Tieba** and **OpenSubtitles** (Lison and Tiedemann, 2016). The Baidu Tieba dataset is composed of single-turn conversations collected from the threads of Baidu Tieba[1], of which the utterance length ranging from 3 to 30 words. The OpenSubtitles dataset contains movie scripts organised by characters, where we follow Li et al. (2016a) to retain subtitles containing 5-50 words in the following experiments. From each of the two datasets, we sample 5,000,000 unique single-turn conversations as the training data, 200,000 additional unique pairs for validation, and another 10,000 as the test set.

4.2 Baselines

To illustrate the performance of the proposed model, we introduce three existing approaches as baselines.

- **Seq2Seq:** the standard sequence-to-sequence model (Sutskever et al., 2014).
- **MMI-anti:** a Seq2Seq model with a Maximum Mutual Information (MMI) criterion (implemented as an anti-language model) (Li et al., 2016a) in the decoding process, which reduces the probability of generating "safe responses".
- **Adver-REGS:** another adversarial strategy proposed by Li et al. (2017)[2], which links the generator and the discriminator together with a reinforcement learning framework, and takes the discriminator's output probability as the reward to train the generator.

[1] https://tieba.baidu.com/index.html

[2] The codes can be accessed at https://github.com/jiweil/Neural-Dialogue-Generation/tree/master/Adversarial

4.3 Evaluation Metrics

For automatic evaluations, the following commonly accepted metrics are employed. Note here, the goal of our model is to obtain responses not only semantically relevant to the corresponding queries, but also of good diversity and novelty. Therefore, in this work, embedding-based metrics (Liu et al., 2016) are adopted to evaluate semantic the relevance between queries and their corresponding generated responses, while **dist-1**, **dist-2** (Li et al., 2016a) are used as diversity measures. In addition, we also introduce a **Novelty** measure as detailed below.

Relevance Metrics: The following three word embedding based metrics[3] are used to compute the semantic relevance of two utterances. The **Greedy** metric is to greedily match words in two given utterances based on the cosine similarities of their embeddings, and to average the obtained scores (Rus and Lintean, 2012). Alternatively, an utterance representation can be obtained by averaging the embeddings of all the words in that utterance, of which the cosine similarity gives the **Average** metric (Mitchell and Lapata, 2008). In addition, one can also achieve an utterance representation by taking the largest extreme values among the embedding vectors of all the words it contains, before computing the cosine similarities between utterance vectors, which yields the **Extreme** metric (Forgues et al., 2014).

Diversity Metrics: To measure the informativeness and diversity of the generated responses, we follow the **dist-1** and **dist-2** metrics proposed by Li et al. (2016a) and Chen et al. (2017), and introduce a **Novelty** metric. The **dist-1** (**dist-2**) is defined as the number of unique unigrams (bigrams for dist-2). A common drawback of dist-1 and dist-2 is that in the computation, less informative words (such as "I", "is", etc.) are considered equally with those more informative ones. Therefore, in this paper, we define an extra **Novelty** metric, which is the number of infrequent words observed in the generated responses. Here we take all the words except the top 2000 most frequent ones in the vocabulary as infrequent words. Note here, the dist-1 and Novelty values are normalised by utterance length, and dist-2 is normalised by the total number of bigrams in the generated response.

Human Evaluation: To evaluate the performance of our model from human perspectives, this paper conducts a human subject experiement by comparing the responses generated by AdverREGS (which is one of the most competitive existing approaches) with those by the proposed model. Three experienced annotators are invited to evaluate 200 groups of examples. In the evaluation, for every given query, the annotators will see 10 generated responses from each model. Since the proposed method aims at improving the diversity of the responses generated by Seq2Seq models, while maintaining their relevance to the input queries, we ask the annotators to evaluate the diversity performance of the two systems only if there is no obvious difference between the performance of their relevance. This experimental setting is due to the following two reasons. Firstly, it is difficult to judge a systems diversity based on one single response (Li et al., 2016a; Zhou et al., 2017). Secondly, the practical deployment of a chat-oriented conversational system will usually decode an N-best list of candidate responses, from which it random samples the final reply. Considering that all the annotators use Mandarin as their first language, the above evaluation is only done on the Tieba dataset.

4.4 Hyperparameters & Training Strategies

Hyperparameter Settings: The hyperparameters of the networks used in all the experiments below are described as follows. The vocabulary sizes for Tieba and OpenSubtitles are truncated to 100,000 and 150,000, respectively. The dimensions of word embedding vectors are set to 100 for Tieba and 300 for OpenSubtitles. The size of the hidden layers in the generator is set to 200 in the all experiments on both datasets. We experiments subsets of $\{1,2,3,4\}$ for the filter sizes of the CNNs, and fixed the filter number to 128. As shown in subsection 5.3, CNNs with filter sizes $\{1,2\}$ are the best choice here. Max-pooling is used in all the CNN settings here. The noise Z is sampled from a normal distribution with 0 mean and 0.1 variance.

Training Strategies: To train the proposed GAN, the parameters of the generator G are initialised based on the pre-training mentioned in subsection 3.3, while those of the discriminator D are randomly initialised. The adversarial training

[3]The implementation of all these metrics follows the code at https://github.com/julianser/hed-dlg-truncated/tree/master/Evaluation.

Model	Relevance			Diversity		
	Average	**Greedy**	**Extreme**	**Dist-1**	**Dist-2**	**Novelty**
Seq2Seq	0.720	0.614	0.571	0.0037	0.0121	0.0102
MMI-anti	0.713	0.592	0.552	0.0127	0.0495	0.0250
Adver-REGS	0.722	0.660	0.574	0.0153	0.0658	0.0392
GAN-AEL	**0.736**	**0.689**	**0.580**	**0.0214**	**0.0963**	**0.0635**

Table 1: Relevance and diversity evaluation on the Tieba dataset.

Model	Relevance			Diversity		
	Average	**Greedy**	**Extreme**	**Dist-1**	**Dist-2**	**Novelty**
Seq2Seq	0.719	0.578	0.505	0.0054	0.0141	0.0045
MMI-anti	0.710	0.569	0.499	0.0175	0.0586	0.0097
Adver-REGS	0.726	0.590	0.507	0.0223	0.0725	0.0147
GAN-AEL	**0.734**	**0.621**	**0.514**	**0.0296**	**0.0955**	**0.0216**

Table 2: Relevance and diversity evaluation on the OpenSubtitles dataset.

starts from pre-training D with the parameters of G fixed. After this, G and D will be trained iteratively with different learning rates, which are 0.0001 for D and 0.00002 for G. In addition, we update D at a frequency of every 5 batches instead of every single batch.

5 Experimental Results

5.1 Automatic Evaluation & Analysis

From Table 1 and 2, it can be observed that the proposed GAN-AEL model outperforms the baselines on both datasets in all metrics, especially for the diversity oriented scores. The improvements can be explained from the following two angles.

a) Since a vanilla Seq2Seq model does not take diversity, novelty or informativeness into account, the discriminator tends to capture such information to distinguish model-generated responses and human responses. By backpropagating the discriminator's feedback to the generator, the adversarially trained generator gains significantly better performance in such aspects. On the other hand, the relevance is also retained during the adversarial training, as one can imagine that the human produced references given to the discriminator are usually semantically highly relevant to the corresponding queries.

b) The proposed approximation layer is an effective way to couple the response generator and the discriminator. Through this differentiable component, the loss of the discriminator is properly propagated to the generator and guide the tuning of the latter's parameters.

It can also be seen from the results that the performance of all the models on the three semantic relevance oriented metrics are comparable to each other. This implies that all the models, including the baseline methods and the proposed model, have the capability to generate responses of reasonable relevance to given queries, which satisfies the primary goal of the response generation task. It further suggests that the Seq2Seq architecture works properly in modelling the semantics of entire utterances. Nevertheless, although the decoder mechanism can select topic-relevant words to construct responses based on the given query, the limitation of naïve Seq2Seq models tend to yield less diverse or informative outputs.

Furthermore, when compared to Adver-REGS, the proposed GAN-AEL gains 30%-60% relative improvement in the dist-1, dist-2 and novelty metrics on both datasets, which indicates that coupling the generator and the discriminator with a differentiable component is a more preferable methodology for text generation tasks, and is a meaningful analogy to standard GANs for image generation. Interestingly, all the models achieve significantly higher novelty scores on the Tieba dataset than on the OpenSubtitle dataset. This is due to the difference of the coverages of high-frequency words in the two corpora. Concretely, since we exclude top 2,000 most frequent words when computing the novelty scores on both datasets, which covers 70% and 82% of the words in Tieba and OpenSubtitle respectively, it is more

likely to observe novel words on the Tieba data.

In addition, it can be seen that GAN-AEL improves the greedy score to a much greater extent than the average and extreme scores, which further suggests that the responses generated by GAN-AEL are more informative. Concretely, the calculations of the average and extreme scores may be dominated by generic non-informative words. By contrast, since the greedy metric is computed based on a (simple and greedy) word-wise semantic alignments between two utterances, the influence of those generic words will be reduced.

5.2 Human Evaluation Results

Table 3 gives the human evaluation results, which indicates that the proposed GAN-AEL is more preferable than Adver-REGS from human perspectives. This again implies that the approximate embedding layer is more effective in propagating the discriminator's feedback to the generator than the reinforcement learning mechanism of (Li et al., 2017). The result is statistically significant with $p < 0.01$ according to sign test.

GAN-AEL vs Adver-REGS		
Wins	Losses	Ties
0.61	0.13	0.26

Table 3: Evaluations of GAN-AEL and Adver-REGS based on human subjects,

5.3 The Influence of the Discriminator to Adversarial Training

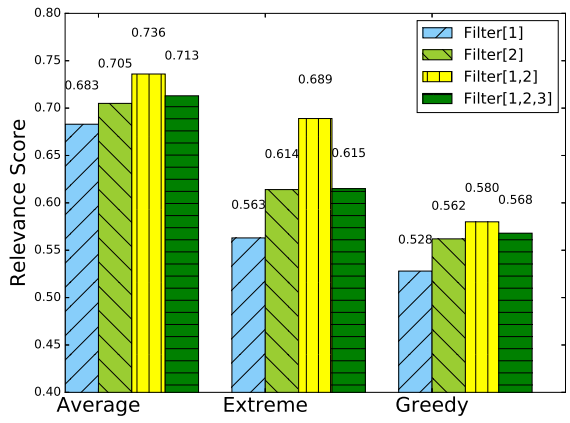

Figure 2: Relevance scores of GAN-AEL on the Tieba corpus with respect to different CNN window sizes.

The discriminator plays an important role in the adversarial training process, which determines whether the GAN model converges to a Nash Equilibrium (Chen et al., 2016). We conduct a set of experiments to explore the influence of the discriminator's capacities to the adversarial training. Figure 2 shows the relevance scores with respect to different convolution window sizes for the CNN discriminator, where "Filter[x]" denotes the CNN with its convolution window(s) set to x.

It can be found that the discriminator with "Filter[1, 2]" achieves the best performance. Two facts based on the principle of GAN could be taken to explain this observation: On one hand, a discriminator with too low capacity (such as "Filter[1]") is less capable in distinguishing human responses from generated ones, which will backpropagate inappropriate signals that misleads the generator. On the other hand, if the capacity of the discriminator is too high (such as "Filter[1, 2, 3]"), in the adversarial training, the training of the discriminator will converge too fast before the generator being sufficiently trained (Durugkar et al., 2016). A similar phenomenon is also observed on the OpenSubtitles dataset. Since the relevance between queries and their corresponding responses is essential in this task, discriminators with "Filter[1,2]" are adopted in the experiments reported in subsections 5.1 and 5.2.

5.4 Case Study

To demonstrate the intuitive performance of the proposed model in comparison with the naïve Seq2Seq model, we provide some example cases in Figure 3, where for each query the response is the top hypothesis obtained via beam search. Especially, we show that when the vanilla approach generates safe responses such as "I don't know what to do", the adversarially trained method can still yield more informative outputs. Furthermore, when both models produces informative responses, the utterance generated by GAN-AEL tends to be of more interestingness, which is a reflection of novelty and diversity.

6 Conclusion

In this paper, we proposed a GAN framework to model single-turn short-text conversations. An approximation embedding layer is introduced to force the entire network differentiable, which significantly overcomes the drawbacks found in

Example responses generated based on the OpenSubtitle data	Example responses generated based on the Baidu Tieba data
Query: You swore an oath when you put that uniform on. **Seq2Seq:** I don't know what to do. **GAN-AEL:** I think I had a good reason.	**Query:** 你喜欢猫还是狗？Do you like cats or dogs? **Seq2Seq:** 喜欢养猫。I Like cats. **GAN-AEL:** 我喜欢金毛。I like golden retriever.
Query: Entire town knows your son is a goon. **Seq2Seq:** What do you mean? **GAN-AEL:** His life is really a sad story.	**Query:** 你像奥巴马的妻子。You look like Obama's wife. **Seq2Seq:** 哈哈哈哈哈哈。Haha… **GAN-AEL:** 我是大众脸。I have a common face.
Query: Turn his face away from me. **Seq2Seq:** I don't know what to do. **GAN-AEL:** Let me explain to you about the mistake.	**Query:** 我不光搭讪，还请吃饭呢。 I wish to invite you to dinner rather than chat up only. **Seq2Seq:** 我也想吃。I want to have dinner too. **GAN-AEL:** 我是胖子，减肥。I am fat, and I am on a diet.

Figure 3: Example responses generated by Seq2Seq and GAN-AEL.

the previous RL-based attempts (Li et al., 2017). The superiority of the proposed method has been demonstrated by empirical experiments based on both automatic evaluation metrics and human judgements. Further explorations of GAN-based techniques to model contextual information in dialogue problems will be addressed in our future research.

Acknowledgments

We thank the anonymous reviewers for their insightful comments. We also thank Dr. Deyuan Zhang and Dr. Xin Wang for their great help. This research is partially supported by National Natural Science Foundation of China (No.61572151, No.61602131, No.61672192) and the National High Technology Research and Development Program ("863" Program) of China (No.2015AA015405).

References

Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. 2016. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in Neural Information Processing Systems 29*, pages 2172–2180.

Xing Chen, Wu Wei, Wu Yu, Liu Jie, Huang Yalou, Zhou Ming, and Ma Wei-Ying. 2017. Topic aware neural response generation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3351–3357.

Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734.

Ishan Durugkar, Ian Gemp, and Sridhar Mahadevan. 2016. Generative multi-adversarial networks. *Proceedings of the 4th International Conference on Learning Representations (ICLR)*.

Gabriel Forgues, Joelle Pineau, Jean-Marie Larchevêque, and Réal Tremblay. 2014. Bootstrapping dialog systems with word embeddings. In *NIPS, Modern Machine Learning and Natural Language Processing Workshop*.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems 27*, pages 2672–2680.

Baotian Hu, Zhengdong Lu, Hang Li, and Qingcai Chen. 2014. Convolutional neural network architectures for matching natural language sentences. In *Advances in Neural Information Processing Systems 27*, pages 2042–2050.

Serban Iulian, Vlad, Klinger Tim, Tesauro Gerald, Talamadupula Kartik, Zhou Bowen, Bengio Yoshua, and C. Courville Aaron. 2017. Multiresolution recurrent neural networks: An application to dialogue response generation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3288–3294.

Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751.

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 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 110–119.

Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, and William B. Dolan. 2016b. A persona-based neural conversation model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Jiwei Li, Will Monroe, Tianlin Shi, Alan Ritter, and Dan Jurafsky. 2017. Adversarial learning for neural dialogue generation. *arXiv preprint arXiv:1701.06547*.

Pierre Lison and Jörg Tiedemann. 2016. Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*.

Chia-Wei Liu, Ryan Lowe, Iulian V Serban, Michael Noseworthy, Laurent Charlin, and Joelle Pineau. 2016. How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2122–2132.

Jeff Mitchell and Mirella Lapata. 2008. Vector-based models of semantic composition. In *Proceedings of ACL-08: HLT*, pages 236–244.

Alan Ritter, Colin Cherry, and William B Dolan. 2011. Data-driven response generation in social media. In *Proceedings of the 2011 Conference on Empirical Methods on Natural Language Processing (EMNLP)*, pages 583–593.

Vasile Rus and Mihai Lintean. 2012. A comparison of greedy and optimal assessment of natural language student input using word-to-word similarity metrics. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 157–162.

Iulian Vlad Serban, Sordoni Alessandro, Lowe Ryan, Charlin Laurent, Pineau Joelle, C. Courville Aaron, and Bengio Yoshua. 2017. A hierarchical latent variable encoder-decoder model for generating dialogues. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3295–3301.

Lifeng Shang, Zhengdong Lu, and Hang Li. 2015. Neural responding machine for short-text conversation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (ACL-IJCNLP)*, pages 1577–1586.

Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan. 2015. A neural network approach to context-sensitive generation of conversational responses. In *Proceedings of the 14th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 196–205.

Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems 27*, pages 3104–3112.

Oriol Vinyals and Quoc Le. 2015. A neural conversational model. *arXiv preprint arXiv:1506.05869*.

Bowen Wu, Baoxun Wang, and Hui Xue. 2016. Ranking responses oriented to conversational relevance in chat-bots. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, pages 652–662.

Lantao Yu, Weinan Zhang, and and Yong Yu Jun Wang. 2017. SeqGAN: Sequence generative adversarial nets with policy gradient. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, volume 31.

Ganbin Zhou, Ping Luo, Rongyu Cao, Fen Lin, Bo Chen, and Qing He. 2017. Mechanism-aware neural machine for dialogue response generation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3400–3407.

Xiangyang Zhou, Daxiang Dong, Hua Wu, Shiqi Zhao, Dianhai Yu, Hao Tian, Xuan Liu, and Rui Yan. 2016. Multi-view response selection for human-computer conversation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 372–381.