Applied Deep Learning
hw3 report

Han Hao Chen, r05922021

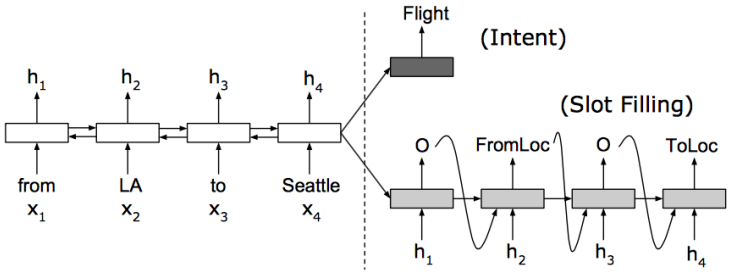
November 29, 2016

1 RNN Model

RNN In this assignment, we have two tasks, both of them can be trained on a single RNN model by combining their loss functions. For intent prediction task. we can use a RNN with sequential input data and output one prediction result at the last time step. Then, use the cross entropy as the loss function with the given ground truth from training data and the prediction result. For tagging prediction task, we can also use a RNN with a encoder and a decoder. Use the RNN encoder to encode the input sequences' information and use the RNN decoder to decode the sequence.

These two RNN models can be combined into a single one. The new model's loss function should minimize the summation of two loss functions from intent prediction task and tagging task. Moreover, the performance of this new model can even do better than the separated models.

Figure 1: Combined RNN Model

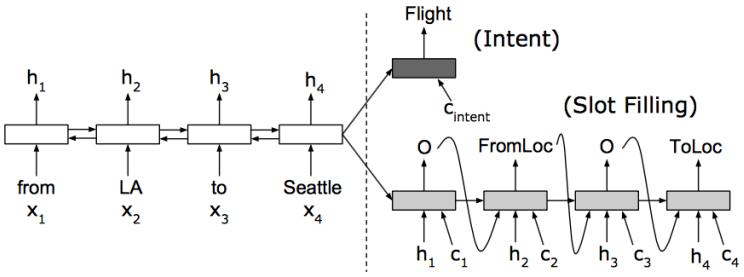


2 Learned and Improvement

Improvement In tagging task, we can use attention-based RNN to get a better performance. By changing the input of the RNN decoder to the weighted linear combination(Eq 1) the the each time step's output, the decoder model will have more informative feature to decode the right prediction result.

$$c_i = \sum_{j=1}^{T} \alpha_{i,j} h_j \tag{1}$$

Figure 2: Attention-Based RNN Model



Learned I found that the sequential information can be well modeled by the Recurrent Neural Network. Moreover, we can use Bidirectional RNN to achieve a higher performance, since it can some how capture more informative info from the both direction of the sequence.

References

[1] Bing Liu, Ian Lane, *Attention-Based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling*, Carnegie Mellon University, 2016.