DP-GAN: Diversity-Promoting Generative Adversarial Network for Generating Informative and Diversified Text

Jingjing Xu*, Xu Sun*, Xuancheng Ren, Junyang Lin, Binzhen Wei, Wei Li

-

Problems Addressed

1. Existing text generation methods tend to produce repeated and "boring" expressions.
2. Directly applying a classifier as the discrimina- tor leads to the problem that the reward given by the classifier doesn't reflect the novelty of text accurately.
3. A simple classifier can reach very high accuracy (almost 99%), which makes most generated text receive reward around zero because the discriminator can identify them with high confidence.
4. The reason for this problem is that the training objective of the classifier-based GAN is minimizing the Jensen-Shannon Divergence (JSD) between the distributions of the real data and the generated data.
5. If the accuracy of classifier is too high, JSD fails to measure the distance between the two distributions, and cannot give reasonable reward to the model for generating real and diverse text.
6. To tackle this problem, the authors propose Diversity-Promoting Generative Adversarial Network (DP-GAN).

Claims

1. The proposed model assigns low reward for repeated text and high reward for "novel" text, encouraging the generator to produce diverse and informative text.
2. The model has a language-model based discriminator, which can better distinguish novel text from repeated text without the saturation problem compared with existing classifier-based discriminators.

Proposed Solution

1. The cross entropy generated by the language-model based discriminator is set to be the reward for the generator.

Generator

- This paper assumes that the output of the model can be long text made up of multiple sentences.
- Given the input sentence $x_{1:m} = (x_1, x_2, x_3, ..., x_m)$ of m words from Γ, the vocabulary of words, the model generates the text of T sentences $Y_{1:T} = (y_1, ..., y_t, ..., y_T)$, where y_t from Λ, the set of candidate sentence. The term $y_t = (y_{t,1}, ..., y_{t,K})$ is the t^{th} sentence, where $y_{t,K}$ is the K^{th} word.

Discriminator

- The cross entropy generated by the language-model based discriminator is set to be the reward for the generator.
- Given a sentence y_t, the term $D_\phi(y_{t,k}|y_{t,<k})$ is a probability indicating how likely y_t, k is as the next token. The cross entropy is then used to calculate the reward:

$$R(y_{t,k}) = -\log D_\phi(y_{t,k}|y_{t,<k}), k = 1, 2, ..., K$$

- To encourage the model to generate novel and diverse text, the discriminator is required to assign higher reward to the real text and lower reward to the generated text. Thus, the reward of the real text is maximized and the reward of the generated text is minimized.
- The loss function of the discriminator is formulated as:

$$J(\phi) = -E_{Y\sim p_{data}}[R(Y)] + E_{Y\sim G_\Theta}[R(Y)]$$

  where $R(Y)$ stands for the averaged reward of Y.
- The model assigns a Word Level reward, assigned as follows:

$$R(y_{t,k}|y_{t,<k}) = -\log D_\phi(y_{t,k}|y_{t,<k})$$

  Similarly, a sentence level reward is defined as:

$$R(y_{t,k}|y_{t,<k}) = -\frac{1}{K}\sum_{k=1}^{K}\log D_\phi(y_{t,k}|y_{t,<k})$$

- The language model is able to assign low reward for the text that appears very frequently and high reward for the text that is uncommon. The reward for novel text is high and does not saturate, while the reward for text with low novelty is small but discriminative.

Key Points

- Policy gradient is used to train the network.
- Teacher forcing is used to train the Generator.

Experiments

1. For review generation, the number of generated sentences is set to 6 with the maximum length of 40 words for each generated sentence.
2. The hidden size is set to 256, embedding size to 128, vocabulary size to 50K, and batch size to 64 for the generator and the discriminator.
3. Adagrad optimizer with the initial learning rate 0.1 is used.
4. In adversarial training, the step for training the generator is 1K, the step for training the discriminator is 5K.
5. Both the generator and the discriminator are pre-trained for 10 epochs before adversarial learning.
6. To be fair, the settings of all sequence-to-sequence models in the baselines are the same with our generator.
7. For PG-BLEU and SeqGAN, before reinforcement learning or adversarial learning, the sequence-to-sequence model is pre-trained for 10 epochs like DP-GAN.
8. For dialogue generation, the settings are the same with review generation, except that the number of generated sentences is set to 1 with the maximum length of 40 words because there is only one sentence in the response.

Results

Yelp	Token	Dist-1	Dist-2	Dist-3	Dist-S
MLE	151.2K	1.2K	3.9K	6.6K	3.9K
PG-BLEU	131.1K	1.1K	3.3K	5.5K	3.1K
SeqGAN	140.5K	1.1K	3.5K	6.1K	3.6K
DP-GAN(S)	**438.6K**	1.7K	7.5K	15.7K	10.6K
DP-GAN(W)	271.9K	2.8K	14.8K	29.0K	12.6K
DP-GAN(SW)	406.8K	**3.4K**	**22.3K**	**49.6K**	**17.3K**
Amazon	Token	Dist-1	Dist-2	Dist-3	Dist-S
MLE	176.1K	0.6K	2.1K	3.5K	2.6K
PG-BLEU	124.5K	0.6K	1.9K	3.5K	2.3K
SeqGAN	217.3K	0.7K	2.6K	4.6K	3.2K
DP-GAN(S)	**467.6K**	0.8K	3.6K	7.6K	7.0K
DP-GAN(W)	279.4K	1.6K	8.9K	18.4K	9.6K
DP-GAN(SW)	383.6K	**1.9K**	**11.7K**	**26.3K**	**13.6K**
Dialogue	Token	Dist-1	Dist-2	Dist-3	Dist-S
MLE	81.1K	1.4K	4.4K	6.3K	4.1K
PG-BLEU	97.9K	1.2K	3.9K	5.5K	3.3K
SeqGAN	83.4K	1.4K	4.5K	6.5K	4.5K
DP-GAN(S)	**112.2K**	1.5K	5.2K	8.5K	5.6K
DP-GAN(W)	79.4K	1.9K	7.7K	11.4K	6.0K
DP-GAN(SW)	97.3K	**2.1K**	**10.8K**	**19.1K**	**8.0K**

Table 1: Performance of the DP-GAN and three baselines on review generation and dialogue generation tasks. Higher is better. DP-GAN(S), DP-GAN(W), and DP-GAN(SW) represent DP-GAN with only sentence-level reward, only word-level reward, and combined reward, respectively. *Token* represents the number of generated words. Dist-1, Dist-2, Dist-3, and Dist-S are respectively the number of distinct unigrams, bigrams, trigrms, and sentences in the generated text. For example, 1.2K in Dist-1 means 1200 distinct unigrams.

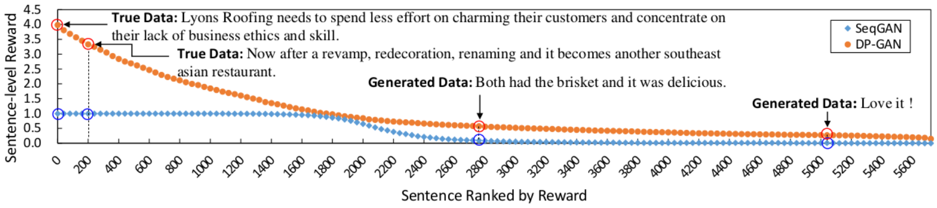

Figure 2: Distribution of rewards between SeqGAN and DP-GAN. The upper two sentences are sampled from the true data and the lower two sentences are sampled from the generated data. It is important to note that the sentence-level reward of DP-GAN is averaged word-level reward and a long sentence does not indicate a high score. As we can see, the reward distribution of SeqGAN saturates and cannot distinguish the novelty of the text accurately. DP-GAN has a strong ability of resisting reward saturation and can give more precise reward for text in terms of novelty.

	Model	Averaged Ranking
Yelp	MLE	1.89
	PG-BLEU	2.22
	SeqGAN	2.12
	DP-GAN	**1.51**
Amazon	MLE	1.93
	PG-BLEU	2.24
	SeqGAN	1.98
	DP-GAN	**1.50**
Dialogue	MLE	2.46
	PG-BLEU	2.40
	SeqGAN	2.17
	DP-GAN	**1.92**

Table 2: Results of human evaluation on three datasets. Lower is better.

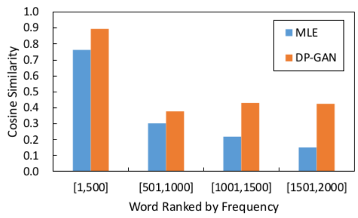

Figure 3: Cosine similarity between true data distribution and generated data distributions of various models. For example, the first column represents the cosine similarity on top 500 words with the highest frequencies in true data. As we can see, the generated data distribution of DP-GAN is closer to true data distribution, especially considering words of low frequency.

Bibliography

1. Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In ICML 2017, pages 214–223, 2017.

2. Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In ICLR 2014, 2014.

3. Dzmitry Bahdanau, Philemon Brakel, Kelvin Xu, Anirudh Goyal, Ryan Lowe, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio. An actor-critic algorithm for sequence pre- diction. In ICLR 2017, 2017.

4. Samy Bengio, Oriol Vinyals, Navdeep Jaitly, and Noam Shazeer. Scheduled sampling for sequence prediction with recurrent neural networks. In NIPS 2015, pages 1171– 1179, 2015.

5. David Berthelot, Tom Schumm, and Luke Metz. BE- GAN: boundary equilibrium generative adversarial net- works. CoRR, abs/1703.10717, 2017.

6. Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable rep- resentation learning by information maximizing generative adversarial nets. In NIPS 2016, pages 2172–2180, 2016.

7. Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In EMNLP 2014, pages 1724–1734, 2014.

8. Emily L. Denton, Soumith Chintala, Arthur Szlam, and Rob Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. In NIPS 2015, pages 1486–1494, 2015.

9. John C. Duchi, Elad Hazan, and Yoram Singer. Adap- tive subgradient methods for online learning and stochas- tic optimization. Journal of Machine Learning Research, 12:2121–2159, 2011.

10. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative adversarial nets. In NIPS 2014, pages 2672–2680, 2014.

11. Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C. Courville. Improved training of wasserstein gans. In NIPS 2017, pages 5769–5779, 2017. Kelvin Guu, Tatsunori B. Hashimoto, Yonatan Oren, and Percy Liang. Generating sentences by editing prototypes. CoRR, abs/1709.08878, 2017.

12. Jiwei Li, Minh-Thang Luong, and Dan Jurafsky. A hierarchical neural autoencoder for paragraphs and documents. In ACL 2015, pages 1106–1115, 2015.

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

14. Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, and Jian feng Gao. Deep reinforcement learning for dialogue gen- eration. In EMNLP 2016, pages 1192–1202, 2016.

15. Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. Adversarial learning for neural di- alogue generation. In EMNLP 2017, pages 2157–2169, 2017.

16. Chia-Wei Liu, Ryan Lowe, Iulian Serban, Michael Nosewor- thy, Laurent Charlin, and Joelle Pineau. How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation.In EMNLP 2016, pages 2122–2132, 2016.

17. Thang Luong, Hieu Pham, and Christopher D. Manning. Effective approaches to attention-based neural machine translation. In EMNLP 2015, pages 1412–1421, 2015.

18. Shuming Ma and Xu Sun. A semantic relevance based neural network for text summarization and text simplification. CoRR, abs/1710.02318, 2017.

19. Shuming Ma, Xu Sun, Jingjing Xu, Houfeng Wang, Wenjie Li, and Qi Su. Improving semantic relevance for sequence- to-sequence learning of chinese social media text summa- rization. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 2: Short Papers, pages 635–640, 2017.

20. Julian John McAuley and Jure Leskovec. From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews. In WWW 2013, pages 897–908, 2013.

21. Sebastian Nowozin, Botond Cseke, and Ryota Tomioka. f- gan: Training generative neural samplers using variational divergence minimization. In NIPS 2016, pages 271–279, 2016.

22. Alec Radford, Luke Metz, and Soumith Chintala. Unsu- pervised representation learning with deep convolutional generative adversarial networks. CoRR, abs/1511.06434, 2015.

23. Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. Sequence level training with recurrent neural networks. In ICLR 2016, 2016.

24. Tim Salimans, Ian J. Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In NIPS 2016, pages 2226–2234, 2016.

25. 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 EMNLP 2017, pages 2210–2219, 2017.

26. Xu Sun, Xuancheng Ren, Shuming Ma, and Houfeng Wang. meprop: Sparsified back propagation for accelerated deep learning with reduced overfitting. In Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017, pages 3299–3308, 2017.

27. Xu Sun, Bingzhen Wei, Xuancheng Ren, and Shuming Ma. Label embedding network: Learning label representation for soft training of deep networks. CoRR, abs/1710.10393, 2017.

28. Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In NIPS 2014, pages 3104–3112, 2014.

29. Richard S. Sutton, David A. McAllester, Satinder P. Singh, and Yishay Mansour. Policy gradient methods for rein- forcement learning with function approximation. In NIPS 1999, pages 1057–1063, 1999.

30. Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Ma- chine Learning, 8:229–256, 1992.

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