Sequence Generative Adversarial Network for Chinese Social Media Text Summarization

北京邮電大學 Beijing University of Posts and Telecommunications

WenChuan Yang, Rui Hua, Qiuhan Zhao

Beijing University of Posts and Telecommunications

Abstract

Although the sequence-to-sequence models have achieved state-of-the-art performance in many summarization datasets, there are still some problems in the processing of Chinese social media text, such as short sentences, lack of coherence and accuracy. These issues are caused by two factors: the principle of the RNN-based sequence-to-sequence model is maximum likelihood estimation, which will lead to gradient vanishing or exploding when generating long summaries; the text in the Chinese social media is long and noisy, for which it is very difficult to generate high-quality summaries. To solve these issues, we apply a sequence generative adversarial network framework. The framework includes generator and discriminator, in which generator is used to generate summaries and discriminator is used to evaluate generated summaries. The softargmax layer is used as a connection layer to guarantee the co-training of generator and discriminator. Experiments are carried out on Large Scale Chinese Social Media Text Summarization Dataset. The length of the sentence, ROUGE score and artificial score of summary's quality are used to evaluate the generated summaries. The result shows that the sentences in the summaries generated by our model are longer and have higher accuracy.

Contributions

(i) A sequence generative adversarial training framework is applied and generates the summaries with longer and more accuracy sentence.

(ii) A softargmax layer is proposed, which is used as a connection layer between discriminator and generator to solve the problem that the back propagation optimization algorithm can not be used to optimize the two models simultaneously because of the discretation of the data between two models.

Proposed Model

1 Sequence Generative Adversarial Network

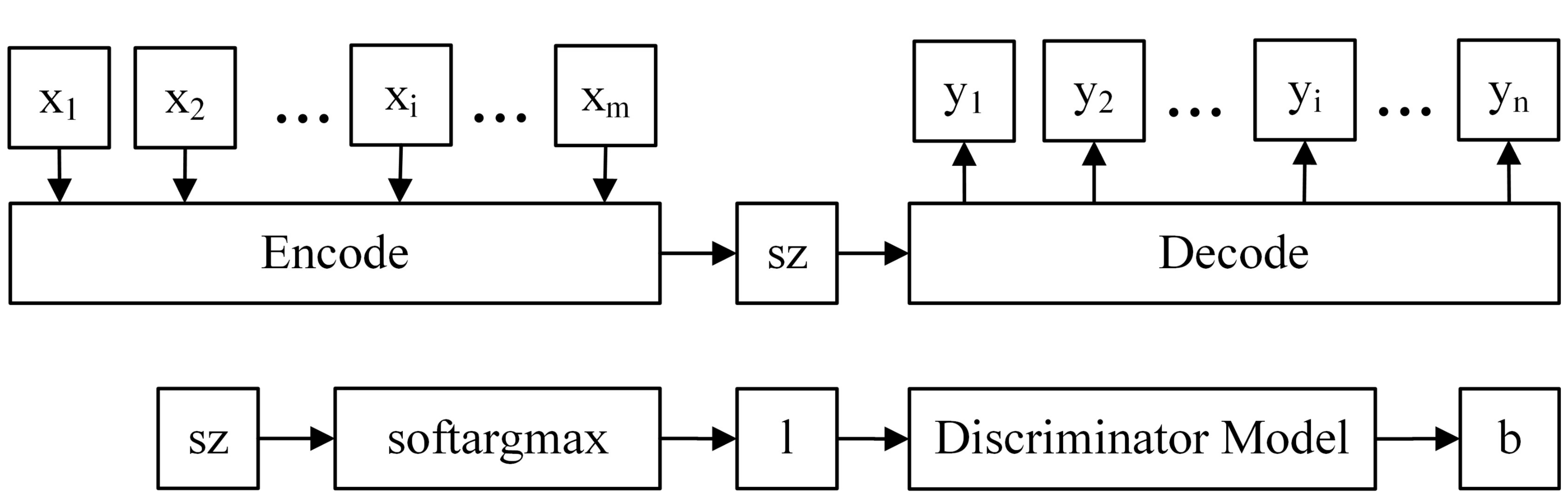

Fig. 1. The overview of our model

In this paper, a sequence generative adversarial network framework is proposed. The framework consists of the Seq2Seq model and the discriminator model. As shown in Fig. 1, the upper part in the graph is Seq2Seq model, and the lower part of the graph is the discriminator model. The softargmax layer is of importance in connecting two different models. At the training stage, we use the whole model. At the test stage, we only use the model Seq2Seq.

2 Softargmax Layer

In the field of text, there are cases where discrete data is not differentiable. Using softarmax layer, the data of the Seq2Seq can be approximately transformed into the ID sequence of the word to the discriminator model, so as to achieve the purpose of optimization. Fig. 2 shows output of the word ID with the maximum probability per line

	修 1	改 2	后 3	… …	的 45	立 46	法 47	公 48	布 49	… …	V |V|		
sz_1	0.5	0.1	0.1	…	0.01	0.02	0.05	0.1	0.1	…	…	l_1	1
sz_2	0.01	0.6	0.1	…	0.02	0.05	0.03	0.11	0.07	…	…	l_2	2
sz_3	0.02	0.05	0.8	…	0.01	0.02	0.04	0.01	0.01	…	…	l_3	3
sz_4	0.01	0.01	0.02	…	0.07	0.03	0.1	0.04	0.01	…	…	l_4	45
sz_5	0.02	0.01	0.01	…	0.01	0.5	0.1	0.03	0.07	…	…	l_5	46
sz_6	0.03	0.04	0.01	…	0.02	0.1	0.6	0.01	0.01	…	…	l_6	47
sz_7	0.02	0.01	0.03	…	0.01	0.05	0.03	0.7	0.01	…	…	l_7	48
sz_8	0.01	0.01	0.01	…	0.01	0.01	0.01	0.01	0.9	…	…	l_8	49

Fig. 2. The softargmax layer diagram

Experiments

In this paper, experiments are carried out on LCSTS data set, and the results are evaluated with ROUGE-1, ROUGE-2, ROUGE-L and other related indicators. The Table 1 shows the ROUGE value compared with different models.

Table 1. Rouge Comparison with Different Models

Models	ROUGE-1	ROUGE-2	ROUGE-L
RNN-character	21.5	8.9	18.6
RNN-word	17.7	8.5	15.8
RNN-context-character	29.9	17.4	27.2
RNN-context-word	26.8	16.1	24.1
SRB	33.3	20.0	30.1
Seq2Seq(this paper)	34.2	17.4	27.5
Our Model(this paper)	38.7	20.1	32.7

For a more intuitive view of the effect of the experimental summary, let's take an example to show in Table 2.

Tabel 2. Comparision of a Summarization Example

Source: 大众集团在2015CES开幕当天，宣布推出一款能够自主学习的半自动泊车系统。它可以学习了解驾驶员的习惯和行为，在场景重复时能够控制方向盘自动完成泊车。

On the opening day of CES in 2015, Volkswagen Group announced the introduction of a self-learning semi-automatic parking system. It can learn to understand the driver's habits and behavior, and can control the steering wheel to complete parking automatically when the scene repeats.

Reference: 懒人的福音：大众推出能“自主学习”的自动泊车系统

The Gospel of Lazy Man: Volkswagen launches auto parking system with self-learning ability.

Seq2Seq: 2015CES开幕开幕

Opening of CES in 2015

Our Model: 大众集团推出自主学习的半自动泊车系统

Volkswagen Group launched self-learning semi-automatic parking system.

Conclusion

In this work, we apply a sequence generative adversarial training framework to improve the summarization quality of Chinese social media text. The softargmx layer connecting generator and discriminator model is proposed to solve the non-differentiable problem of discrete text data. Experiments in dataset LCSTS shows that the ROUGE score of summarization given by our model can get higher than most models in dataset LCSTS. In addition, our model can give longer sentence length, more coherent and more accuracy summarization.