Yaniv Sheena

308446764

1) I generated training and test sets, each consist of 10,000 examples.

2) The average length of the sequences is ~140, where the size of each part of the sequence (i.e. [1-9] sequences and letter sequences) is randomly chose in the range: [1,30].

3) My network managed to distinguish the two languages (I.e. it reached 100% accuracy on the test-set) after one iteration, that took 58 seconds (I'm running on a VM with relatively low resources).

4) In order to make my network work:
   a. I used a 1-layer LSTM with input of 50 dimensions and output of 50 dimensions as well.
   b. On top of the LSTM, I built a MLP with 1 hidden layer of 20 dimensions (using tanh as activation function), and output layer consists 2 outputs (binary classifier).
   c. I used learning-rate = 1