R11942152 Yu Wang

1

(1) intent_cls.sh:

First, I split each sentence into words and used GLoVe pre-trained embeddings to transform each word into a vector. Next, the pre-trained embeddings were updated during model training. Finally, I used the embeddings trained by my model to transform each word into a vector.

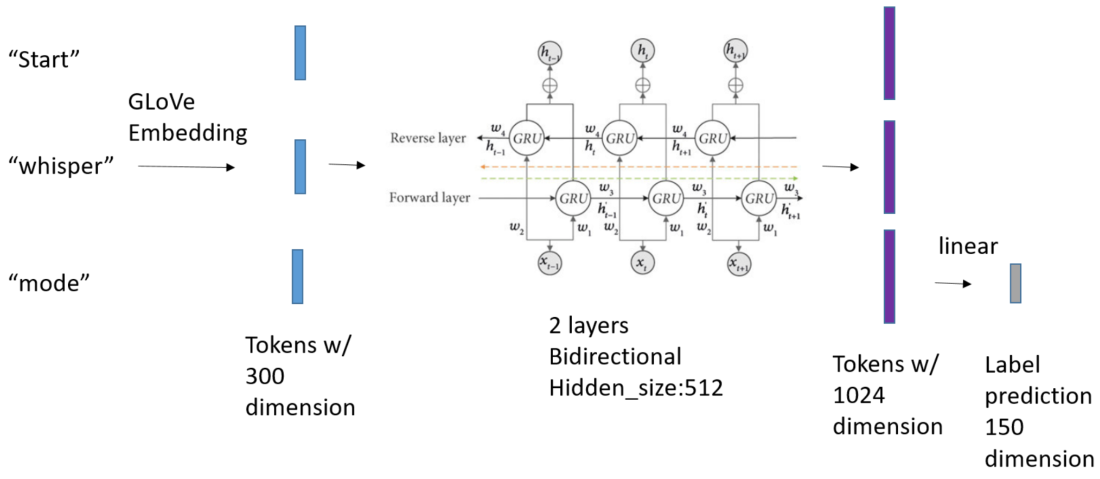
(2) slot_tag.sh:

The embeddings I used in this problem were similar to those in the intent classification problem. I started with GLoVe pre-trained embeddings as my initial embeddings and updated them during model training.

I used the sample codes for both problems.

2

a.

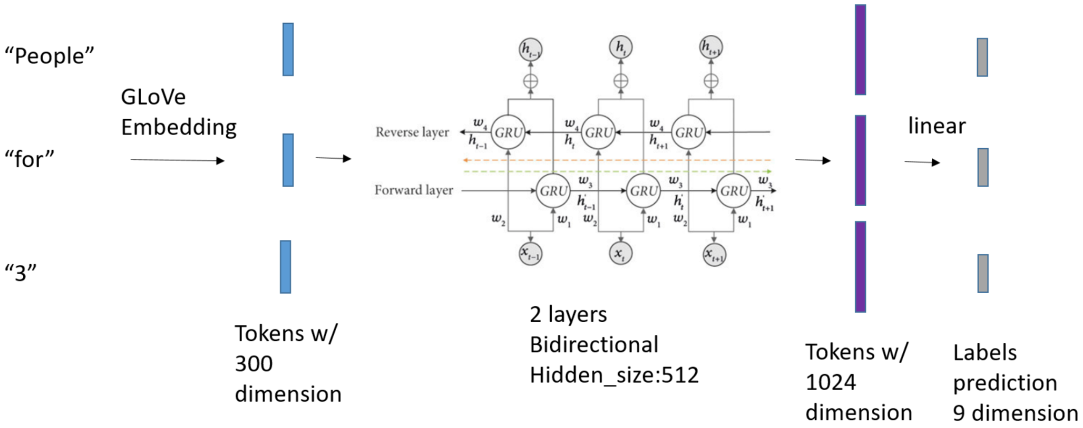


c. I use cross entropy loss.

d. I use AdamW optimization algorithm with learning rate 0.001 and batch size 128.

3

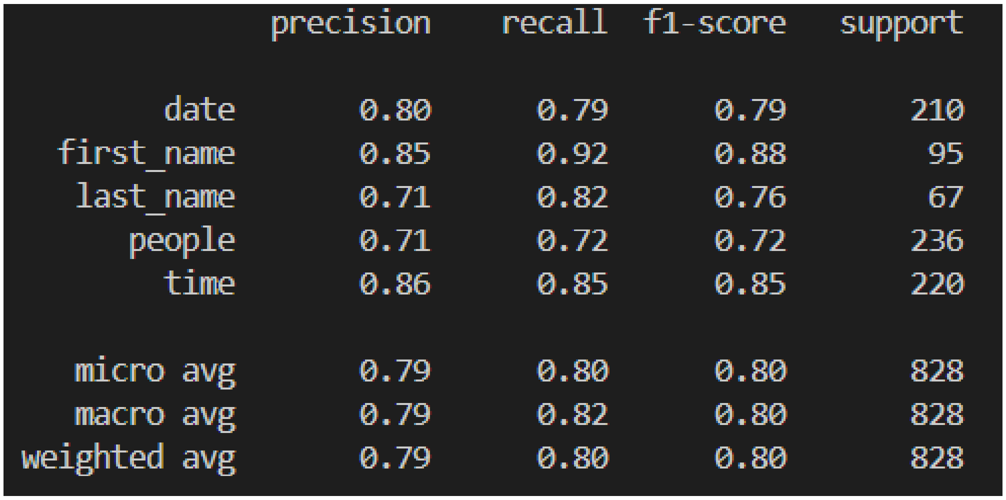
a.



c. I use cross entropy loss.

d. I use AdamW optimization algorithm with learning rate 0.001 and batch size 128.

4



```
              precision    recall  f1-score   support

        date       0.80      0.79      0.79       210
  first_name       0.85      0.92      0.88        95
   last_name       0.71      0.82      0.76        67
      people       0.71      0.72      0.72       236
        time       0.86      0.85      0.85       220

   micro avg       0.79      0.80      0.80       828
   macro avg       0.79      0.82      0.80       828
weighted avg       0.79      0.80      0.80       828
```

The precision of label 1 is $\frac{\text{the number we predict to label1 and being right.}}{\text{the number of predicting to label1}}$.

The recall of label 1 is $\frac{\text{the number we predicting to label1 and being right.}}{\text{the number of data whose label are 1}}$.

The F1-score of label 1 is harmonic mean of the precision and the recall of label 1, as we can see is slightly less than average of the precision and the recall when the precision and the recall are close to each other.

Token accuracy is the ratio of correctly predicted tokens to the total number of tokens, while joint accuracy is the ratio of correctly predicted sequences to the total number of sequences.