Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2024.080705
ISSN 2371-8838 Vol. 8 Num. 7

Conditional Question Generation Model Based on Diffusion Model

Yujie Wei[1,a,*], Kui Jin[1,b]

[1]*School of Information, Yunnan Normal University, Kunming, China*
[a]*wsamrtyujie@163.com*, [b]*1459871084@qq.com*
[*]*Corresponding author*

Keywords: Deep Learning, Question Generation, Diffusion Model

Abstract: The ultimate goal of conditional question generation is to generate high-quality questions with diversity, and the classifier-based diffusion model for conditional problem generation mainly categorizes the source data through classifiers so that high-quality problems can be generated. However, this kind of generation method has the drawback of complex joint training process and over-dependence on labeled data, which can lead to the lack of diversity and quality of generation. To tackle this problem, we propose a novel classifier-free diffusion model for conditional question generation. First, discrete text data are mapped into continuous vector data as input of the model in terms of an embedding function. Second, we design a classifier-free training method, which embeds the condition into the data fitting process, and the vector data completes the training under the condition. Finally, with the aid of rounding function, the samples generate the discrete text problem data. Experiments show that our proposed approach achieves a relatively decent average score and realizes better problem diversity than other state-of-the-art methods.

1. Introduction

Conditional question generation is an important yet challenging problem, which aims to generate natural and relevant questions by extracting relevant information from natural language text. The approaches of conditional question generation can be broadly classified into traditional rule-based question generation methods and deep neural network-based methods. The former heavily relies on heuristic rules to constrain question generation, which can be divided into three categories: template-based [1], semantic-based methods [2] and syntax-based [3]. Template-based approaches utilize templates extracted from the training set to design questions. For instance, Wolfe et al. [4] first proposed a template-based educational system using pattern matching for question generation. Semantic-based approaches analyze text semantically to create questions. For example, Yao and Zhang [5] proposed a question generation methodology based on Minimal Recursion Semantics representation, which is unnecessary the templates or syntax information. Syntax-based approaches first determine the syntactic structure of a given text and then mix syntactic question word placements and transformation rules to obtain the questions. As an exemplar, Straach and Truemper [6] took advantage of knowledge base and logic programming techniques as an expert system, which was used to implement question generation. However, all three methods require manual constraints design to accomplish question generation, which heavily relies on specialized knowledge and experience, therefore, they perform poorly in terms of model generalization and extensibility.

Deep neural network-based methods tackle the problem of poor scalability of the above methods, providing a fully data-driven end-to-end trainable framework. These approaches have improved the scalability as well as the quality and diversity of question generation, and could be roughly divided into Recurrent Neural Networks-based models (RNN) [7] [8], RNN fused with Self-Attention Mechanism [9], Answer Position Features-based models [10], deep model-based approaches such as VAE [11], GAN [12], DDPM [13], Graph Neural Networks-based [14] models, and pre-training model frameworks models [15] [16] [17]. For instance, Du et al. [7] applied RNN-based Seq2Seq architecture to conditional question generation and achieved better results than traditional methods. Later on, Liu et al. [10] integrated control elements of answer location features to further improve the generation quality and effectiveness, and Chai et al. [9] incorporated the coverage mechanisms into the RNN-based Seq2Seq framework, which sentences are utilized into input text. However, the inherent sequential nature of the RNN-based framework leads to problems of increased computational cost and long-time dependency, and hence novel approaches are needed to enhance the effectiveness of conditional question generation.

To cope with foregoing problem, generative models are applied to question generation tasks as well. For instance, Wang et al. [11] applied VAE to question generation task, which maps variables into latent variables by utilizing an encoder, and then a decoder is used to map latent variables to output variables. The model is trained by a maximum variational lower bound. Dathathri et al. [17] proposed a plug-and-play conditional question generation approach by means of external classifier control, which controls the final question generation by analyzing the extent to which the classifier conditions are satisfied in the generated text. More frameworks for textual attribute-level control [18][19] emerged. For example, Li et al. [13] introduced a classifier-based diffusion model for conditional question generation called Diffusion-LM. This method is built on the framework of the diffusion model incorporating an external classifier. After converting the mapped data into a series of Gaussian vectors in the forward process, the external control conditions of the classifier are integrated into the intermediate variable transformation process, and then the noise is reduced to word vectors corresponding to the conditions in the reverse process. It takes full advantage of the characteristic of high-quality generation of the diffusion model, and achieves better generation quality and diversity than other methods.

However, the classifier design of classifier-based methods in more complex conditional question generation tasks will be more redundant. At the same time, different control conditions lead to an increase of control attributes and labeled data, and eventually, the joint train process is complex, thus affecting the diversity effect of question generation.

To relieve the limitations stated above, we propose a classifier-free diffusion model for conditional question generation. Our primary contributions are listed:

1) We design a conditional question generation-based model on classifier-free diffusion model, which incorporates conditional control into the generation process of diffusion models, without designing separate classifiers. This technology reduces the extra losses from training classifiers and reduce the training time.

2) After the comparison experiments with the three aspects of question generation methods, by comparing the evaluation metrics that impact the quality and diversity of generation, our methodology all show superior question diversity and quality of generation.

2. Problem Description

The objective of the classifier-free conditional question generation diffusion is to generate fixed-length target question data according to the features of the source sequence after learning the features of the given-length source sequence data. The input source sequences are described as tuples $\{w^x, w^y\}$, where $w^x = \{w_1^x, ..., w_n^x\}$ denotes a sequence of questions with n-length and $w^y = \{w_1^y, ..., w_n^y\}$ denotes a sequence of answers with n-length. We design an embedding function $EMB(\mathbf{W})$ to transform the binary data into a continuous vector, where $\mathbf{W}^{\mathbf{x} \oplus \mathbf{y}}$ is the cascade vector of w^x and w^y. The forward process gradually adds Gaussian noise and subsequently learns the features between w^x and w^y. In the reverse process, a prediction function $f_\theta(\mathbf{z}_t, t)$ is designed to predict an answer vector $\mathbf{x}_t$ incorporating textual information of $\mathbf{y}_t$. The answer vector $\mathbf{x}_t$ was rounded by the Rounding($\mathbf{z}_t$) function to obtain a word vector w^{out}. Table 1 shows an example of data.

Table 1: Example of the model data

Inputs: $\{w^x, w^y\}$	Outputs: w^{out}
{How can I contact her? From writing letters.}	{How do I contact her?}
{What color is the book? It is red.}	{What color is that book?}
{When did you buy this dress? This Friday.}	{When did you get this dress}

3. Proposed framework

The procedure for conditional question generation based on classifier-free diffusion is a gradual process of transforming sequential data into Gaussian noise, then learning the features, and sampling the relevant data from the Gaussian noise. The data simulation process in the classifier-free condition is demonstrated. Therefore, in order to realize the conditional question generation based on the classifier-free method, we devise a conditional question generation framework based on the classifier-free diffusion, where discrete data are initially embedded into the model to obtain continuous vector data. The continuous vector data are used to obtain the target question vectors after the diffusion model forward and reward processes, which are finally fed into the rounding function to obtain the word vectors. Where w denotes the group $\{w^x, w^y\}$ that is injected into the model to obtain the intermediate variable $x_0, x_1, ..., x_T$. In this paper, the Improved Denoising Diffusion Model (DDIM) [20] is used as the basic model.

4. Experiment

This section outlines the datasets used for the experiments and the parameter settings, and the evaluation criteria for the experimental comparisons.

4.1. Dataset

The Stanford Question and Answer Dataset (SQuAD) is the most prevalently used reading comprehension dataset for question generation. It is divided into two datasets SQuAD1.1 [21] and SQuAD2.0 [22], which are composed of a set of questions and answers from Wikipedia articles.

Since a certain percentage of unanswerable questions were added to the SQuAD2.0, the possibility that an answer may be empty also occurs.

4.2. Parameter Settings and Training Environment

The data used in our experiment first required fine-tuning in HuggingFace. Specifically, we fine-tune BART-based model 2 epochs on the SQuAD2.0 dataset to obtain the checkpoint BART-base-SQuAD2.0-2 epoch (BbS 2), which was then used to initialize the diffusion. The encoder of the diffusion is initialized two times by the BbS 2. In addition, linear layers that do not exist in BbS 2 but exist in the diffusion will be randomly initialized. We set gradient descent by virtue of the Adam Optimizer with the coefficient set to 0.7. Table 2 details the parameter settings for our experiment.

Table 2: Parameter settings for our experiment.

Parameter	Value
batch size	20
epoch	3
learning rate	0.00002
dropout	0.2
beam search size	10
maximal input length	1024
maximal problem size	20
minimal problem size	3

5. Results

5.1. Baseline

Based on the methods frequently used in the field of conditional question generation, we classify our comparison experiments into three main categories, i.e., the traditional end-to-end approach, conditional question generation based on the deep generative modeling framework, question generation model architectures built on autoregressive (AR) methods and non-autoregressive methods (NAR). We hence compare the following three sets of models as a baseline for our experiments:

1) Transformer and GRU with Attention: this set of methods is based on the encoder-decoder architecture; which is well-behaved in SEQ2SEQ sequential task solving, and we conduct parameter comparison experiments on the more practical of these models.

2) GPVAE: A model modified on the basis of deep generative model VAE. The comparison experiments between ours and GPVAE are used to demonstrate the superiority of the diffusion model as a generative model over other generative models in terms of generation speed and generation quality.

3) Strongly iterative NAR model Lev: For the last batch of baselines, we utilize a widely used method for comparison of autoregressive methods. All baselines were trained for comparison as described in their papers.

5.2. Comparison Results

Based on the evaluation criteria, we summarize the results of the comparison in Table 3.

Table 3: Comparison results of modeling methods

Methods	BLEU	ROUGE	Score	dist-1	Len
Ours	0.1671	0.3666	0.6023	0.9065	11.9
GRU-attention	0.0651	0.2617	0.5222	0.7930	10.1
Transformer-base	0.1663	0.3441	0.6307	0.9309	10.3
NAR	0.0930	0.2893	0.5491	0.8914	6.93
GPVAE	0.1251	0.3390	0.6308	0.9381	11.4

Referring to the indicators affecting the quality and diversity of question generation, we conduct experiments with the above models and compared the indicators. It can be seen that our model performs better in terms of generation quality and diversity than the question generation model based on encoder-decoder structure and non-autoregressive methods, while in comparison with the GPVA method, ours is worse than GPVAE in terms of diversity, but still outperforms that method in terms of generation quality.

6. Conclusion

Nowadays, the conditional question generation task has been a prominent problem in AI, and there are various kinds of methods used to implement the question generation, yet, due to the existence of various problems in the current methods, which lead to the shortage of generation quality and diversity as well as the complexity of the training process. Meanwhile, for the limited methods for the conditional question generation task resolved by applying the diffusion model, this paper proposes a classifier-free diffusion model for conditional question generation. We first utilize the embedding function to embed the discrete text data sets into the continuous diffusion model to obtain the continuous vector data that can be processed by the continuous model. Then adds noise to the vector data and learns the features of the data in the forward process, and in the reverse process the model designs a prediction function to generate the answer vectors constrained by the input text. Subsequently, the answer vectors are input the rounding function to obtain the final word vectors, and completing the whole question generation. The model is constructed on the framework of a classifier-free diffusion, which does not undergo a separate classifier treatment thus improving the diversity and quality of the generated questions. In the experimental period, we compare three different types of question generation models longitudinally, conducting several comparative experiments affecting the quality and diversity metrics of questions. The final experiment indicates that our classifier-free diffusion model can achieve better generation quality and diversity. In the future, it is noteworthy to continue use the diffusion to generate problems in specialized fields and related disciplines on this basis, which can be of greater benefit to related subjects.

Acknowledgements

The work is supported by the Graduate Research Innovation Fund of Yunnan Normal University (No.YJSJ23-B113)

References

[1] Mostow J, Chen W. Generating instruction automatically for the reading strategy of self-questioning[C]. Artificial Intelligence in Education. IOS Press, 2009: 465-472.

[2] Kunichika H, Katayama T, Hirashima T, et al. Automated question generation methods for intelligent English learning systems and its evaluation[C]. Proc. of ICCE. 2004, 670.

[3] Huang Y, He L. Automatic generation of short answer questions for reading comprehension assessment[J]. Natural Language Engineering, 2016, 22(3): 457-489.

[4] Du X, Shao J, Cardie C. Learning to ask: Neural question generation for reading comprehension[J]. arXiv preprint arXiv:1705.00106, 2017.
[5] Hu, W., Liu, B., Ma, J., Zhao, D., & Yan, R. (2018). Aspect-based Question Generation. International Conference on Learning Representations.https://openreview.net/forum?id=rkRR1ynIf
[6] Chai Z, Wan X. Learning to ask more: Semi-autoregressive sequential question generation under dual-graph interaction[C]. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 225-237.
[7] Liu B, Zhao M, Niu D, et al. Learning to generate questions by learning what not to generate[C]. The world wide web conference. 2019: 1106-1118.
[8] Wang W, Feng S, Wang D, et al. Answer-guided and semantic coherent question generation in open-domain conversation[C]. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 5066-5076.
[9] Bao J, Gong Y, Duan N, et al. Question generation with doubly adversarial nets[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018, 26(11): 2230-2239.
[10] Chen Y, Wu L, Zaki M J. Reinforcement learning based graph-to-sequence model for natural question generation[J]. arXiv preprint arXiv:1908.04942, 2019.
[11] Dong L, Yang N, Wang W, et al. Unified language model pre-training for natural language understanding and generation[J]. Advances in neural information processing systems, 2019, 32.
[12] Johansen S, Juselius K. Maximum likelihood estimation and inference on cointegration—with appucations to the demand for money[J]. Oxford Bulletin of Economics and statistics, 1990, 52(2): 169-210.
[13] Dathathri S, Madotto A, Lan J, et al. Plug and play language models: A simple approach to controlled text generation[J]. arXiv preprint arXiv:1912.02164, 2019.
[14] Krause B, Gotmare A D, McCann B, et al. Gedi: Generative discriminator guided sequence generation[J]. arXiv preprint arXiv:2009.06367, 2020.
[15] Yang K, Klein D. FUDGE: Controlled text generation with future discriminators[J]. arXiv preprint arXiv:2104.05218, 2021.
[16] Li X, Thickstun J, Gulrajani I, et al. Diffusion-lm improves controllable text generation[J]. Advances in Neural Information Processing Systems, 2022, 35: 4328-4343.
[17] Wolfe J H. Automatic question generation from text-an aid to independent study[C]. Proceedings of the ACM SIGCSE-SIGCUE technical symposium on Computer science and education. 1976: 104-112.
[18] Yao X, Zhang Y. Question generation with minimal recursion semantics[C]. Proceedings of QG2010: The Third Workshop on Question Generation. 2010: 68-75.
[19] Straach J, Truemper K. Learning to ask relevant questions[J]. Artificial Intelligence, 1999, 111(1-2): 301-327.
[20] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851.
[21] Rajpurkar P. Squad: 100,000+ questions for machine comprehension of text[J]. arXiv preprint arXiv:1606.05250, 2016.
[22] Rajpurkar P, Jia R, Liang P. Know what you don't know: Unanswerable questions for SQuAD[J]. arXiv preprint arXiv:1806.03822, 2018.