**CS 205 FINAL PROJECT PROPOSAL**

Parallelizing Neural Network with Improved Performance

Linglin Huang, Chang Liu, Greyson Liu, Kamrine Poels

**Background**

Despite the availability of high-frequency stock market data, its use in forecasting stock prices is studied to a lesser extent. Similarly, despite the recent success of neural network on as a forecasting method, its power in forecasting high-frequency dynamics has been relatively overlooked. In addition, most of the studies in the literature have been focused on stock market indices instead of individual stocks. A possible explanation is the intractable computational intensity of training neural networks on the massive volume of high-frequency data of individual stocks. This motivates our study on applying parallelism to the training task and evaluate its performance to demonstrate weak and strong scaling.

**Methodology**

We formulate the task as a prediction problem, using lagged previous prices of individual stocks to predict future prices at the minute level. The high-frequency consolidated trade data for the US equity market comes from NYSE Trade and Quote (TAQ) database, available by the WRDS research center.

For the prediction method, multi-layer Artificial Neural Networks (ANN) using back-propagation algorithm has shown promising results in stock index prices compared with traditional methods [1]. Note that the gradient descent algorithm of back-propagation is sequential by nature. We will therefore apply a technique that uses MPI and OpenMP to parallelize the training process: mini-batch or Stochastic Gradient Descent method using the SIMD-BP algorithm [2]. The optimization metric for this ANN will be the squared error loss between predicted price and true stock market price. Using a similar method to Downpour SGD [3], we store parameter values in a parameter server while model replicas (each replica in one node) execute asynchronously and fetch/push updated parameters from and to the parameter server. Within each model replica, we intend to further parallelize the training of the ANN by implementing OpenMP. Lastly, we benchmark our approach against an analogous sequential version using various performance metrics on accuracy and efficiency.

**References**

[1] Selmi, N., Chaabene, S., & Hachicha, N. (2015). Forecasting returns on a stock market using Artificial Neural Networks and GARCH family models: Evidence of stock market S&P 500. Decision Science Letters,4(2), 203-210. doi:10.5267/j.dsl.2014.12.002

[2] Valafar, Faramarz, and Okan K. Ersoy. (1993). A Parallel Implementation of Backpropagation Neural Network on MASPAR MP-1.

[3] Dean, J., et al. (2012). Large scale distributed deep networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, Curran Associates Inc.**:** 1223-1231.