CoE202 - Basics of Artificial Intelligence “Big data analysis and machine learning”

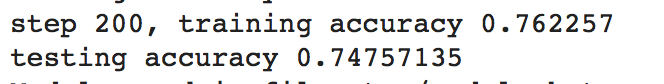
Mukanov Ayan 20170881

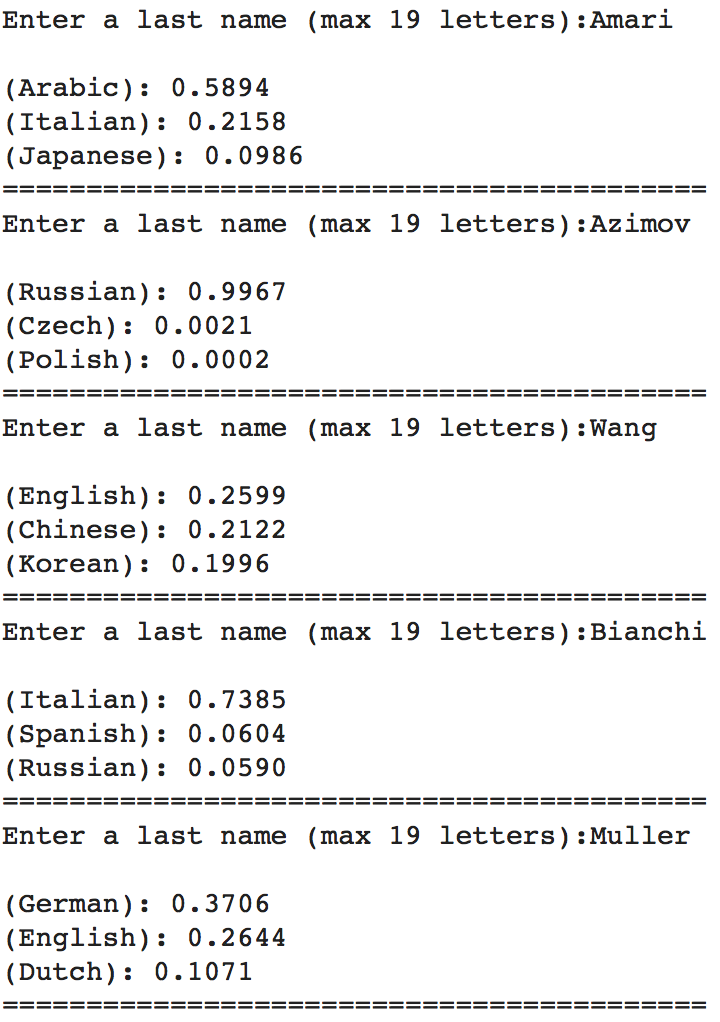
Activity #5 Report

**Name Classification**

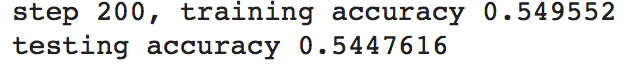
**Part a: Use GRU, LSTM and Simple RNN functions for training. Compare each of results**

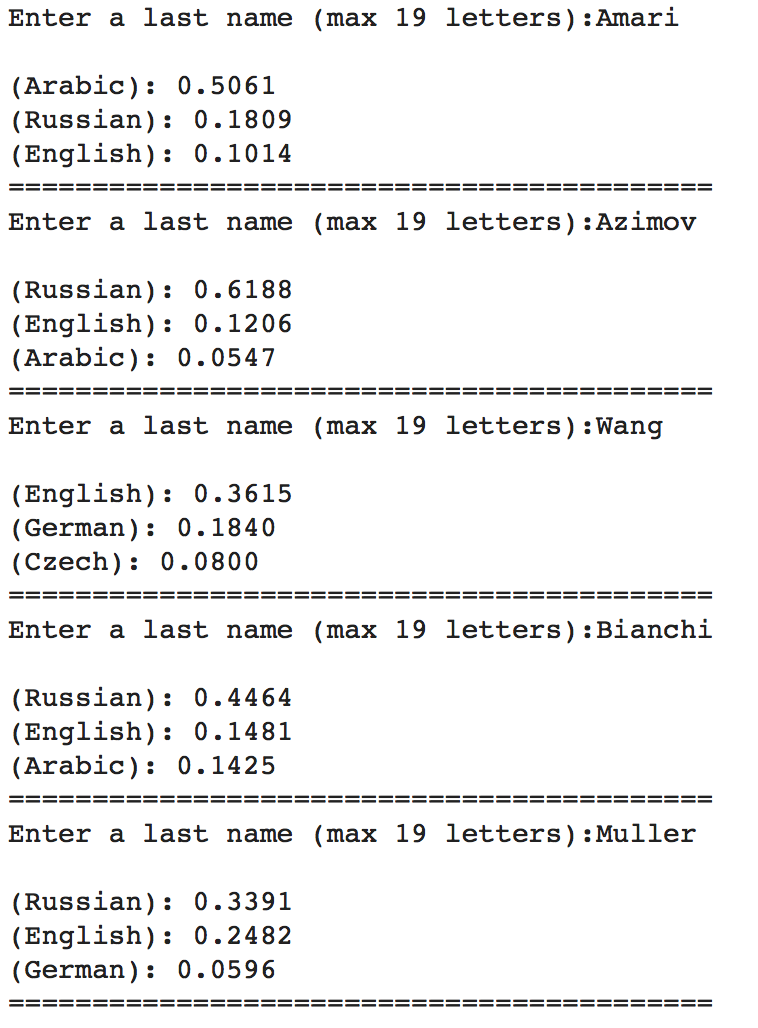
1. GRU



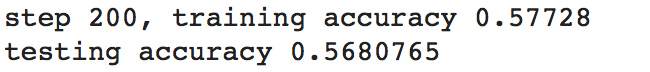


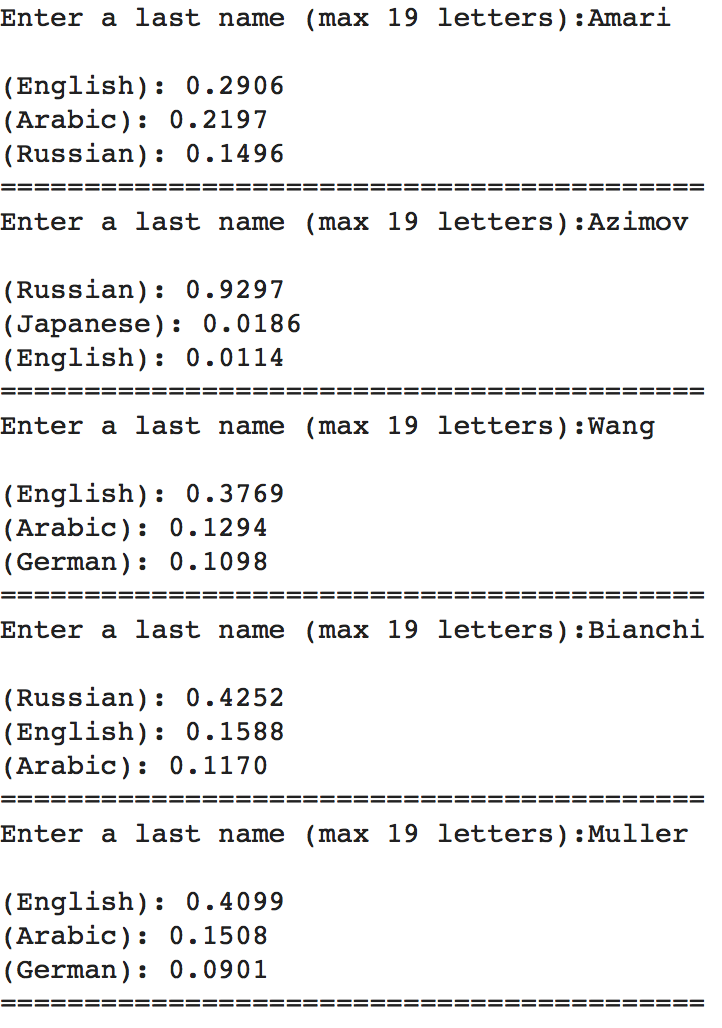
1. Simple RNN





1. LSTM





Discussion:

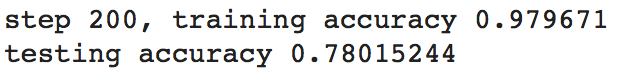
It is clear that GRU outperformed all other models. The second best model was LSTM and the worst – RNN.

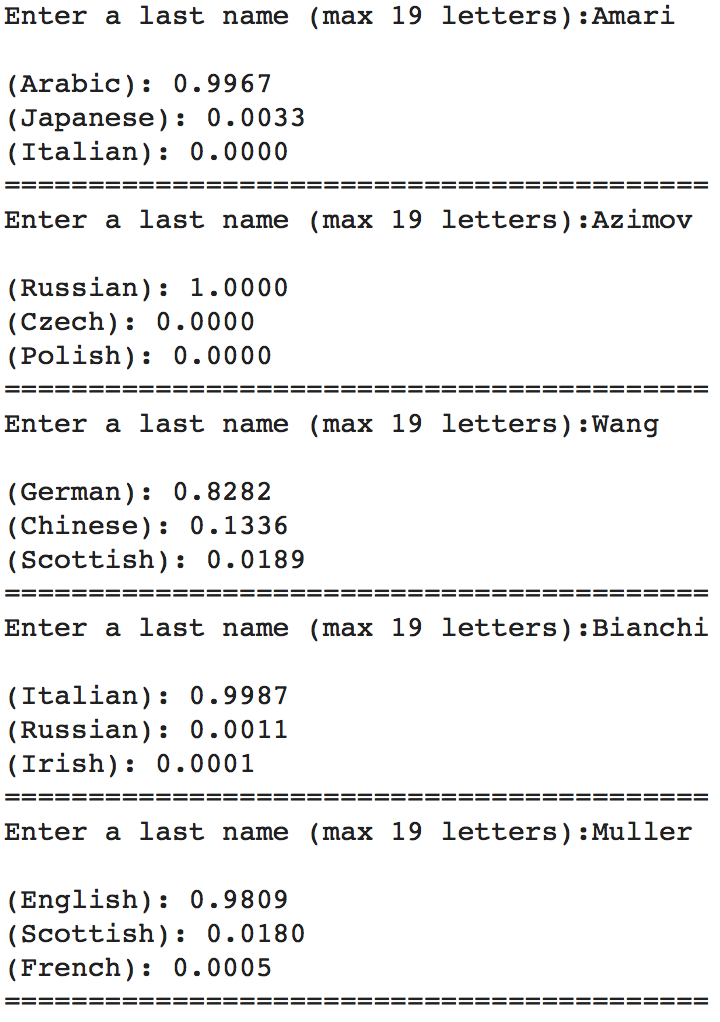
The reasons for such results could be numerous. For RNN, the short-term memory might have resulted in worse results as when sequence becomes sufficiently long, RNN fails to hold all data in memory. Also, the learning process could also be influenced negatively when gradient value becomes small.

Given the simplistic nature of our RNN model, the results obtained are within our expectations, performance wise. LSTM model performs a tad better as it addresses some of RNN’s shortcomings, such as shorter memory problem and the gradient problem, since it has gates that determine what data to hold or not.

Finally, GRU, being quite similar in nature to LSTM, now has fewer number of operations and also has a reset and update gate. So, in my project, possible due to these reasons GRU outperformed other 2 models.

**Part b: Replace RNN with DNN**

****



Conclusion:

It is strikingly clear that the DNN that we implemented shows the best results out of all models so far. There’s also the mere fact of overfitting, which is observable from the testing accuracy being quite less than the training accuracy.