

flexible and creative chinese poem generation using neural memory

Anonymous ACL submission

Abstract

Recently using neural network in the generation of chinese poem has been some achievements ,e.g., be able to produce relatively fluent Poetry. However, these poems have just reached the basic standard,and have a lot of room for improvement.In this paper, we demonstrate that a new neural model can generate flexible and creative chinese poem. the model adds memory module on the basis of attention-based recurrent neural network, which accepts a set of keywords as the theme and generates poems by looking at each keyword during the generation. We particularly focus on traditional Chinese poetry, and after expert evaluation show that machines can not only produce different styles of poem ,but also more stably generate better than the poems by the attention-based poem generation in terms of rhythm, fluency, and mood, theme-consistent.

1 Introduction

The classical Chinese poetry is a special cultural heritage with over 2,000 years of history and is still fascinating many contemporary poets. In history, Chinese poetry flourished in different genres at different time, including Tang poetry, Song iambics and Yuan songs. Different genres possess their own specific structural, rhythmical and tonal patterns. The structural pattern regulates how many lines and how many characters per line; the rhythmical pattern requires that the last characters of certain lines hold the same or similar vowels; and the tonal pattern requires characters in particular positions hold particular tones, i.e., ‘Ping’ (level tone), or ‘Ze’ (downward tone). A good poem should follow all these pat-

tern regulations (in a descendant order of priority), and has to express a consistent theme as well as a unique emotion. For this reason, it is widely admitted that traditional Chinese poetry generation is highly difficult and can be only performed by a very few knowledged people.

Among all the genres of traditional Chinese poetry, perhaps the most popular is the quatrain, a special style with a strict structure (four lines with five or seven characters per line), a regulated rhythmical form (the last characters in the second and fourth lines must follow the same rhythm), and a required tonal pattern (tones of characters in some positions should satisfy some predefined regulations). This genre of poems flourished mostly in Tang Dynasty, so often called ‘Tang poem’ . An example of quatrain written by Lun Lu, a famous poet in Tang Dynasty (Wang, 2002), is shown in Table 1.

Due to the stringent restriction in rhythm and tone, it is not trivial to create a fully rule-compliant quatrain. More importantly, besides such strict regulations, a good quatrain should also read fluently, hold a consistent theme, and express a unique affection. This is like dancing in fetlers, hence very difficult and can be performed only by knowledged people with long-time training.

We are interested in machine poetry generation, not only because of its practical value in entertainment and education, but also because it demonstrates an important aspect of artificial intelligence: the creativity of machines in art generation. We hold the belief that poetry generation (and other artistic activities) is a pragmatic process and can be largely learned from past experience. In this paper, we focus on traditional Chinese poetry generation, and demonstrate that machines can generate flexible and creative chinese poem.

Recently using neural network in the generation of chinese poem has been some achievements

塞下曲 Frontier Songs 月黑雁飞高, (* Z Z P P) The wild goose flew high to the moon shaded by the cloud, 单于夜遁逃。 (P P Z Z P) With the dark night' s cover escaped the invaders crowd, 欲将轻骑逐, (* P P Z Z) I was about to hunt after them with my cavalry, 大雪满弓刀。 (* Z Z P P) The snow already covered our bows and swords.
--

Table 1: An example of a quatrain. The rhyming characters are in boldface, and the tonal pattern is shown at the end of each line, where ‘P’ indicates level tone and ‘Z’ indicates downward tone, and ‘*’ indicates the tone can be either.

,e.g., be able to produce relatively fluent Poetry. There have been some attempts in this direction, e.g., by machine translation models (He et al., 2012) and recurrent neural networks (RNN) (xing Zhang and Lapata, 2014). However, these poems have just reached the basic standard, and have a lot of room for improvement.

In this paper, we demonstrate that a new neural model can generate flexible and creative chinese poem. This model is inspired by the fact that a poet should not only consider the existing rules, but also refer to the poems he has seen in the past in the creation of poetry. the model adds memory module on the basis of attention-based recurrent neural network, which accepts a set of keywords as the theme and generates poems by looking at each keyword during the generation. We particularly focus on traditional Chinese poetry, and after expert evaluation show that machines can not only produce different styles of poem ,but also more stably generate better than the poems by the attention-based poem generation in terms of rhythm, fluency, and mood, theme-consistent.

2 Related Work

.....

3 Method

In this section, we first present the new model that Chinese poetry generation framework of combin-

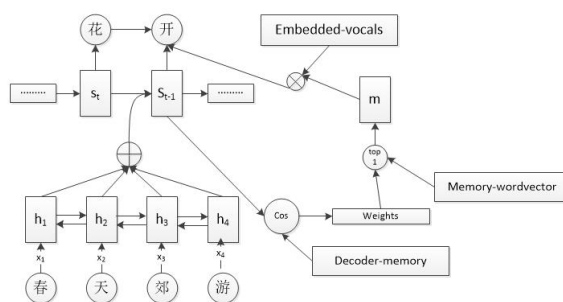


Figure 1: the new model that Chinese poetry generation framework of combining memory mechanism and attention mechanism

ing memory mechanism and attention mechanism, and then describe the implementation of the new model that have been tailored for our task, shown in Figure 1.

3.1 The New Model

As shown in Figure 1, the new model consists of two parts, one is attention-base recurrent neural network, the other is memory module that perfects the attention-base recurrent neural network.

the attention-based model applied to Chinese poetry generation is illustrated in the left part of Figure 1. The encoder (a bi-directional GRU that will be discussed shortly) converts the input keywords, a character sequence denoted by (x_1, x_2, \dots) , into a sequence of hidden states (h_1, h_2, \dots) . The decoder then generates the whole poem character by character, denoted by (y_1, y_2, \dots) . At each step t , the prediction for the next character y_t is based on the ‘current’ status s_t of the decoder as well as all the hidden states (h_1, h_2, \dots, h_n) of the encoder. Each hidden state h_i contributes to the generation according to a relevance factor $\alpha_{t,i}$ that measures the similarity between s_t and h_i .

the main role of the memory mechanism in the right of Figure 1 perfects the attention-base recurrent neural network.

3.2 Model Training

The new model just need to train the attention-base recurrent neural network, then based on which add memory mechanism.

The goal of the attention-base recurrent neural network training is to let the predicted character sequence match the original poem. We chose the cross entropy between the distributions over

Method	num	Compliance	Fluency	Mood	Imagery-Consistence	Overall
Mem+Frontier style	52	3.73	2.42	2.46	2.76	2.84
Mem+Pastoral style	42	4.02	2.88	3.00	3.14	3.26
Mem+Romantic style	52	3.98	2.69	2.51	2.96	3.03
Mem+all styles	146	3.90	2.65	2.63	2.94	3.03
Attention	62	4.08	2.95	2.59	2.87	3.12

Table 2: Attention denotes the poems by attention-based recurrent neural network;Mem+Frontier style denotes the poems by the new model using Frontier style of poems as the content of memory mechanism;Mem+Pastoral style denotes the poems by the new model using Pastoral style of poems as the content of memory mechanism;Mem+Romantic style denotes the poems the new model using Romantic style of poems as the content of memory mechanism;Mem+all styles denotes the poems that is the sum of Mem+Frontier style, Mem+Pastoral style and Mem+Romantic style

Chinese characters given by the decoder and the ground truth (essentially in a one-hot form) as the objective function. To speed up the training, the minibatch stochastic gradient descent (SGD) algorithm was adopted. The gradient was computed sentence by sentence, and the AdaDelta algorithm was used to adjust the learning rate (Zeiler, 2012). Note that in the training phase, there are no keyword input, so we use the first line as the input to generate the entire poem.

4 Experiments

We describe the experimental settings and results in this section. Firstly the datasets used in the experiments are presented, and then we report the evaluation in two phases: (1) the first phase focuses on generating different styles of poems, then let experts evaluate if achieving the goal(2) the second phase focuses on more stably generating better than the poems by the attention-based poem generation in terms of rhythm, fluency, and mood, theme-consistent,then let experts evaluate if achieving the goal.

4.1 Datasets

Two datasets are used to conduct the experiments. Firstly a Chinese quatrain corpus was collected from Internet. This corpus consists of 13, 299 5-char quatrains and 65,560 7-char quatrains. As far as we know, this covers most of the quatrains that are retained today. We filters out some poems which contains 100% low frequency words. Through corpus cleaning, a corpus which contains 9,195 5-char quatrains and 49,162 7-char quatrains was obtained. 9, 000 5-char and 49, 000 7-char quatrains are used to train the LSTM model of the attention model and the rest poems are used as the test datasets.

The second dataset was used to train and derive

character vectors for attention model initialization. This dataset contains 284, 899 traditional Chinese poems in various genres, including Tang quatrains, Song iambics, Yuan Songs, Ming and Qing poems. This large amount data ensures a stable learning for semantic content of most characters.

4.2 Different Styles of Poems

In the phase, we use three different styles of poems, which includes about 100 Frontier style’s, about 100 Pastoral style’s and about Romantic style’s, as the content of memory mechanism. The experiment is to prove that the new model can generate the style of poems without sacrificing the poems quality via using a certain style of content of memory mechanism.

We randomly picked out some of the poems generated by the new model with different styles of poems as content of memory mechanism and attention-based recurrent neural network,then let expert evaluate the poems in two aspects:one is scoring in rhythm, fluency, mood and Imagery-consistent,in the scale from 0 to 5; the other is judging the style of these poems. the results of evaluation are presented in Table 2 and Table 3.

Method	corr rate
Mem+Frontier style	0.80
Mem+Pastoral style	0.95
Mem+Romantic style	0.63
Mem+all styles	0.79

Table 3: corr rate denotes the proportion of the style of poems in all poems by the new model with a certain style of poems as the content of memory mechanism.

4.3 Better, More Creative Poems

In the phase, we randomly picked out about 5000 poems in training data as the content of memory mechanism.The experiment is to prove that the new model can more generate better, more creative

VS	Compliance		Fluency		Theme-Consistence		Mood		Imagery-Consistence	
	better	worse	better	worse	better	worse	better	worse	better	worse
Best+Mem vs Best	0.63	0.37	0.63	0.37	0.61	0.39	0.68	0.32	0.57	0.43
Last+Mem vs Last	0.61	0.39	0.73	0.27	0.70	0.30	0.70	0.30	0.66	0.34
Last vs Best	0.43	0.57	0.35	0.65	0.31	0.69	0.32	0.68	0.33	0.67
Last+Mem vs Best	0.57	0.43	0.49	0.51	0.44	0.56	0.52	0.48	0.45	0.55

Table 4: Demographic Prediction performance comparison by three evaluation metrics.

poems than attention-based recurrent neural network.

We randomly picked out some of poem pairs in all poem pairs combined by the same theme poems respectively generated by the new model with 5000 poems as content of memory mechanism and attention-based recurrent neural network, then let expert judge which better in rhythm, fluency, Theme-consistence, mood and Imagery-consistence.

References

- Jing He, Ming Zhou, and Long Jiang. 2012. Generating chinese classical poems with statistical machine translation models. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*. pages 1650–1656.
- Li Wang. 2002. *A Summary of Rhyming Constraints of Chinese Poems (Shi Ci Ge Lv Gai Yao)*, volume 1. Beijin Press.
- Xing xing Zhang and Mirella Lapata. 2014. Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. pages 670–680.