

Generating Chinese Classical Poems with LSTM Algorithm

Ruibo Liu

*Ira A. Fulton Schools of Engineering
Arizona State University
Tempe, Arizona 602338-0983
Email: rliu68@asu.edu*

Abstract—Sequence-to-Sequence (seq2seq) modeling, which has been proven as an effective NLP tool for many text-generation and sequence-labeling tasks, has rapidly become a very hot research topic. Sutskever et al. [1] applied Long-Short Term Memory (LSTM) to address seq2seq problems, since the traditional RNN model has trouble dealing with long-term dependencies, and DNN can only be applied to problems whose inputs and targets are of fixed length. Compared with RNN and DNN, LSTM is more flexible that can largely alleviate the quick-forgetting problem associated with the traditional RNN model. In encoder-decoder architecture for seq2seq learning, the neural network usually needs to compress all necessary source information into a fixed-length vector, which still makes it difficult to cope with long sentences. To solve this issue, Bahdanau et al. [2] proposed an attention-based mechanism, which achieved an impressive result. Also, Cheng et al. [3] introduced agreement-based learning into attention-based neural model to improve alignment quality for seq2seq learning tasks. Wiseman and Rush [4] extends seq2seq to learn global sequence scores to avoid classical biases associated with local training. The goal of my project is to utilize LSTM algorithm to generate Chinese classical poems.

1. Scheduling

I am going to implement a LSTM model based on encoder and decoder structure before March. Then, I plan to add attention-based mechanism to the LSTM model, and test if there is any improvement in performance around March mid. More tests and debug would extend this period of time. Before April, I would find a optimal hyper-parameter strategy and start to writing the final report. Finally, I would finish the report before the final exam.

2. Related Works

This project is inspired by Zhe et al. [5] 's work on the similar goal, and LSTM algorithm is implemented by myself roughly in CSE 575 course. But in that course, I used it to solve a QA problem. Now I want to add attention-based mechanism and try to do some alaysis on the results. Also, a different application is my challenge.

References

- [1] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems* (pp. 3104-3112).
- [2] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- [3] Cheng, Y., Shen, S., He, Z., He, W., Wu, H., Sun, M., & Liu, Y. (2015). Agreement-based joint training for bidirectional attention-based neural machine translation. *arXiv preprint arXiv:1512.04650*.
- [4] Wiseman, S., & Rush, A. M. (2016). Sequence-to-Sequence Learning as Beam-Search Optimization. *arXiv preprint arXiv:1606.02960*.
- [5] Zhe Wang, Wei He, Hua Wu, Haiyang Wu, Wei Li, Haifeng Wang, Enhong Chen (2016). Sequence-to-Sequence Learning as Beam-Search Optimization. *arXiv preprint arXiv:1610.09889*.