**Project Notes**

Jashua Luna

August 30, 2021

At this point in the project several models have been trained and their accuracy evaluated. Independent of individual model’s performances, performance is better on data used for training, than on data outside of the training set. With a trading performance simulator, the best performing model has a trade accuracy of 55-60% when decisions are made with the value taken at -prediction length + 1 from the prediction. The same model has a prediction accuracy of 30% in data from outside the training set. This is mitigated by changing the value with which decisions are made, though best accuracy is still only 40%.

Perhaps this is because the model is overfit to training data. 500,000 samples were used to train the best performing model and it is possible trades are made too frequently. Redundant predictions might alleviate some losses resulting from trades made from lapses in prediction accuracy, and a larger stopping function window in training data generation.

Also consider training a model on t – 7days and evaluating that model on t + 1 day data. That way, the data the model is trained on is like the data it performs on. If this proves viable, a new network can be trained daily for use on that day. Ideally, the model will continuously learn as it makes trades thus preventing other problems previously encountered (more on that in a later report, or earlier).