Journal of Machine Learning Research 22 (2021) 1-29 Submitted 1/20; Revised 3/21; Published 5/21

LocalGAN: Modeling Local Distributions for Adversarial Response Generation

Baoxun Wang ASULEWANG@TENCENT.COM
Zhen Xu ZENXU@TENCENT.COM
Platform & Content Group, Tencent Technology Co., Ltd
No.8 Dongbei wang west Road, Beijing, China 100080

Huan Zhang* ZHANGHUAN123@PKU.EDU.CN
Peking University
No.5 Yiheyuan Road, Beijing, China 100871

Kexin Qiu KQ2131@COLUMBIA.EDU
Columbia University
500 W 120th St No.510, New York, NY 10027, United States

Deyuan Zhang DYZHANG@SAU.EDU.CN
College of Artificial Intelligence, Shenyang Aerospace University
No.37 Daoyi South Avenue, Shenyang, China 110136

Chengjie Sun CJSUN@INSUN.HIT.EDU.CN
School of Computer Science & Technology, Harbin Institute of Technology
No.92 West Dazhi Street, Harbin, China 150001

Editor: Amos Storkey

Abstract

This paper presents a new methodology for modeling the local semantic distribution of responses to a given query in the human-conversation corpus, and on this basis, explores a specified adversarial learning mechanism for training Neural Response Generation (NRG) models to build conversational agents. Our investigation begins with the thorough discussions upon the objective function of general Generative Adversarial Nets (GAN) architectures, and the training instability problem is proved to be highly relative with the special local distributions of conversational corpora. Consequently, an energy function is employed to estimate the status of a local area restricted by the query and its responses in the semantic space, and the mathematical approximation of this energy-based distribution is finally found. Building on this foundation, a local distribution oriented objective is proposed and combined with the original objective, working as a hybrid loss for the adversarial training of response generation models, named as LocalGAN. Our experimental results demonstrate that the reasonable local distribution modeling of the query-response corpus is of great importance to adversarial NRG, and our proposed LocalGAN is promising for improving both the training stability and the quality of generated results.

Keywords: Neural Response Generation, Adversarial Learning, Local Distribution, Energy-based Distribution Modeling, Conversational Agents

∗. The work was done when Huan Zhang was an intern in PCG, Tencent.

1. Introduction

End-to-End generative conversational agents (a.k.a., generative Chat-bots) are believed to be practicable on the basis of the Sequence-to-Sequence (Seq2Seq) architecture (Sutskever et al., 2014) trained with large amounts of human-generated conversation sessions (Shang et al., 2015; Sordoni et al., 2015), and this task is named as Neural Response Generation (NRG). Similar to the Neural Machine Translation (NMT) approaches (Bahdanau et al., 2015; Wu et al., 2016), the NRG models are expected to directly generate appropriate and meaningful responses according to the input query. That is, such models formally aim at estimating the conditional probability $p(r|q)$, where q is the input query and r indicates the corresponding response. From the perspective of Seq2Seq, which is typically utilized in NMT, the estimation can be conducted via $\prod_{t=1}^{T_r} p(w_t^r|v, w_1^r, ..., w_{t-1}^r)$, where w_t^r is the t-th word of the T_r-word response $r = (w_1^r, ..., w_{T_r}^r)$ to be generated, and this operation is known as the decoding process. It should be noted that the variable v indicates the semantic representation obtained based on q, by performing the encoding procedure $p(v|q)$.

Compared to the success of NMT systems, the application progress of NRG models is not satisfying at present due to the "safe response" problem (Li et al., 2016). That is, most of the generated responses are boring and meaningless, which blocks the continuation of conversations. Indeed, eliminating "safe responses" is the essential task of NRG models. Thus, various methods have been considered to address this problem (Li et al., 2016; Xu et al., 2017; Li et al., 2017; Xing et al., 2017; Pandey et al., 2018; Zhang et al., 2018a; Du et al., 2018; Gao et al., 2019; Du and Black, 2019).

More recently, Generative Adversarial Nets (GAN) (Goodfellow et al., 2014) have been introduced to eliminate safe responses (Li et al., 2017; Xu et al., 2017, 2018; Zhang et al., 2018b; Zhu et al., 2019). Basically, this methodology is reasonable since the GAN framework involves an adversarial discriminator that helps NRG models leap out of the shortsighted state of minimizing the empirical risk on word distribution, by providing feedback on real samples from the model generated ones. Despite the improvement on the diversity, the adversarial training process of GAN based response generation models is generally unstable and sensitive to the training strategy (Yu et al., 2017).

The unstable convergence problem is largely ascribed to the complicated data distribution in practical scenarios (Arora et al., 2017; Arora and Zhang, 2018). For the response generation oriented GAN models, in particular, the data distribution appears to be much more complicated. Fundamentally, an essential characteristic of conversation data is that, for each given query, there always exists a group of semantically-diverse responses, rather than the semantically-unified ones. Furthermore, the response groups of two different queries tend to keep great divergences in the semantic space. In this scenario, the discriminator needs to consider the distributions of the generated result in the semantic space, rather than simply examining whether one single sample comes from the generator or the original dataset, so as to make the generator sense the distribution of conversation dataset in the adversarial training procedure.

This paper aims at presenting a specific adversarial training schema for neural response generation. Beginning with the investigation on the reason for the unsatisfying performance of GANs on the NRG task, we find the upper-bound of the current GAN learning strategies

taking query-response pairs as the independent training samples. On this basis, we claim that the training schema, including the adversarial strategy and the overall loss function, should be re-defined to agree with the distribution of NRG training samples in the semantic space, rather than roughly adopting the GAN framework designed for generating images.

Consequently, we describe the distributional state of the given query and the corresponding responses with the free energy defined on the basis of Deep Boltzmann Machines (DBM) (Salakhutdinov and Hinton, 2009). In this way, we can quantify the formation process of generating the response set with the topic restriction of the given query. From the perspective of free energy, this paper proposes a new cost function to measure the expansion degree of the responses in the local area of the real-valued semantic space. Cooperating with the traditional implicit density discriminating loss of GAN, the proposed cost actually provides an explicit density approximation for the local distribution of each response cluster. Thus, the adversarial learning procedure can be expected to be more stable with better response generation results obtained.

2. Related Work

Building deep neural nets with the ability of directly generating responses to a given query is of great significance to linguistic intelligence, and meanwhile, takes the tremendous challenge to the studies of conversational agents. Early models for Neural Response Generation (NRG) are inspired by neural machine translation architectures (Bahdanau et al., 2015; Wu et al., 2016). By adopting Sequence-to-Sequence (Seq2Seq) models (Sutskever et al., 2014), the query is encoded with a sequence model into a semantic vector, and the response is generated based on the vector via a sequence decoding procedure (Shang et al., 2015; Sordoni et al., 2015).

With the thorough analysis on generated results, it is widely accepted that classic NRG models are very likely to produce uninformative generic results with highly homogeneous patterns, called "safe responses" (Li et al., 2016; Xu et al., 2017), which makes the diversity of responses is even more difficult to be guaranteed than the relevance. A number of models have been explored to address the safe response problem, such as the ones based on diversity-oriented training goal (Li et al., 2016), topic/structure constrains (Xing et al., 2017; Du and Black, 2019), and variational auto-encoders (Du et al., 2018; Gao et al., 2019), etc.

The development of Generative Adversarial Nets (GAN) (Goodfellow et al., 2014) has brought a special perspective to address the safe response issue. Compared to the traditional methodologies, the NRG models with adversarial architectures are found to have the capability of leaping out of the safe-response status, by conducting a more aggressive optimization strategy. To adapt the adversarial response generation scenario, which is quite different from GAN-based image generation, some typical techniques are employed, such as policy gradient (Li et al., 2017), differentiable connection layer (Xu et al., 2017), specially designed optimization goal (Zhang et al., 2018b), etc.

Some very recent studies have further shown the potential of GAN-based NRG architectures. The informativeness and diversity of generated results are improved with various reasonable methods, including introducing multi-objective into the discriminator (Zhang et al., 2020), leveraging the information of future conversations (Feng et al., 2020), and adopting new generating policy (Zhu et al., 2020). The above studies indicate that, for

the NRG task, GAN is a promising and flexible architecture with effective variants to be designed. Meanwhile, it can be seen that, currently, little attention has been paid on investigating the local distribution of the responses to a given query and its influence to the GAN-based NRG model, which is the very focus of our work.

3. The Limitation of General GAN in the NRG Scenario

According to (Goodfellow et al., 2014), the standard GAN framework contains a generator G and a discriminator D, which are trained by an iterative adversarial learning procedure based on the following objective function:

$$J^{(D)} = \mathbb{E}_{x\sim p_d}\left[\log D(x)\right] + \mathbb{E}_{z\sim p_z}\left[\log(1 - D(G(z)))\right] \tag{1}$$

$$J^{(G)} = \mathbb{E}_{z\sim p_z}\left[\log D(G(z))\right] \tag{2}$$

where p_z denotes the prior on input noise variables, and p_d is the true data distribution.

It should be noted that the GAN tries to learn the manifold of a given dataset (Khayatkhoei et al., 2018; Kumar et al., 2017), and the discriminator D actually provides a metric for judging whether the results generated according to z fits the expected manifold or not. In the NRG scenario, a naive Seq2Seq model without the guidance signal from D can not capture the data manifold of the real query-response corpus, which is one of the major facts the safe-response problem can be ascribed to. Assuming that there exists an oracle discriminator with the ability of distinguishing the generated fake samples from the ground-truth ones, by mapping each query-response pair (q, r)[1] to a confidence score s, and meanwhile, it can be assume that any practically existing discriminator of GAN gives the confidence $\tilde{s}$ to (q, r). If the practical discriminator can make $\tilde{s} \to s$, the generator will obtain more meaningful guidance for the better generation. That is, to improve the capability of GAN-NRG, it is wise to construct more powerful discriminators for more reasonable $J^{(G)}$.

Now let's pay attention to the actual change of NRG models brought by GAN. In the generative conversation agent scenario, $G(z)$ is corresponding to a generated response $\tilde{r}$ to a given query q. Thus, the objective of the generator in the GAN based NRG model can be simply formulated as:

$$J^{(G)} = \mathbb{E}_{(q,\tilde{r})\sim p_g}\left[\log D(q, \tilde{r})\right] \tag{3}$$

To generate realistic responses according to given queries, the training of the generator is actually the procedure to maximize $J^{(G)}$ toward $\mathbb{E}_{(q,r)\sim p_d}\left[\log D(q, r)\right]$. Thus, in the context of adversarial learning, $J^{(G)}$ should satisfy the following inequality:

$$J^{(G)} \leqslant \mathbb{E}_{(q,r)\sim p_d}\left[\log D(q, r)\right] \tag{4}$$

However, it is well known that a conversational dataset should not be simply taken as a collection $\{(q, r)\}$ composed of independent query-response pairs. Instead, to each given query q, there exists a finite set of corresponding responses $R_q = \{r_i\}$. In this case, it is of

1. Here q and r represent the vectorized query and its response. We take the simple embedding-averaging based method to transform texts into vectors. Besides, due to our adopted text vectorizing method, the pre-training phase of the Deep Boltzmann Machine takes some specified trick detailed in Appendix A.

great necessity to consider the whole training dataset as a collection of R_q, which takes the form of a number of clusters with their own local distributions in the semantic space. And we can rewrite the joint distribution p_d in (4) as $p(q)p(r|q)$ and assume every corresponding response to a query follows equal-probability distribution[2], which means that $p(r|q) = \frac{1}{|R_q|}$. Thus, in the real NRG scenario, on the basis of Inequality 4, $J^{(G)}$ follows the inequality as below:

$$\begin{aligned} J^{(G)} &\leqslant \sum_{(q,r)} \mathbb{E}_{(q,r)\sim q,R_q} \left[\log D(q,r)\right] \\ &= \sum_{q} \sum_{r\in R_q} p(q) \frac{1}{|R_q|} \left[\log D(q,r)\right] \leqslant \sum_{q} p(q) \log \left[\frac{1}{|R_q|} \sum_{r\in R_q} D(q,r)\right] \end{aligned} \tag{5}$$

where $p(q)$ denotes the probability of the query q, and R_q is defined above.

According to Equation 5, the upper bound of $J^{(G)}$ is obtained, in which the essence is the $\log\left[\frac{1}{|R_q|}\sum_{r\in R_q} D(q,r)\right]$ part. Apparently, the expression $\frac{1}{|R_q|}\sum_{r\in R_q} D(q,r)$ indicates the mean value of the confidence scores given by the discriminator to each member of the response set R_q to a given query q. Moreover, it should be noted that current studies tend to utilize semantic relevance oriented models to build the discriminators of GAN (Xu et al., 2017; Li et al., 2017). Consequently, $D(q,r)$ can be actually considered as the spatial relationship of q and r in the semantic space. In the situation that the discriminator D is continuous, which is a generic precondition of the NRG oriented GAN frameworks discussed in this work, the scalar $\frac{1}{|R_q|}\sum_{r\in R_q} D(q,r)$ actually corresponds to a vector representing the semantic center of R_q. That is, the optimization process of adversarial learning upon conversational datasets will make the generated responses approach to the center of each local distribution of R_q to each given dependent query.

The practical value of this change lies in that, intuitively, the GAN architecture forces the generator to pay attention to the local distributions of the individual response clusters, rather than taking the (q,r)-pairs as an entirety. According to the thorough studies on the safe responses of NRG models (Li et al., 2016; Xu et al., 2017; Zhang et al., 2018a; Pandey et al., 2018), it can be inferred that the general Seq2Seq will fall into the divergence state of generating the patterns with the maximum probabilities taking account of the entire dataset, ignoring the individual-difference of each query. By introducing the implicit loss focusing on the response clusters, GAN makes the divergence of the generator much closer to the 'local patterns' rather than the general patterns, and thus the higher diversity can be expected and observed (Xu et al., 2017; Li et al., 2017).

The problem turns to: Is the upper bound in Equation 5 powerful enough? Apparently, there exists an obvious gap between the 'local patterns' and the responses with the satisfying informativeness and diversity. The upper bound only focuses on the mean of the outputs of given by the discriminator. That is, on the basis of Equation 5, the classic GAN only pays attention to the semantic center of each response set as mentioned above, but the local distribution (or the actual "shape") of each cluster has not been taken into account.

2. Actually, the equal-probability distribution of responses corresponds to the common sense that, without any prior hypotheses, all the responses to a given query share the same probability.

This situation does not change when the cost function of adversarial training is defined by Wasserstein GANs (WGAN) (Arjovsky et al., 2017) with 1-Lipschitz function f:

$$J_W^{(G)} = \mathbb{E}_{(q,\tilde{r})\sim p_g}\left[f(q,\tilde{r})\right] \tag{6}$$

Intuitively, it is of paramount importance to estimate both the "location" and the "shape" of the response set R_q in the semantic space (indeed, $\frac{1}{|R_q|}\sum_{q,r} D(q,r)$ is only relative with "location"), so as to determine the optimization objective of adversarial training. Consequently, we have two critical problems to be discussed and addressed in the following sections:

- How to describe the state of the response set R_q with a given query q in the semantic space?
- Taking account of the reasonable state modeling of (q, R_q), what is the loss function for adversarial training to generate responses?

4. Modeling the State of the Local Distribution for Responses

As mentioned above, the semantic one-to-many relationship between queries and responses makes it necessary to model the local distribution of the response cluster R_q corresponding to each query, and it is paramount to turn to the fitting of each local distribution in the adversarial learning procedure, rather than considering each (q, r)-pair as an independent sample. Basically, this issue equals to the task of reasonably modeling the state of (q, R_q) in the semantic space, by considering each (q, R_q) as a systematic entirety and assigning the state of the entirety with probabilistic distribution. The additional major challenge of this task is, indeed, we have to infer the state of a local area in the semantic space from a group of finite samples, since it is impossible to sample all the possible responses to a query from the given corpus, regardless of the corpus size.

4.1 Representing Local Distributions with Query-Response Oriented Free Energy

In this part, we typically take an energy based statistical model, the Average Free Energy (Hinton and Zemel, 1994; Friston et al., 2006; Ngiam et al., 2011; Friston et al., 2012), to describe the state of the local distribution of (q, R_q) in the semantic space, for the reasons that: a) energy based models are considered as a promising avenue towards learning explicit generative models (LeCun et al., 2006; Le Roux and Bengio, 2008), by representing data distributions without any prior assumptions; and b) energy based models can be trained in the unsupervised way, and the energy functions of such models have the potential to estimate the state of generative models (Zhao et al., 2017).

At first, the free energy of a given query-response pair (q, r) can be defined as following:

$$F(q,r) = -\log\left[\sum_H \exp(-E(q,r,H))\right] \tag{7}$$

where $E(q, r, H)$ stands for the energy function defined according to the relationship of the query q and its response r via the hidden variable H.

We employ the Deep Boltzmann Machine (DBM) (Salakhutdinov and Larochelle, 2010; Smolensky, 1986; Hinton and Salakhutdinov, 2006) to implement $E(q, r, H)$, as illustrated by Figure 1. The reason for this choice lies in that, from the view of conversational agents, the meaningful query and the corresponding response are generally considered to maintain strong semantic relevance. Thus, the query and response can be mutually transformed into each other, which is supported by the considerable amount of studies on response generation (Shang et al., 2015; Shao et al., 2017; Zhang et al., 2018a; Baheti et al., 2018) and question generation (Du et al., 2017; Zhao et al., 2018; Sun et al., 2018). Without loss of generality, the pairwise semantic relationship of the query q and the corresponding response r can be modeled by a two-layer DBM. The bottom layer is actually an abstract version of Seq2Seq models, in which a response r can be generated based on a hidden variable h_q, and h_q depends on the given query q theoretically. In the top layer, h_q conditionally depends on a hyper hidden variable h. Figure 1 illustrates the Deep Boltzmann Machine for modeling the semantic relationship of a query and its responses.

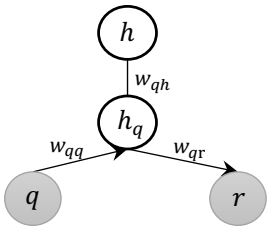

Figure 1: The DBM for modeling the semantic relationship of Query-Response pairs.

Following (Salakhutdinov and Hinton, 2009), on the basis of the DBM in Figure 1, the energy of the state $\{(q, r), H\}$ is defined as:

$$E(q, r, H) = E\left(q, r, h_q, h\right) = -h_q^T W_{qr} r - h_q^T W_{qq} q - h^T W_{qh} h_q \tag{8}$$

where q denotes the query, r stands for the response and $H = \{h_q, h\}$ represents the hidden units. W_{qr}, W_{qq} and W_{qh} stand for the weights on the corresponding connections of the query, response and the hidden variables respectively in the graph model shown by Figure 1.

Consequently, we can define the average free energy of the query q and its response set R_q as follows:

$$F(q, R_q) = \frac{1}{|R_q|} \sum_{r_i \in R_q} F(q, r_i) \tag{9}$$

For better conducting the following discussion, we further define the energy difference between response r_i and r_j as:

Definition 4.1 (Scaled Energy Difference)

$$\Delta_{q, r_i, r_j} = \frac{F(q, r_i) - F(q, r_j)}{F(q, R_q)} \tag{10}$$

Meanwhile, it is necessary to assign a spatial intuition to R_q in the semantic space by defining:

Definition 4.2 (Response Cluster) *In the semantic space, the meaningful responses to the given query q lie in a restricted region (e.g., a hyper sphere), which can be named as the **Response Cluster**, in which $F(q, r)$ can be taken as the distance from a response r to the cluster center.*

4.2 Estimation of $F(q, R_q)$

Basically, the DBM in Figure 1 provides the definition of the energy function $E(q, r, H)$ of the free energy given in Equation 8. Thus, the local distribution state of the responses R_q with the given query q can be mathematically described by $F(q, R_q)$ based on Equation 7 - 9. In practice, however, this procedure is not operable yet because the computation of $F(q, R_q)$ requires all the response in R_q, and it is impractical to perform the exhaustive enumeration over all the possible responses of the given query, regardless of the amount of the training query-response pairs. Consequently, it is highly necessary to approximate $F(q, R_q)$ under some reasonable assumptions.

Considering the response cluster defined in Definition 4.2, we can naturally assume that the response random variable r to a given query q follows multivariate normal distribution with mean r_c and covariance matrix Σ, where r_c and Σ is only determined by query q. Afterwards, the observed R_q can be considered as a realization of the $|R_q|$-variate random variable $(r_1, \cdots, r_{|R_q|})$ for $r_i, 1 \leq i \leq |R_q|$, i.i.d random variables drawn from the distribution $\mathcal{N}(r_c, \Sigma)$. Consequently, an executable approximation of $F(q, R_q)$ can be obtained based on the following Lemma 1 and Theorem 1[3].

Lemma 1 *Given the free energy $F(q, r)$ defined in Equation 7, we have*

$$\lim_{\mathbb{E}\|r-r_c\|_2 \to 0} |\mathbb{E}[F(q, r)] - F(q, r_c)| = 0 \tag{11}$$

Practically, the expected Euclidean distance between random variable r and r_c can not be zero. Thus, the fact conveyed by Lemma 1 is that, actually, if the expected Euclidean distance is small enough, the difference between $\mathbb{E}[F(q, r)]$ and $F(q, r_c)$ can be controllable (or even close to zero). Based on the lemma we have:

Theorem 1 *Given $\mathbb{E}[F(q, r)] < \infty$, we have*

$$|F(q, R_q) - F(q, \hat{r}_c)| \xrightarrow{a.s.} 0 \qquad \textit{when } |R_q| \to \infty \textit{ and } \mathbb{E}\|r - r_c\|_2 \to 0 \tag{12}$$

where $\hat{r}_c$ is the estimation of r_c based on the well-trained DBM.

Proof

$$|F(q, R_q) - F(q, \hat{r}_c)| \leq |F(q, R_q) - \mathbb{E}[F(q, r)]| + |\mathbb{E}[F(q, r)] - F(q, r_c)| + |F(q, r_c) - F(q, \hat{r}_c)| \tag{13}$$

According to the strong law of large numbers, when $|R_q|$ goes to infinity, the sample average $F(q, R_q) = \frac{1}{|R_q|} \sum_{r_i \in R_q} F(q, r_i)$ converges to the expected value $\mathbb{E}[F(q, r)]$.

Following Lemma 1, $|\mathbb{E}[F(q, r)] - F(q, r_c)|$ goes to zero when $\mathbb{E}\|r - r_c\|_2 \to 0$.

As discussed in (Wang et al., 2010; Srivastava and Salakhutdinov, 2012), a DBM sufficiently trained with large amounts of (q, r) pairs actually guarantees the estimation of any response r_i (denoted by $\hat{r}_i$), and thus the statement "well-trained" indicates that the estimate $\hat{r}_c$ is very close to the parameter r_c. Since the function $F(q, r)$ is continuous, the difference $|F(q, r_c) - F(q, \hat{r}_c)| \to 0$ when $||r_c - \hat{r}_c||_2 \to 0$.

3. The detailed proof of Lemma 1 is given in Appendix B. Moreover, based on a different assumption, we also provide an intuitive and simple proof of Theorem 1 in Appendix C, which does not depend on any lemma.

To conclude, $|F(q, R_q) - F(q, \hat{r}_c)| \to 0$ when the conditions mentioned in Theorem 1 are satisfied. ■

5. The Hybrid Loss of Adversarial Response Generation

As discussed in Section 3, the ability of the response generator in the general GAN architecture is limited to learning the dense distribution around $\frac{1}{|R_q|}\sum_{r\in R_q} D(q, r)$ (see Equation 5), which is composed of the most frequent patterns in the semantic space. By contrast, it is difficult for the general architecture to sense the remaining sparse space containing high-quality diverse responses. Therefore, reasonably describing the local distribution of the responses to a given query is highly necessary. According to the analysis in Section 4, the average free energy can be taken to model the state of the local area of the responses to a query, and such energy can be reasonably approximated via the DBM defined on the query-response pairs. On the basis of the previous sections, this section will finally propose the new hybrid loss function to force the generator to produce responses with better diversity through the more stable adversarial training process.

5.1 The Radial Distribution Function of the Response

The analysis in Section 4 have shown that the local distribution state of the responses R_q to the given query q can be modeled by the average free energy $F(q, R_q)$. On this basis, it is possible to propose the description of the spatial state of a single response r in the semantic space, and consequently, we can give a new adversarial loss indicating the cost of simulating the local distribution of R_q.

According to Definition 4.1 and 4.2, in each response cluster R_q, the distance from a response r to the cluster center r_c is actually equivalent to the scaled energy difference between them, that is,

$$\Delta_{q,r,r_c} = \frac{F(q, r) - F(q, r_c)}{F(q, R_q)} \tag{14}$$

Meanwhile, on the basis of Theorem 1, $F(q, R_q)$ can be approximated by $F(q, \hat{r}_c)$, and $\hat{r}_c$ is modeled from training data, and thus we have:

$$\Delta_{q,r,r_c} \approx \frac{F(q, r) - F(q, \hat{r}_c)}{F(q, \hat{r}_c)} = \alpha_{(q,r)} \tag{15}$$

Here we approximate Δ_{q,r,r_c} with $\alpha_{(q,r)}$, and formally call $\alpha_{(q,r)}$ as the Radial Distribution Function (RDF), indicating the relative cost ratio to $F(q, r_c)$ for obtaining r from a given q (also the distinctiveness of r, actually).

5.2 The Hybrid Objective Function

Based on the previous discussions, for the adversarial response generation methodology, the essence is to reasonably describe the state of the local distribution of the response cluster given by Definition 4.2, and further more, to take this important element into account in the final optimization.

Especially, in Subsection 5.1, we have defined the Radial Distribution Function in Equation 15 to quantify the distinctiveness of a response, the very basis of which is the description of the local state $F(q, R_q)$ in the semantic space. Thus, we can further build a mechanism to quantify the difference between the generated response and the golden response as follows:

$$\delta\alpha = \alpha_{(q,r)} - \alpha_{(q,\tilde{r})} \tag{16}$$

where $\tilde{r}$ is the generated response given by the generator and r comes from the original data. If $\delta\alpha$ moves toward zero, $\alpha_{(q,\tilde{r})}$ would be close to $\alpha_{(q,r)}$ sharing the same $F(q, r_c)$.

Consequently, a new expectation comes out. That is, the generator needs to provide results that can minimize $\delta\alpha$, so as to fit the local distribution of the existing responses to a given query. Thus, a hybrid objective of the generator can be finally defined as:

$$\min J^G = -\mathbb{E}\left[\log D(q, r)\right] + ReLU(\delta\alpha) \tag{17}$$

A hinge loss, conducted by the $ReLU$ function $ReLU(\delta\alpha) = \max(0,\ \delta\alpha)$, is especially introduced to reform $\delta\alpha$. The primary reason of this operation is that the $ReLU$ function has positive output only if $\delta\alpha \geqslant 0$, according to the definition of $ReLU(\delta\alpha)$. Apparently, $\delta\alpha < 0$ indicates that the generated response $\tilde{r}$ is too far from the center of the response cluster in the semantic space, so that its relevance may be highly questionable. Meanwhile, minimizing a negative variable is against the optimization direction. After the ReLU transformation, there remains valid loss only when $\delta\alpha \geqslant 0$, and thus both the diversity and the relevance of generated results are taken into account. On the basis of the objective, Algorithm 1 details the training process of LocalGAN[4].

Algorithm 1: The Training of LocalGAN

Data: A dialogue set $\mathcal{S} = \{(q_i, r_i)\}_{i=1}^{|\mathcal{S}|}$

Pre-train the generator G with $\mathcal{S}$;

Pre-train the discriminator D with positive samples (q_i, r_i) and negative samples $(q_i, G(q_i))$, $i = 1, ..., |\mathcal{S}|$;

while *not convergence* **do**

  for $\mathcal{S}$ *has unsampled batches* **do**

    Sample a batch of K instances $\{(q_j, r_j)\}_{j=1}^{K}$ from $\mathcal{S}$;

    Generate $\tilde{r}_j = G(q_j)$;

    Update θ_D by the gradient descent on the discriminator loss $\mathcal{L}^D = \frac{1}{K}\sum_{j=1}^{K}[\log D(q_j, r_j) + \log(1 - D(q_j, \tilde{r}_j))]$;

    Generate $r_c = DBM(q_j, r_j)$;

    Generate $\tilde{r}_c = DBM(q_j, \tilde{r}_j)$;

    Estimate RDF $\alpha_{(q_j,r_j)}$ and $\alpha_{(q_j,\tilde{r}_j)}$ with Equation 15 ;

    Update θ_G by the gradient descent of minimizing the hybrid loss $\mathcal{L}^G = \frac{1}{K}\sum_{j=1}^{K}\left[\log D(q_j, r_j) + ReLU(\delta\alpha_j)\right]$;

  end

end

4. The code of our proposed model LocalGAN can be found in `https://github.com/Kramgasse49/local_gan_4generation`

5.3 The Phase-wise Intuition of the Optimization

According to the analysis in Section 3, the trivial adversarial training directed by $-\mathbb{E}\left[\log D(q, r)\right]$ can only determine the form of general responses to a given query. From the spatial perspective in the semantic space, the original adversarial objective is helpful to roughly locate the response cluster to be generated. However, the local distribution can not be captured by this procedure.

By contrast, according to the discussions above, the proposed hybrid objective in Equation 17 actually provides a way to force the generated responses, originally gathering around the general form, to expand into the expected local shape described by the golden truth. The whole optimization procedure can be detailed in an intuitive way:

Foundation: Once the DBM in Figure 1 is well-trained with the query-response corpus, the semantic center r_c of a Response Cluster can be determined by the given query q.

Phase-1: In the early stage of the adversarial training, a generated response $\tilde{r}$ is not semantically relevant to the query q. Thus, it can be inferred that $\tilde{r}$ is radially farther from the cluster center r_c than the golden response r. In this situation, according to Equation 15 and Equation 16, we can claim that $\delta\alpha \leq 0$. In this phase, the hyper objective goes back to the general adversarial objective due to the ReLU function. Thus, the model is trying to force the generated samples to approach the center of each cluster, ignoring local distributions.

Phase-2: During the adversarial training in *Phase-1*, the generated result $\tilde{r}$ will go approaching to the cluster center r_c, which means $\alpha_{(q,\tilde{r})} \to 0$. It should be noted that, for any meaningful existing training sample r, $\alpha_{(q,r)} > 0$. Therefore, at some point, it turns to $\delta\alpha > 0$ and the right part of the hybrid objective in Equation 17 takes effect. Consequently, for each given query, the distribution of the generated results will expand to fit the local distribution of the golden samples.

The DBM is pre-trained on the query-response corpus independently without supervision. Besides, phase-1 and phase-2 are actually expected behaviors occurring during the optimization, and thus they are not performed sequentially.

6. Experiments

This section gives the experimental results of our proposed LocalGAN, which are analyzed and compared to those of the baseline models on the widely applied metrics.

6.1 Experimental Setups

Datasets. Our experiments are conducted on two main stream open-access conversation corpora: The Opensubtitles corpus and the Sina Weibo corpus. The OpenSubtitles dataset contains 5,200,000 movie dialogues, where we extract query-response pairs following (Xu et al., 2018; Li et al., 2016). The Sina Weibo Corpus (Shang et al., 2015) contains 2,500,000 single-turn Chinese dialogues, in which the length of the query and response ranges from 4 to 30. We sample 100,000, and 2,000 unique query-response pairs as validation and testing dataset respectively from both of the corpora[5].

5. Both the English and the Chinese datasets used in our experiments are uploaded to `https://www.dropbox.com/sh/k8i079gd2111lsb/AACLLtlNAzile543Da8Qs9tFa?dl=0`.

Baselines. For meaningful comparison, we introduce the following models as baselines:

(1) ***Seq2Seq***: a Seq2Seq model trained with maximum likelihood estimation (MLE).

(2) ***Seq2Seq-MMI***: the NRG model with a Maximum Mutual Information criterion (Li et al., 2016).

(3) ***Adver-REGS***: the NRG model trained using adversarial framework, in which the policy gradient was employed to transfer the reward of the discriminator to the generator (Li et al., 2017).

(4) ***GAN-AEL***: an adversarial framework with an approximate embedding layer for connecting the generator with the discriminator directly (Xu et al., 2017).

(5) ***AIM / DAIM***: the adversarial training strategy allowing distributional matching of synthetic and real responses, and explicitly optimizing a variational lower bound on pairwise mutual information between the query and response, so as to improve the informativeness and diversity of generated responses (Zhang et al., 2018b)[6].

(6) ***BigGAN***: According to (Brock et al., 2019), a group of operations upon GANs, including enlarging the batch size (from 128 to 1024), doubling the parameter number of the entire model, adopting orthogonal regularization, etc., are very helpful for improving the quality of generated images. For the comparison of response generation, we apply the operations of BigGANs to AdverREGS to build a new baseline model.

(7) ***LocalGAN-SE***: To further illustrate the effect of the local distribution modeling, a simpler alternative, named as LocalGAN-SE (SE is short for Simpler Edition), is designed and employed as a baseline. Basically, LocalGAN-SE intuitively takes the ratio $\frac{d(q,r)-d(q,\hat{r}_c)}{d(q,\hat{r}_c)}$ to compute $\alpha_{(q,r)}$ in Equation 15, where $d(q,r)$ is the Euclidean distance between q and r in the semantic space, replacing the energy function F. Apparently, this baseline actually models the variance of each r in the response cluster as LocalGAN does.

Evaluation Metrics. To evaluate the diversity, we adopt three widely-applied metrics: Distinct-1 (**Dist-1**), Distinct-2 (**Dist-2**), and Entropy (**Ent4**) (Li et al., 2016; Zhang et al., 2018b; Jost, 2006). Besides, the relevance (**Rel.**) is measured by summing three embedding-based similarities (greedy, average, extreme) (Liu et al., 2016) upon the ground-truth and generated responses, which are verified to be more coherent with the human evaluation on relevance(Wu et al., 2019; Zhou et al., 2019; Ghandeharioun et al., 2019).

Training Details. The vocabulary size of both datasets is 40,000. The embedding layer of OpenSubtitles and Sina Weibo is initialized using 200-dimensional Glove vectors (Pennington et al., 2014) and 300-dimensional Weibo vectors (Li et al., 2018) respectively. All the models are first pre-trained by MLE, and then the models including Adver-REGS, GAN-AEL, AIM, DAIM, BigGAN, LocalGAN-SE and LocalGAN are trained with adversarial learning. The discriminator of Adver-REGS and GAN-AEL are based on CNN following (Yu et al., 2017; Xu et al., 2017), in which the filter sizes are set to (1,2,3,4) and the filter number is 128, while that of LocalGAN adopts DBM with (2×embedding_size, 128, 128) to represent the semantic of queries and responses. The hidden size of the generator is set to 256 and 512 in GAN-based models and Seq2Seq respectively. To guarantee the performance consistency of AIM and DAIM, we adopt the recommended parameter settings given by Zhang et al. (2018b). The experiments are conducted on the Tesla K80 GPU.

6. We have taken the codes of AIM and DAIM from `https://github.com/dreasysnail/converse_GAN` implemented by the authors of this work for comparisons.

Model	Dist-1	Dist-2	Ent4	Rel.
Seq2Seq	0.025 (-)	0.081 (-)	5.650 (-)	1.090 (-)
Seq2Seq-MMI	0.027 (-)	0.086 (-)	5.698 (-)	1.067 (-)
Adver-REGS	0.0296(±0.0012)	0.098(±0.0027)	5.701(±0.086)	1.113(±0.029)
GAN-AEL	0.030(±0.003)	0.100(±0.0046)	5.733(±0.183)	1.106(±0.0.067)
AIM	0.0292(±0.0019)	0.095 (±0.0040)	5.783(±0.135)	1.120(±0.0.050)
DAIM	0.031(±0.0017)	0.103(±0.0039)	5.873(±0.124)	1.098(±0.0423)
BigGAN	0.0336(±0.0004)	0.108(±0.0018)	6.020(±0.047)	1.117(±0.013)
LocalGAN-SE	0.0314(±0.0013)	0.104(±0.0032)	5.922(±0.115)	1.103(±0.0372)
LocalGAN	**0.036**(±0.0006)	**0.110**(±0.0024)	**6.073**(±0.068)	**1.132**(±0.019)

Table 1: Performances of LocalGAN and Baselines on the Opensubtitles Datasets.

Model	Dist-1	Dist-2	Ent4	Rel.
Seq2Seq	0.055 (-)	0.153 (-)	6.400 (-)	0.315 (-)
Seq2Seq-MMI	0.059 (-)	0.172 (-)	6.860 (-)	0.309 (-)
Adver-REGS	0.061(±0.0019)	0.181(±0.0037)	7.658(±0.1850)	0.320(±0.0128)
GAN-AEL	0.062(±0.0048)	0.183(±0.0152)	7.765(±0.6512)	0.318(±0.0317)
AIM	0.064(±0.0028)	0.189(±0.0074)	7.833(±0.2945)	0.321(±0.0244)
DAIM	0.067(±0.0023)	0.195(±0.0067)	8.042(±0.2262)	0.316(±0.0199)
BigGAN	0.0685(±0.0007)	0.204(±0.0031)	8.273(±0.0847)	0.321(±0.0052)
LocalGAN-SE	0.0678(±0.0018)	0.201(±0.0061)	8.134(±0.1764)	0.311(±0.0145)
LocalGAN	**0.071**(±0.0011)	**0.212**(±0.0057)	**8.561**(±0.1279)	**0.327**(±0.0066)

Table 2: Performances of LocalGAN and Baselines on the Weibo Datasets.

6.2 Results & Analysis

Table 1 and Table 2 list the best quantitative results on the diversity and relevance of responses generated by all the models on the Opensubtitles and Weibo dataset respectively. As shown by the results, compared to Seq2Seq and Seq2Seq-MMI, the GAN-based methods give better results on the diversity oriented metrics, including Dist-1, Dist-2 and Ent4. This observation indicates that adversarial learning does provide the meaningful guidance to NRG models to avoid some of the safe-responses.

It can be observed that LocalGAN outperforms the baselines with adversarial learning architecture (Adver-REGS, GAN-AEL, AIM/DAIM, and BigGAN) on both the diversity metrics and the relevance metrics. Generally, a notable improvement on diversity may lead to some negative influence on relevance, and thus promoting the diversity of generated response while maintaining their relevance is essentially desired for any methodologies, which has been achieved by our LocalGAN. The performances of LocalGAN can be attributed to the fact that LocalGAN has taken the local distribution of responses to a given query into account. By adopting the hybrid objective function, the proposed adversarial model gets to capture the spatial characteristics of response clusters, and the generation process is consequently forced to fit the semantic distributions of response clusters. Especially, we can see that LocalGAN-SE has outperformed the other baselines on some of the metrics with the comparable stability, even though this model takes a rather simple way to depict the

shape of local distributions. This observation indicates that modeling the local distribution of response clusters is potentially valuable.

Broadly, the safe response problem can also be considered as a special type of mode collapse problem in the NRG scenario, even though the safe response problem exist with no dependence on GANs. According to Table 1 and 2, it can be observed that compared to Seq2Seq models, the baselines with the GAN architecture get notable improvements on the diversity oriented metrics, which shows the capability of GANs on reducing safe responses. However, as discussed in Section 3, the responses generated by general GAN architectures are limited by the local patterns of each response cluster, and thus the response quality can be further improved. LocalGAN is able to fit the local distribution of each response cluster and extend the semantics of generated results on the basis of the local patterns. Thus, the improvements on diversity are improved with the relevance is also guaranteed.

The training stability is a tough issue to be addressed for adversarial learning (Yu et al., 2017), and as discussed in the previous sections, one of the motivations of our LocalGAN is to make adversarial learning more stable. To verify this aspect, we first obtain the standard deviations of all the adversarial learning based models' best results, including ours, on each metric (given in parentheses in Table 1 and 2), by independently repeating the training process for 10 times on them. It can be observed that, compared to the adversarial baselines, our proposed LocalGAN has the lowest standard deviations, indicating that the stability of LocalGAN is better than the others.

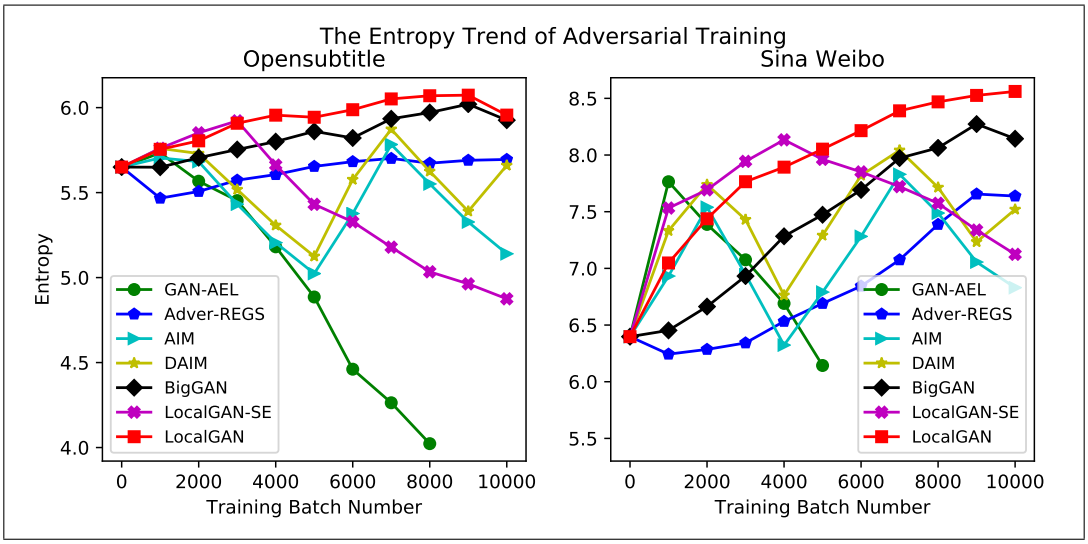

Figure 2: The Entropy Trend of adversarial learning based models in the Training Process.

In addition to the comparisons of the standard deviations, it is also necessary to observe the stability of models with the training epoch increasing. Thus, we track the changing of the Entropy (Ent4) of results given by GAN-AEL, Adver-REGS, AIM, DAIM, BigGAN, LocalGAN-SE, and LocalGAN, as shown in Figure 2. It can be observed that the training procedures of LocalGAN, Adver-REGS and BigGAN are relatively stable. By contrast, we can see that there exist obvious fluctuations on the curves of AIM and DAIM, and

GAN-AEL rapidly gets out of control after 1000 batch. For AIM, DAIM and GAN-AEL, it is rather difficult to grasp the best status. This group of results indicates the necessity of introducing additional restrictions into adversarial learning processes. For this purpose, Adver-REGS introduces a teacher-forcing loss (Li et al., 2017), while AIM and DAIM have taken the informativeness oriented constraints to partially control the stability (Zhang et al., 2018b). However, GAN-AEL only takes the Wasserstein distance as the objective (Xu et al., 2017), and thus the entropy goes down rapidly. Among the baselines, BigGAN has the best stability. Meanwhile, BigGAN actually outperforms the other baseline models according to Table 1 and 2. This observation shows the modifications of BigGAN are valuable for enhancing the generation in practice, and especially, scaling up the batch size is helpful for improving the stability. Compared to Adver-REGS and BigGAN, our LocalGAN achieves better diversity with even a more smooth entropy curve. The training of LocalGAN benefits from the phase-wise optimization driven by the hybrid loss, and its stability also indicates the meaningfulness of modeling and utilizing local distributions of responses. As a simpler alternative of LocalGAN, the baseline LocalGAN-SE has a good trend similar with that of LocalGAN at the early stage, but the curves go down with the training batch number increasing. This observation indicates that modeling the local variance is promising, however, the local distribution can not be reasonably depicted by the simple spatial relationship.

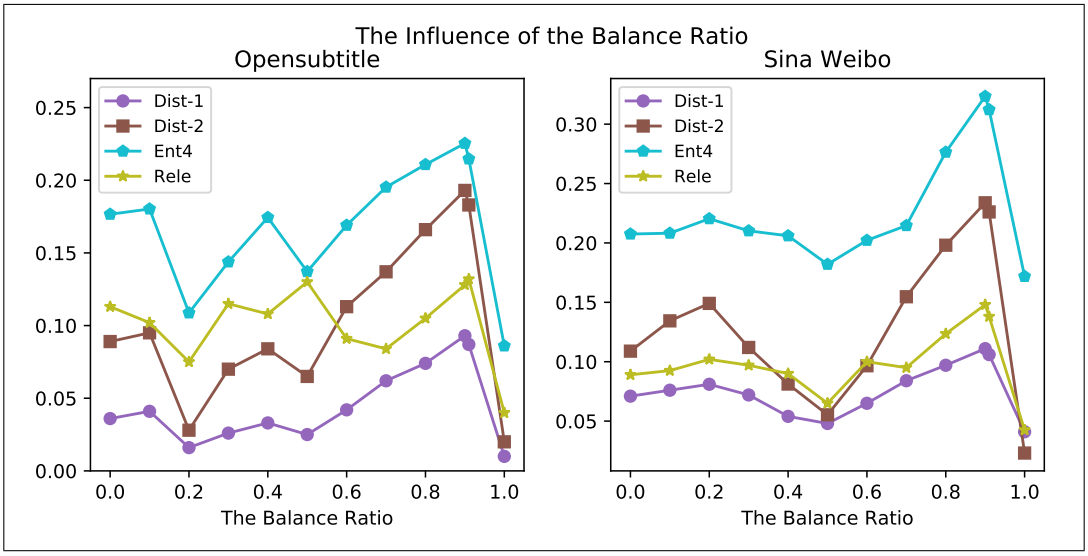

Figure 3: The Influence of the Balancing Ratio for the two terms in the hybrid objective.

According to Equation 17, the final hybrid objective of the generator consists of two parts, and it is necessary to investigate whether the proportion of the two terms will influence the results of our model. For this purpose, a balancing ratio β is introduced as a hyper-parameter, and the objective can be reformed as $-(1-\beta)\mathbb{E}\left[\log D(q,r)\right] + \beta ReLU(\delta\alpha)$, where $\beta \in [0,1]$[7]. By setting the interval with 0.1, we get the performance curves of our

7. Actually, this balancing ratio originally works in Equation 17, since we have observed the scales of the two terms and set the balancing ratio to about 0.91.

model on the metrics Dist-1, Dist-2, Ent4 and Relevance in Figure 3. From this figure, it can be observed that our model's performances on all the metrics get the highest point when $\beta = 0.9$. This trend can be partially attributed to the fact that, approximately, the value of $\mathbb{E}[\log D(q, r)]$ is 10 times larger than that of $ReLU(\delta\alpha)$, and thus β actually works as a scaling factor balancing the two terms. Moreover, when $\beta = 0$, the hybrid objective regresses to the classic GAN objective, and the considerable improvements can be expected. Meanwhile, all the curves drops significantly when $\beta = 1.0$, which shows that $ReLU(\delta\alpha)$ can not work alone as the optimization objective.

6.3 Human Evaluation

Method	Relevance	Informativeness
Seq2Seq	0.738	0.25
Seq2Seq-MMI	0.67	0.336
Adver-REGS	0.702	0.398
GAN-AEL	0.696	0.41
AIM	0.746	0.294
DAIM	0.768	0.45
BigGAN	0.763	0.487
LocalGAN-SE	0.751	0.437
LocalGAN	**0.784**	**0.536**

Table 3: Human Evaluation Results of Models.

To further conduct intuitive comparisons among the NRG models, we perform human evaluations on 500 testing samples. Five annotators are asked to judge whether a response is relevant to the given query and whether the response is informative or not respectively. Both human metrics "Relevance" and "Informativeness" solely have two labels 0 and 1 (0: irrelevant or non-informative; 1: relevant or informative) for simplicity. The human evaluation results are listed in Table 3, and the results basically match the observations in Table 1 and 2. The Kappa of relevance and informativeness annotations are 0.734 and 0.561 respectively.

6.4 Case Study

The effectiveness of the NRG models can be illustrated by the responses generated by them intuitively. Table 4 lists the sampled results given by all the baselines and our proposed LocalGAN, trained on the OpenSubtitle corpus. It can be observed that, compared to the basic Seq2Seq and Seq2Seq-MMI, the models incorporating adversarial learning have generated results with varying degrees of semantic diversity, as expected. Among the baselines, DAIM produces more diverse responses, which matches the quantified evaluation results in Table 1. The samples from LocalGAN have shown better semantic relevance, and meanwhile the diversity and informativeness are also guaranteed. The same observations can be get on the samples based on the Sina Weibo corpus, given in Figure 4.

Query: you know you're a strong woman and i admire that. **Ground-truth:** i'm sorry for what i'm about to do.	**Query:** you are a perfect girlfriend. **Ground-truth:** some times i wish that you would go out more.
Seq2Seq: you know what i mean. **Seq2Seq-MMI:** you know what you're doing.	**Seq2Seq:** what are you talking about? **Seq2Seq-MMI:** you have no idea what you're talking about.
Adver-REGS: i don't know how to do that. **GAN-AEL:** what are you talking about? **AIM:** what are you doing here? **DAIM:** i had a plan for you. **BigGAN:** i don't even know what it is.	**Adver-REGS:** do you have a girlfriend? **GAN-AEL:** what are you doing here? **AIM:** you are the only one. **DAIM:** you are a good person. **BigGAN:** why are you so nice to me?
LocalGAN: i thought it was your idea.	**LocalGAN:** you are a part of me.
Query: maybe cause you keep marrying for money. **Ground-truth:** i did not marry for money.	**Query:** you're just my roommate eric. **Ground-truth:** millie is much much more than that.
Seq2Seq: what do you mean? **Seq2Seq-MMI:** what do you want me to do?	**Seq2Seq:** i don't know what you're talking about. **Seq2Seq-MMI:** i need to talk to you.
Adver-REGS: what do you want to say? **GAN-AEL:** what is the difference? **AIM:** what are you talking about? **DAIM:** what is the problem? **BigGAN:** i don't want to say it.	**Adver-REGS:** i want to talk to you about this. **GAN-AEL:** can you help me please? **AIM:** what are you talking about. **DAIM:** its a pleasure to meet you. **BigGAN:** why are you telling me this?
LocalGAN: i don't have a choice.	**LocalGAN:** no you are my friend.
Query: he says some kind of sea creatures washed up on his beach. **Ground-truth:** how could it be a sea creature?	**Query:** would you excuse us for just a minute? **Ground-truth:** ted i'm gonna be honest.
Seq2Seq: i don't know what to do. **Seq2Seq-MMI:** what are we supposed to do?	**Seq2Seq:** why are you doing this? **Seq2Seq-MMI:** i don't want to see you again.
Adver-REGS: he wants to go to the sandglass. **GAN-AEL:** he told us to come back. **AIM:** i thought what was going to happen. **DAIM:** he said it was a gift. **BigGAN:** He is a very brave man.	**Adver-REGS:** what are you going to do? **GAN-AEL:** i have to go to the bathroom. **AIM:** what are you supposed to do? **DAIM:** why are you calling me? **BigGAN:** why are you so worried?
LocalGAN: i thought it looks like a fish.	**LocalGAN:** please give me your reason for absence.

Table 4: Sample results given by different models trained on the OpenSubtitle corpus.

Query: 凡事不要想的太复杂，手握的太紧，东西会碎，手会疼。 You don't need to make things too complicated, and it's just like something you are holding, if you use too much power, it will break and hurt you. **Ground-truth:** 没办法啊。为什么我那么敏感呢? I feel very helpless. Why am I so sensitive?	**Query:** 河豚君。生气的样子真呀么真可爱。 The globefish looks so cute when it gets angry. **Ground-truth:** 我还以为他吞了一个球呢。 I thought it had swallowed up a ball.
Seq2Seq: 这句话说的很好! That was spoken well. **Seq2Seq-MMI:** 这句话说得太好了。 That was spoken quite well.	**Seq2Seq:** 我也是这样的啊。 So am I. **Seq2Seq-MMI:** 原来是这样的啊! So that's it!
Adver-REGS: 这句话很有道理。 It quite make sense. **GAN-AEL:** 这句话是什么意思? What does that mean? **AIM:** 这句话说得太对了! These words speak the truth. **DAIM:** 那是我自己想要的东西。 That's what I want.	**Adver-REGS:** 我也有一个这样的。 I also have one like this. **GAN-AEL:** 我也有这样的感觉。 I also have the same feeling. **AIM:** 是啊，我也是这样的。 Yeah, so am I. **DAIM:** 好可爱的样子! It looks so cute!
LocalGAN: 放下，放下，立地成佛。 Let it go and free yourself.	**LocalGAN:** 这就是传说中的河豚? So this is the legendary globefish?
Query: 给所有的胖纸讲个鬼故事，夏天来了。 A terrible story to all the fatties: the summer is coming. **Ground-truth:** 现在也就这只肥猫能让我笑一笑。 This fatty cat looks so funny.	**Query:** 远离城市，才得安宁。 You can feel peaceful only if you are far from cities. **Ground-truth:** 我也是这么想的，但真正有几个人能做到呢。 Agree with that, but I think few people can really do it.
Seq2Seq: 我只想说这句话我很喜欢 I just wanna to say that I like this sentence. **Seq2Seq-MMI:** 我只想说这句话是什么意思? I just wanna say what do you mean?	**Seq2Seq:** 我也是这么认为 I also think so . **Seq2Seq-MMI:** 我也有这种感觉 I have the same feeling as well.
Adver-REGS: 我只能说，你是我的骄傲 I can only say that I am proud of you. **GAN-AEL:** 我不会也有这样的一天吧。 May I have such a day ? **AIM:** 我只能说，我也是这样的。 I can only say so do I. **DAIM:** 你是我的唯一，我的瘦子。 You are my only, skinny.	**Adver-REGS:** 这是什么意思? What does this mean? **GAN-AEL:** 又不是一个人的城市。 This is not a city for any single person. **AIM:** 我也想，但我不会去。 I also want it, but I will not take action. **DAIM:** 这就是我们的生活! This is our life!
LocalGAN: 这就是传说中的瘦子的心声。 This must come from the heart of a skinny guy.	**LocalGAN:** 心若不静，耳根亦不清静。 If you don't get inner peace, you will hear the boring noise.
Query: 只要生命还在，就可以从头再来。 Never give up as long as you are alive. **Ground-truth:** 有些东西. 不能重来。 But there always exist something not repeatable.	**Query:** 喜欢这样被婴儿的小手抓住手指的童鞋举手! Just let me know if you also love to be holding on your finger by the tiny hand of a baby! **Ground-truth:** 喜欢那带着稚气声音喊我姐姐、娜姨 I love to be called as the sister or auntie with baby voice.
Seq2Seq: 加油，永远支持你! Cheer up, and I am always supporting you. **Seq2Seq-MMI:** 每一天都是最好的。 Every day is the best day.	**Seq2Seq:** 我也想要这样的! I also want this! **Seq2Seq-MMI:** 我喜欢这种感觉! I love this feeling!.
Adver-REGS: 新的一天，加油! A new day, cheer up! **GAN-AEL:** 新的一天，新的起点! New day, new beginning! **AIM:** 加油，继续加油! Cheer up, forever! **DAIM:** 一切都会过去的。 Everything will pass.	**Adver-REGS:** 我家也是这样的! The one in my home is like this as well! **GAN-AEL:** 有木有想要一个? Do you want one like this? **AIM:** 好喜欢这种感觉。 Love this feeling so much. **DAIM:** 我也想要一个这样的男朋友 I also want a boyfriend like this.
LocalGAN: 淡定吧，一切都会好起来的。 Be calm, and everything will be OK.	**LocalGAN:** 好有爱啊，我也想要一个宝宝。 So cute, I also want a baby.

Figure 4: Sample Results given by different models trained on the Sina Weibo corpus.

7. Conclusions

This paper has given the theoretical proof of the upper bound of the adversarial training leveraged models on the Seq2Seq-based neural response generation task. The proof indicates that, due to the local distribution nature of query-response corpora, the GAN based NRG models will converge to the states mostly generating specialized patterns corresponding to given queries. To address this issue, we proposed to model the local distribution of the queries and its responses in the semantic space by adopting energy-based function, and found the approximation of this function. According to this approximated distribution representation, a new loss function describing the local expansion cost in the fitting of response distribution is presented and finally combined with the traditional GAN loss to form a hybrid training objective for the GAN based NRG model. This paper provides a reasonable explanation to the unstable training process and unsatisfying results of GAN based NRG approaches, and meanwhile gives a different perspective to leverage the local data distribution to enhance classic GAN approaches.

Acknowledgments

We thank the editor and anonymous reviewers for their insightful comments. This research is partially supported by the National Key R&D Program of China via grant 2018YFC0830700.

Appendix A. The Graphical Model for Response Distribution Modeling

Different from the computer vision related scenario, training a DBM on a set of text vectors is not trivial, since the training procedure is difficult to converge due to the value scale of text vectors is much larger than image vectors. Moreover, the simple scaling methods are not effective enough for this issue. For this purpose, this paper adopts standard-scaler[8] to remove the mean and scale to unit variance. To valid the effectiveness of standard-scaler, We conduct experiments on the query-response matching task using normalized vectors. The experiment result show that the matching performance based normalized vectors is similar to that of CNN based architecture (Kim, 2014).

Appendix B. Detailed Proof of Lemma 1

Proof [**Proof of Lemma 1**] For a fixed query, $F(q, r)$ can be seen as the the scalar function of vector r. For simplicity, we denote $F(q, r)$ as $f(r)$.

Taylor expansions for the first moment of function of random variables are as follows.

$$\begin{aligned}\mathbb{E}[f(r)]] &= \mathbb{E}\left[f\left(r_c + (r - r_c)\right)\right] \\ &= \mathbb{E}\left[f\left(r_c\right) + (r - r_c)^T Df(r_c) + \frac{1}{2}(r - r_c)^\top \left\{D^2 f(r_c)\right\}(r - r_c) + R_{r_c,2}(r - r_c)\right]\end{aligned}$$

where $Df(r_c)$ is the gradient of f evaluated at r_c, $D^2 f(r_c)$ is the Hessian matrix and $R_{r_c,2}(r - r_c)$ is the Lagrange remainder. Since $\mathbb{E}r = r_c$, the second term $\mathbb{E}[(r - r_c)^T Df(r_c)]$ disappears.

After that, we try to find the upper bound of the third term and the remainder term. The relevant theorems used in the proof are listed as follows.

- According to (Petersen et al., 2008), assuming that the matrix A is symmetric, $c = \mathbb{E}[\mathbf{x}]$ and $\mathbf{\Sigma} = \mathrm{Var}[x]$, then

$$\mathbb{E}\left[x^T \mathrm{A} x\right] = \mathrm{Tr}(\mathrm{A}\mathbf{\Sigma}) + c^T \mathrm{A} c.$$

- (Mirsky, 1975) states the following theorem: If A, B are complex $n \times n$ matrices with singular values $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ respectively, then

$$|\,\mathrm{Tr}(\mathrm{AB})| \leq \sum_{i=1}^{n} \alpha_i \beta_i$$

Firstly, since $r - r_c \sim N(0, \Sigma)$ and Σ is positive semi-definite matrix, the third term can be simplified as follows.

$$\begin{aligned}|\mathbb{E}[(r - r_c)^\top \left\{D^2 f(r_c)\right\}(r - r_c)]| &= |\,\mathrm{Tr}(\{D^2 f(r_c)\}\Sigma) + \mathbf{0}^\top D^2 f(r_c)\mathbf{0}| \\ &\leq \sum_{i=1}^{n} \alpha_i \beta_i \\ &\leq \alpha_1 \mathrm{Tr}(\Sigma) = \alpha_1 \mathbb{E}\left[\|r - r_c]\|_2^2\right]\end{aligned}$$

8. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

where $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ denote the singular value of matrix $\{D^2 f(r_c)\}$ and Σ respectively.

Meanwhile, according to the Claim 1 in (Folland, 2005), we have that

$$|R_{r_c,2}(r - r_c)| \leq \frac{\tilde{M}}{3!}\|r - r_c\|_1^3$$

where $\tilde{M}$ is the upper bound for absolute value of third-order partial derivatives of f.

Next, we show that the α_1 and $\tilde{M}$ can be bounded by M, where $M = \max_{h_q} ||h_q^T W_{qr}||_\infty$ (h_q follows multinomial distribution).

Substituting the definition of $E(q, r, H)$ into $F(q, r)$, we have following equation:

$$f(r) = F(q, r) = -\log \sum_{h_q,h} exp(h_q^T W_{qr} r + h_q^T W_{qq} q + h^T W_{qh} h_q).$$

Its first-order, second-order and third-order partial derivative are calculated as follows:

$$\frac{\partial f(r)}{\partial r_i} = -\frac{\sum_{h_q,h} exp(h_q^T W_{qr} r + h_q^T W_{qq} q + h^T W_{qh} h_q) \times (h_q^T W_{qr})_i}{\sum_{h_q,h} exp(h_q^T W_{qr} r + h_q^T W_{qq} q + h^T W_{qh} h_q)}$$

$$\frac{\partial^2 f(r)}{\partial r_i \partial r_j} = \sum_{h_q,h} a(h_q, h) \times b(h_q, i) \times \left[\sum_{\tilde{h}_q,\tilde{h}} a(\tilde{h}_q, \tilde{h}) \times b(\tilde{h}_q, j) - b(h_q, j) \right]$$

$$\begin{aligned}\frac{\partial^3 f(r)}{\partial r_i \partial r_j \partial r_k} =& \sum_{h_q,h} c(h_q, h, k) \times b(h_q, i) \times \left[\sum_{\tilde{h}_q,\tilde{h}} a(\tilde{h}_q, \tilde{h}) \times b(\tilde{h}_q, j) - b(h_q, j) \right] \\ &+ \sum_{h_q,h} a(h_q, h) \times b(h_q, i) \times \left[\sum_{\tilde{h}_q,\tilde{h}} c(\tilde{h}_q, \tilde{h}, k) \times b(\tilde{h}_q, j) \right]\end{aligned}$$

where

$$a(h_q, h) = \frac{exp(h_q^T W_{qr} r + h_q^T W_{qq} q + h^T W_{qh} h_q)}{\sum_{\tilde{h}_q,\tilde{h}} exp(\tilde{h}_q^T W_{qr} r + \tilde{h}_q^T W_{qq} q + \tilde{h}^T W_{qh} \tilde{h}_q)}$$

$$c(h_q, h, k) = \frac{\partial a(h_q, h)}{\partial r_k} = a(h_q, h) b(h_q, k) - a(h_q, h) \sum_{\tilde{h}_q,\tilde{h}} a(\tilde{h}_q, \tilde{h}) b(\tilde{h}_q, k)$$

and $b(h_q, i) = (h_q^T W_{qr})_i$ representing the i-th element of the vector $h_q^T W_{qr}$. According to the definition of $a(h_q, h)$, it is obvious that $a(h_q, h) > 0$ and $\sum_{h_q,h} a(h_q, h) = 1$.

The upper bounds for the second-order and third-order partial derivative are shown as follows.

$$\left| \frac{\partial^2 f}{\partial r_i \partial r_j} \right| \leq \sum_{h_q,h} |a(h_q, h)| \times |b(h_q, i)| \times \left[\sum_{\tilde{h}_q,\tilde{h}} \left| a(\tilde{h}_q, \tilde{h}) \times b(\tilde{h}_q, j) \right| + |b(h_q, j)| \right] \leq 2M^2$$

$$|c(h_q,h,k)| \le |a(h_q,h)| \times |b(h_q,k)| + |a(h_q,h)| \times \sum_{\tilde{h}_q,\tilde{h}} |a(\tilde{h}_q,\tilde{h})| \times |b(\tilde{h}_q,k)| \le 2M|a(h_q,h)|$$

$$|\frac{\partial^3 f(r)}{\partial r_i \partial r_j \partial r_k}| \le \sum_{h_q,h} |c(h_q,h,k)| \times M \times [M+M] + M^2 \sum_{\tilde{h}_q,\tilde{h}} |c(\tilde{h}_q,\tilde{h},k)| \le 6M^3$$

Based on the upper bound above, we can see that $\tilde{M}$ can be $6M^3$ and the upper bounds for α_1 and $|R_{r_c,2}(r-r_c)|$ are as follows.

$$\alpha_1 = \sigma_{\max}(\{D^2 f(r_c)\}) \le \|\{D^2 f(r_c)\}\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |\frac{\partial^2 f}{\partial r_i \partial r_j}|^2} \le 2nM^2$$

$$|R_{r_c,2}(r-r_c)| \le \frac{\tilde{M}}{3!}\|r-r_c\|_1^3 = M^3\|r-r_c\|_1^3 \le (\sqrt{n}M\|r-r_c\|_2)^3$$

Hence, we have

$$\begin{aligned} |\mathbb{E}[f(r)] - f(r_c)| &= |\frac{1}{2}\mathbb{E}\big[(r-r_c)^\top \{D^2 f(r_c)\}(r-r_c)\big] + \mathbb{E}\big[R_{r_c,2}(r-r_c)\big]| \\ &\le \frac{1}{2}\alpha_1 \mathbb{E}\|r-r_c]\|_2^2 + \mathbb{E}|R_{r_c,2}(r-r_c)| \\ &\le \mathbb{E}(\sqrt{n}M\|r-r_c]\|_2)^2 + \mathbb{E}(\sqrt{n}M\|r-r_c\|_2)^3 \end{aligned}$$

Therefore,

$$\lim_{\mathbb{E}\|r-r_c\|_2 \to 0} |\mathbb{E}[f(r)] - f(r_c)| = 0$$

■

Appendix C. The Proof of Theorem 1 Based on A Different Assumption

Let's consider a simpler situation for Theorem 1 and give an intuitive explanation as follows.

Remark 1 *In the semantic space, if the response* $r \sim \mathcal{N}\left(r_c, \begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2\end{bmatrix}\right)$*, a randomly sampled response* r *is intuitively very close to the center* r_c *with high probability when each* σ_j *is small enough. Correspondingly,* $F(q,r)$ *is very likely close to* $F(q,r_c)$*. That is, if* $r \sim \mathcal{N}\left(r_c, \begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2\end{bmatrix}\right)$ *and* $\mathbb{E}||r-r_c||_2$ *is small enough, the average* $F(q,R_q)$ *is very close to* $F(q,\hat{r}_c)$*, where* $\hat{r}_c$ *is the estimation of* r_c *based on the well-trained DBM.*

Proof [**Proof of Remark 1**] Since $F(q,r)$ is a continuous function of variable r when q is given, for any $\varepsilon > 0$ there exists a positive value δ such that $||r - r_c||_2 \leq \delta$ implies $|F(q,r) - F(q,r_c)| \leq \varepsilon$. Next, we will look at the lower bound of $P(||r - r_c||_2 \leq \delta)$.

Considering that the event $\{|r^{(j)} - r_c^{(j)}| \leq \frac{\delta}{\sqrt{n}}, 1 \leq j \leq n\}$ ($r^{(j)}$ is the j-th element of the n-dimensional vector r) is the subset of the event $\{||r - r_c||_2 \leq \delta\}$, we have that

$$P(||r - r_c||_2 \leq \delta) \geq P\left(\{|r^{(j)} - r_c^{(j)}| \leq \frac{\delta}{\sqrt{n}}, 1 \leq j \leq n\}\right)$$

Under the assumption $r_i \sim \mathcal{N}\left(r_c, \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}\right)$, the probability can be written as

$$\begin{aligned} P\left(\{|r^{(j)} - r_c^{(j)}| \leq \frac{\delta}{\sqrt{n}}, 1 \leq j \leq n\}\right) &= \prod_{j=1}^{n} P\left(|r^{(j)} - r_c^{(j)}| \leq \frac{\delta}{\sqrt{n}}\right) \\ &= \prod_{j=1}^{n} P\left(\left|\frac{r^{(j)} - r_c^{(j)}}{\sigma_j}\right| \leq \frac{\delta}{\sigma_j\sqrt{n}}\right) \\ &= \prod_{j=1}^{n} P\left(|Z| \leq \frac{\delta}{\sigma_j\sqrt{n}}\right) \end{aligned}$$

where Z follows the standard normal distribution. In other words, the following inequality holds

$$P(||r - r_c||_2 \leq \delta) \geq \prod_{j=1}^{n} P\left(|Z| \leq \frac{\delta}{\sigma_j\sqrt{n}}\right)$$

Similarly, due to the statement that $||r - r_c||_2 \leq \delta$ implies $|F(q,r) - F(q,r_c)| \leq \varepsilon$, the following inequality holds

$$P(|F(q,r) - F(q,r_c)| \leq \varepsilon) \geq P(||r - r_c||_2 \leq \delta) \geq \prod_{j=1}^{n} P\left(|Z| \leq \frac{\delta}{\sigma_j\sqrt{n}}\right)$$

Based on the inequality above, we can show that

$$P(|F(q,r) - F(q,r_c)| \leq \varepsilon) \to 1 \text{ when } \sigma_j \to 0 \ (j = 1, \cdots, n).$$

Under the assumption $r \sim \mathcal{N}\left(r_c, \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}\right)$, the condition $\sigma_j = 0$ $(j = 1, \cdots, n)$ is equivalent to the condition $E||r - r_c||_2 = 0$. Therefore, it can be concluded that the average $F(q, R_q)$ is intuitively close to $F(q,r_c)$ when $E||r - r_c||_2$ is very small. As discussed in (Wang et al., 2010; Srivastava and Salakhutdinov, 2012), a DBM sufficiently trained with large amounts of (q,r) pairs actually guarantees the estimation of any response r_i (denoted by $\hat{r}_i$), and thus the statement "well-trained" indicates that the estimate $\hat{r}_c$ is very close to r_c. Hence, the final conclusion holds. ■

Appendix D. The Illustration of the RDF Estimation

According to Equation 15, the proposed Radial Distribution Function (RDF) actually quantifies the radial offset from the center of the response cluster in the semantic space. Thus, a smaller RDF value indicates that the corresponding response is more generic, and apparently, an informative and meaningful response is expected to be assigned to a relatively larger RDF value.

	RDF (LocalGAN)	RDF (LocalGAN-SE)
Ground-truth: i did not marry for money.	0.3691	0.0522
what do you mean ?	0.0982	0.0139
what do you want me to do ?	0.0128	0.0174
what do you want to say ?	0.1508	0.0108
what is the difference ?	0.1595	0.0611
what are you talking about ?	0.2642	0.0651
what is the problem ?	0.2406	0.0426
i don't have a choice .	0.3917	0.0413

Table 5: A Toy Example, with the query ***maybe cause you keep marrying for money***, for Illustrating the RDF Estimation.

To illustrate the effect of the RDF estimation, as shown in Table 5 we employ a toy example with a given query utterance and compute the RDF values of the ground-truth and generated candidates. For comparison, we introduce the radial function $\frac{d(q,r)-d(q,\hat{r}_c)}{d(q,\hat{r}_c)}$ of the baseline LocalGAN-SE, where $d(q,r)$ is the Euclidean distance between q and r. From the values in Table 5, it can be seen that LocalGAN assigns much lower RDF values to the generic responses, and the most reasonable and informative response "*i don't have a choice*" gets the highest value. By contrast, the values given by LocalGAN-SE can only be used to roughly distinguish the more generic responses from the others. Especially, on the best candidate, the estimation given by LocalGAN is much closer to that of the ground-truth, compared to the estimation from LocalGAN-SE. The observations on this toy example match the experimental results in Section 6.

References

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning*, pages 214–223, 2017.

Sanjeev Arora and Yi Zhang. Do gans learn the distribution? some theory and empirics. In *ICLR*, 2018.

Sanjeev Arora, Rong Ge, Yingyu Liang, Tengyu Ma, and Yi Zhang. Generalization and equilibrium in generative adversarial nets (GANs). In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 224–232. PMLR, 06–11 Aug 2017.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In *Proceedings of the 2015 ICLR*, September 2015.

Ashutosh Baheti, Alan Ritter, Jiwei Li, and Bill Dolan. Generating more interesting responses in neural conversation models with distributional constraints. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3970–3980, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis. In *International Conference on Learning Representations*, 2019.

Jiachen Du, Wenjie Li, Yulan He, Ruifeng Xu, Lidong Bing, and Xuan Wang. Variational autoregressive decoder for neural response generation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3154–3163, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

Wenchao Du and Alan W Black. Top-down structurally-constrained neural response generation with lexicalized probabilistic context-free grammar. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3762–3771, 2019.

Xinya Du, Junru Shao, and Claire Cardie. Learning to ask: Neural question generation for reading comprehension. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1342–1352, 2017.

Shaoxiong Feng, Hongshen Chen, Kan Li, and Dawei Yin. Posterior-gan: Towards informative and coherent response generation with posterior generative adversarial network. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 7708–7715. AAAI Press, 2020. URL `https://aaai.org/ojs/index.php/AAAI/article/view/6273`.

G. B. Folland. Higher-order derivatives and taylor's formula in several variables. `https://sites.math.washington.edu/~folland/Math425/taylor2.pdf`, 2005. Accessed Nov 8, 2019.

Karl Friston, James Kilner, and Lee Harrison. A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3):70–87, 2006.

Karl Friston, Christopher Thornton, and Andy Clark. Free-energy minimization and the dark-room problem. *Frontiers in psychology*, 3:130, 2012.

Jun Gao, Wei Bi, Xiaojiang Liu, Junhui Li, Guodong Zhou, and Shuming Shi. A discrete cvae for response generation on short-text conversation. 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 1898–1908, 2019.

Asma Ghandeharioun, Judy Hanwen Shen, Natasha Jaques, Craig Ferguson, Noah Jones, Agata Lapedriza, and Rosalind Picard. Approximating interactive human evaluation with self-play for open-domain dialog systems. *arXiv preprint arXiv:1906.09308*, 2019.

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

Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimeoality of data with neural networks. *science*, 313(5786):504–507, 2006.

Geoffrey E Hinton and Richard S Zemel. Autoencoders, minimum description length and helmholtz free energy. In *Advances in neural information processing systems*, pages 3–10, 1994.

Lou Jost. Entropy and diversity. *Oikos*, 113(2):363–375, 2006.

Mahyar Khayatkhoei, Maneesh K Singh, and Ahmed Elgammal. Disconnected manifold learning for generative adversarial networks. In *Advances in Neural Information Processing Systems*, pages 7343–7353, 2018.

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

Abhishek Kumar, Prasanna Sattigeri, and Tom Fletcher. Semi-supervised learning with gans: Manifold invariance with improved inference. In *Advances in Neural Information Processing Systems*, pages 5534–5544, 2017.

Nicolas Le Roux and Yoshua Bengio. Representational power of restricted boltzmann machines and deep belief networks. *Neural computation*, 20(6):1631–1649, 2008.

Yann LeCun, Sumit Chopra, Raia Hadsell, M Ranzato, and F Huang. A tutorial on energy-based learning. *Predicting structured data*, 1(0), 2006.

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of NAACL-HLT*, pages 110–119, 2016.

Jiwei Li, Will Monroe, Tianlin Shi, Alan Ritter, and Dan Jurafsky. Adversarial learning for neural dialogue generation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2157–2169, 2017.

Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, and Xiaoyong Du. Analogical reasoning on chinese morphological and semantic relations. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 138–143, 2018.

Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. 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*, pages 2122–2132, 2016.

L. Mirsky. A trace inequality of john von neumann. *Monatshefte für Mathematik*, 79(4): 303–306, Dec 1975. ISSN 1436-5081. doi: 10.1007/BF01647331. URL `https://doi.org/10.1007/BF01647331`.

Jiquan Ngiam, Zhenghao Chen, Pang W Koh, and Andrew Y Ng. Learning deep energy models. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 1105–1112, 2011.

Gaurav Pandey, Danish Contractor, Vineet Kumar, and Sachindra Joshi. Exemplar encoder-decoder for neural conversation generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1329–1338, Melbourne, Australia, July 2018. Association for Computational Linguistics.

Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.

Kaare Brandt Petersen, Michael Syskind Pedersen, et al. The matrix cookbook. *Technical University of Denmark*, 7(15):510, 2008.

Ruslan Salakhutdinov and Geoffrey E. Hinton. Deep boltzmann machines. In *International Conference on Artificial Intelligence and Statistics*, pages 448–455, 2009.

Ruslan Salakhutdinov and Hugo Larochelle. Efficient learning of deep boltzmann machines. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 693–700, 2010.

Lifeng Shang, Zhengdong Lu, and Hang Li. 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, 2015.

Yuanlong Shao, Stephan Gouws, Denny Britz, Anna Goldie, Brian Strope, and Ray Kurzweil. Generating high-quality and informative conversation responses with sequence-to-sequence models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2210–2219, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1235.

P. Smolensky. *Parallel Distributed Processing: Foundations*, volume 1, pages 194–281. MIT Press, Cambridge, 1986.

Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan. 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, 2015.

Nitish Srivastava and Ruslan R Salakhutdinov. Multimodal learning with deep boltzmann machines. In *Advances in neural information processing systems*, pages 2222–2230, 2012.

Xingwu Sun, Jing Liu, Yajuan Lyu, Wei He, Yanjun Ma, and Shi Wang. Answer-focused and position-aware neural question generation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3930–3939, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

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

Baoxun Wang, Xiaolong Wang, Chengjie Sun, Bingquan Liu, and Lin Sun. Modeling semantic relevance for question-answer pairs in web social communities. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1230–1238. Association for Computational Linguistics, 2010.

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean. Google's neural machine translation system: Bridging the gap between human and machine translation. *CoRR*, abs/1609.08144, 2016. URL `http://arxiv.org/abs/1609.08144`.

Yu Wu, Furu Wei, Shaohan Huang, Yunli Wang, Zhoujun Li, and Ming Zhou. Response generation by context-aware prototype editing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 7281–7288, 2019.

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

Jingjing Xu, Xuancheng Ren, Junyang Lin, and Xu Sun. Diversity-promoting gan: A cross-entropy based generative adversarial network for diversified text generation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3940–3949, 2018.

Zhen Xu, Bingquan Liu, Baoxun Wang, SUN Chengjie, Xiaolong Wang, Zhuoran Wang, and Chao Qi. Neural response generation via gan with an approximate embedding layer. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 628–637, 2017.

Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. Seqgan: Sequence generative adversarial nets with policy gradient. In *AAAI*, pages 2852–2858, 2017.

Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu, and Xueqi Cheng. Learning to control the specificity in neural response generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1108–1117, Melbourne, Australia, July 2018a. Association for Computational Linguistics.

Xuemiao Zhang, Zhouxing Tan, Xiaoning Zhang, Yang Cao, and Rui Yan. Adaptively multi-objective adversarial training for dialogue generation. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 2872–2878. International Joint Conferences on Artificial Intelligence Organization, 7 2020. URL `https://doi.org/10.24963/ijcai.2020/397`.

Yizhe Zhang, Michel Galley, Jianfeng Gao, Zhe Gan, Xiujun Li, Chris Brockett, and Bill Dolan. Generating informative and diverse conversational responses via adversarial information maximization. In *Advances in Neural Information Processing Systems*, pages 1810–1820, 2018b.

Junbo Zhao, Michael Mathieu, and Yann LeCun. Energy-based generative adversarial network. In *Proceedings of the 2017 ICLR*, 2017.

Yao Zhao, Xiaochuan Ni, Yuanyuan Ding, and Qifa Ke. Paragraph-level neural question generation with maxout pointer and gated self-attention networks. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3901–3910, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

Kun Zhou, Kai Zhang, Yu Wu, Shujie Liu, and Jingsong Yu. Unsupervised context rewriting for open domain conversation. 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 1834–1844, 2019.

Qingfu Zhu, Lei Cui, Wei-Nan Zhang, Furu Wei, and Ting Liu. Retrieval-enhanced adversarial training for neural response generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3763–3773, 2019.

Qingfu Zhu, Weinan Zhang, Ting Liu, and William Yang Wang. Counterfactual off-policy training for neural dialogue generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3438–3448, 2020.