Report1

* Model Architecture:

1. Embedding with dimension 4x4
2. LSTM 4x8
3. Fully Connected Layer 8 neurons
4. Activation function Tanh
5. Fully Connected 2 neurons
6. Log SoftMax

* Optimizer: Adam with learning rate 0.001
* Loss function: Cross Entropy Loss

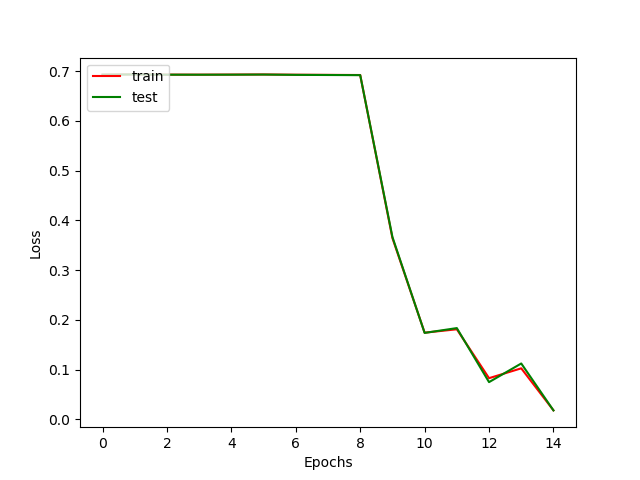
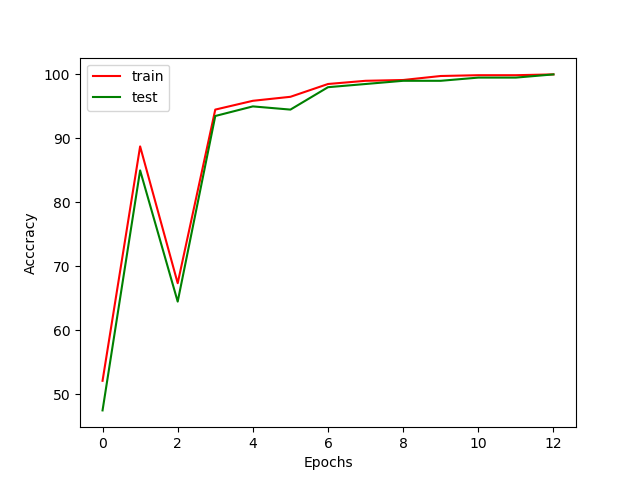
Our model was learned by 1000 different sequences half of them labeled as bad and another one is good. Each sequence had a different length without padding technique was inserted one by one to the model and after each example there was a correction of the weight.

The model succeeded to distinguish two languages in training and in testing sets. We run it multiple times and sometimes it finds pattern in 8 epoch and next time it takes a little bit longer.

First run it takes 12 epochs to separate the data and it took 1 minute and 48 seconds on GPU

Loss

Accuracy



Second run it takes 10 epochs to separate the data and it took 1 min and 16 second on GPU

Loss

Accuracy

