Tutorial 8

Question 1

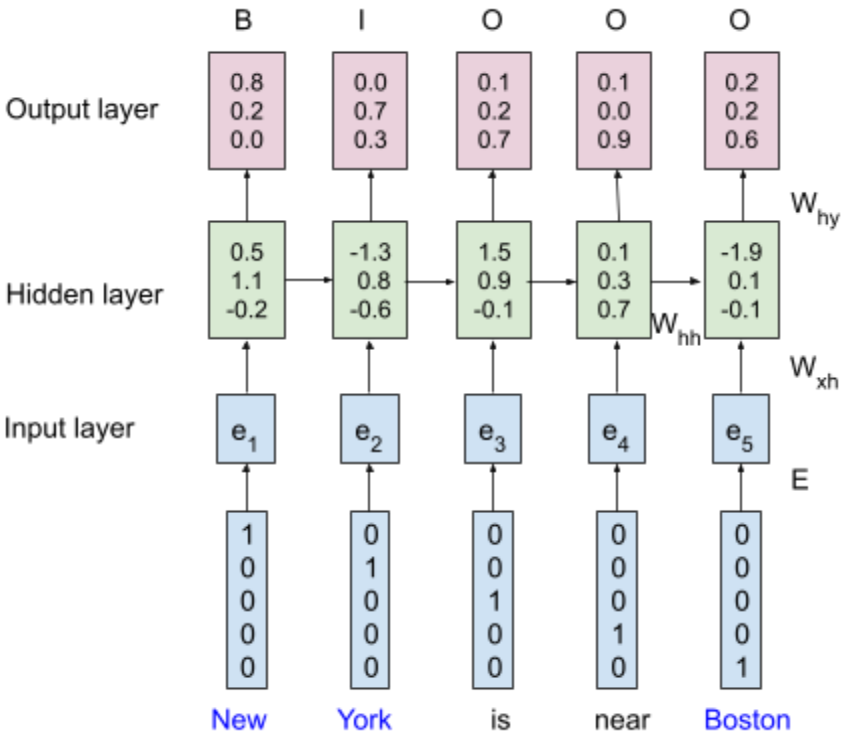
RNNs have been used for many problems in NLP. For this exercise, we will consider RNNs for named entity recognition as shown in the Figure below, assuming there is only one type of named entity (person).

1) If W_{xh} and W_{hh} are the weight matrices for input to hidden layer and hidden to hidden layer, write down the composition function of a vanilla unidirectional RNN model.

2) Write the composition functions for a bidirectional RNN.

3) Describe the number of classes in this NER task and how the final label is predicted in the output layer given the hidden representation for each word.

4) F1-score is a commonly used metric to evaluate the NER performance. It is computed based on recall and precision: $F1 = \frac{2 \times recall \times precision}{recall + precision}$. Suppose your model predicts "B I O O O" whereas the gold label should be "B I O O B". What is the F1 score in this case?



Question 2

To actually implement RNNs in python, there are several points to take note of.

1) For mini-batch training, the input to RNNs contains a number of sentences with different lengths. In order to perform efficient batch training, how to handle variable-length inputs?

2) In language model training with RNNs, the model uses the correct previous tokens from the dataset to predict the next one. However, during testing, it must rely on its own predictions to generate future tokens. How might this difference between training and testing impact the model's performance? Can you think of a solution to alleviate this issue?

Question 3

Imagine a simplified neural machine translation model that translates sentences from English to French. Consider that the model uses a single-layer RNN for both the encoder and the decoder. For simplicity, assume that each word in both languages has already been represented as a fixed-size vector (e.g., word2vec embedding).

1) Describe the role of the encoder RNN in the neural machine translation process.

2) What is the purpose of the decoder RNN, and how does it differ in functionality from the encoder?

3) Imagine you have an English sentence: "I am learning". You wish to translate it into French: "Je suis en train d'apprendre". Describe the steps the NMT model would take to perform this translation.

4) How would the process change if the encoder is a bi-directional RNN?

5) What could be the problem of greedy decoding during translation generation? Can you come up with any possible solution?

6) What are some potential drawbacks or limitations of using RNNs for NMT?

Coding exercises (for your own interest)
https://drive.google.com/file/d/1aDFQtSCYBpgRwkj9eq8Rao9eq5_68C31/view?usp=sharing