NLU project exercise lab: 10

Marco Carraro (240642)

University of Trento

marco.carraro-1@studenti.unitn.it

1. Introduction

Within this exercise, two features had to be implemented at the baseline of an IAS Model in the first part, namely:

- addition of bidirectionality;
- addition of a dropout layer;

The dataset used is that of ATIS. The second part of the exercise involved the implementation of a BERT model. I attempted to reach a solution by working on it for quite some time, however, I was not able to get it, so I found it of no use to send a part of the code that did not work.
Thus, below are the results obtained from the first part of the exercise.

2. Implementation details

In both cases, for the addition of both bidirectionality and dropout layers, I used the following dimensions: hid_size = 200, emb_size = 300.
The optimizer selected was Adam, with a learning rate value of 0.0001.
However, I built two different models for the two case histories, in the first applying only bidirectionality to it, and in the second adding the dropout layer to it.
I thus obtained two distinct results that I could compare with each other.
For bidirectionality, I specified that the utt_encoder layer was, in fact, bidirectional. This means that the layer processes input both from left to right and vice versa, to allow the model to take advantage of both past and future information for each word.
Regarding the dropout layer, in the second case, I added precisely a Dropout type layer with probability 0.1 to my model. I then specified the Dropout layer activation at the output of the LSTM encoder within the forward function of the model itself.

3. Results

Additions	Slot F1	Intent Accuracy
Bidirectionality	0.937 $\pm$0.003	0.948 $\pm$0.003
Dropout	0.941 $\pm$0.003	0.953 $\pm$0.003

As can be seen from the table, good performance results are obtained already by adding only bidirectionality. These results undergo a slight increase with the addition of the dropout level, which probably decreases the overfitting on the training data.