Lea Setruk
Yoel Benabou

Deep Learning - 89687

Ass3 – Part 1

We generate 1000 examples - 500 examples for each language.

We shuffled these examples and split them to two part:
1. Train : 80% (=800)
2. Dev : 20% (=200)

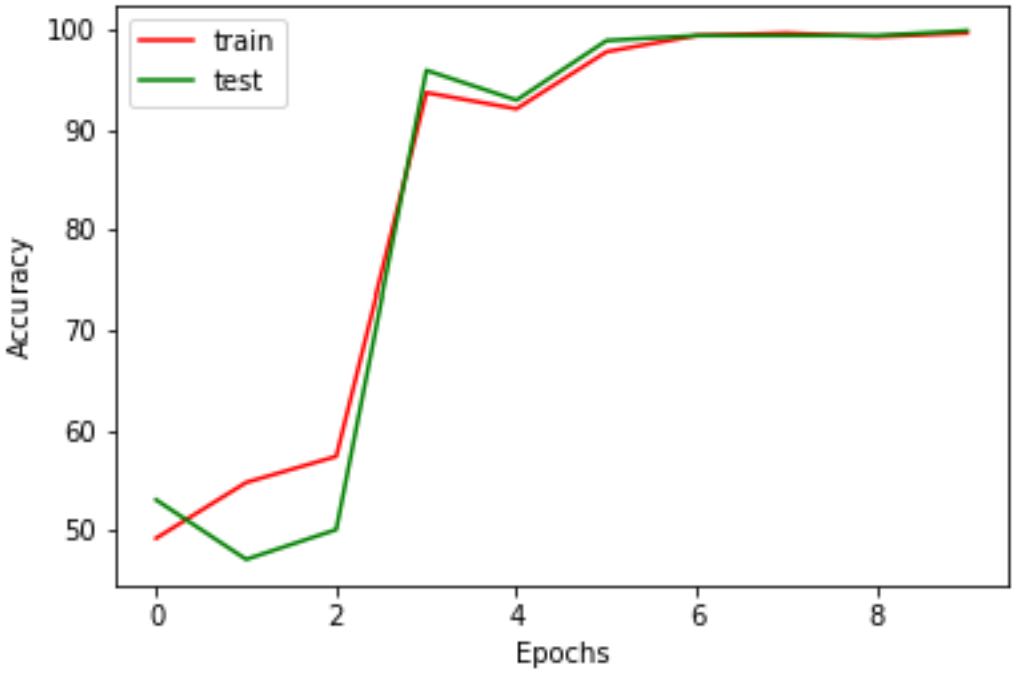
Model

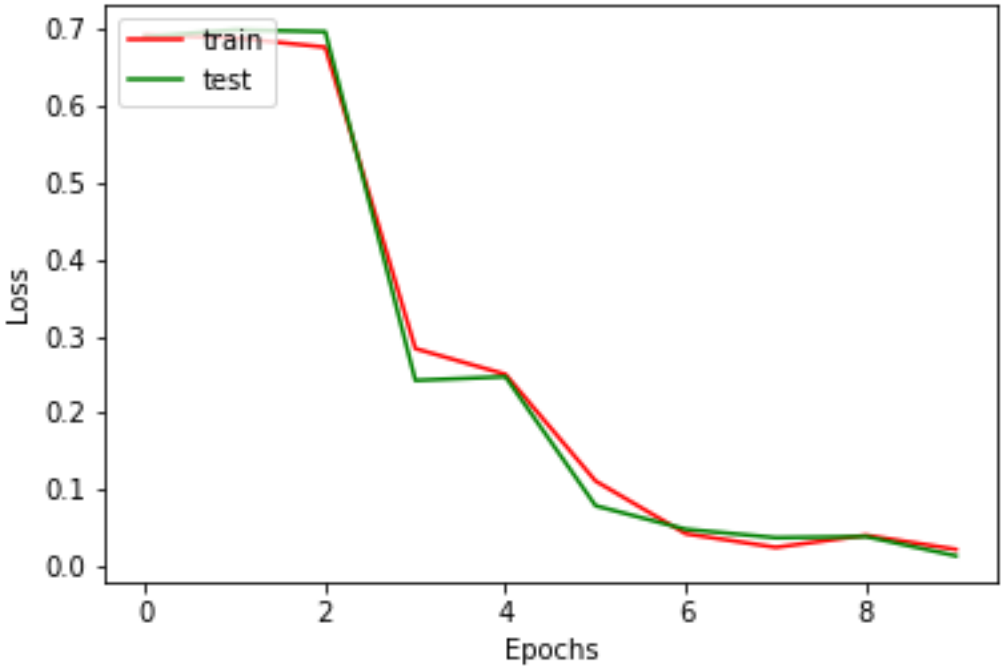
The model is composed by :
1. Embedding (13x13)
2. LSTM (13x8)
3. Fully Connected layer – 8 neurons
4. Tanh activation function
5. Fully Connected layer – 2 neurons
6. Log Softmax function

We use an Adam optimizer with a learning rate 0.001 and cross entropy as loss function.

Results

Our model succeeded in distinguishing the two languages after 10 iterations done in 39.52s on GPU. He got 100% accuracy on the train and the dev data.


train
test
Accuracy
Epochs
100
90
80
70
60
50
0
2
4
6
8


train
test
Loss
Epochs
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8