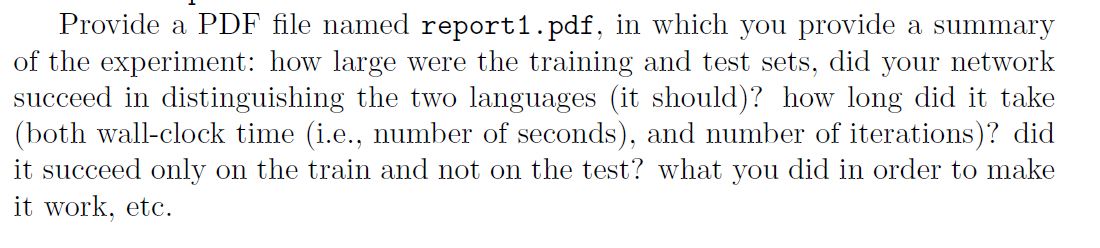
**ID: 011862141**



I used a train files with 1000 samples and a test files with 100 samples. The network was able to capture the pattern as long as the training data segments were approx.. 20 characters long. Its predictions were perfect most of the time in these cases. I did not have to fiddle around with any settings to make it work. The LSTM hidden layer has 10 nodes and the MLP hidden layer has 15 nodes.

With longer segments, when used in the training set, the network failed to make correct predictions. However, when the network was trained on up-to-20-character-long segments, it was able to predict much longer segments – even 50 character segments - in the test. This shows the network training with intermediate-length sequences learned enabled it to learn the correct induction rule. When the training sequences were too short (e.g. 4 characters), such induction did not materialize – the network correctly predicted similar short-segment sequences but failed on longer sequences.

Training time was 33 seconds for 6 epochs (6000 cases). When the network succeeded, it was usually close to 100% success rate by the end of the second epoch.