The training set contained 10000 generated examples 5000 were positive examples and the other 5000 were negative examples. The test set contained 1000 generated examples, 500 were positive examples the other 500 were negative. Before training and testing the data was shuffled. The network succeed in distinguishing the two languages perfectly (1.0 accuracy on training and test) on the training and the validation after 2 iteration on the training data, which took about 21.5 seconds. To make the network work, each character in the sequence was mapped to an embedding vector of dimension of 30, each sequence of embedding vectors was fed into an lstm network with 1 layer in the dimension size of 2, the output of the lstm was fed as input to an mlp network with one hidden layer of size 2 with tanh activation function and an output layer of size 1 with a sigmoid activation function.