

Title: ML Crypto-Trading Algorithm

Team: Raphael Iskra (519, Yang), Brendan McCaffrey (519, Yang), Karan Jaisingh (419, Shubham).

Task T: We propose to predict crypto performance in intervals (10 mins, 1 hour, 1 day), translated into binary output (0/1 equating to sell/buy), aiming to perform better than HODL strategy.

Experience E: To do this, we will import time series data (open, high, low, close, volume, network news) from Coinbase Pro API as well as BTC futures data (potentially also pull LIBOR/treasury yield data from Yahoo Finance). We will train our model with past 7 years of historical data for several cryptocurrencies (BTC, ETH, LTC, DOGE, etc). We will use k-fold cross validation across this time series, alternating our test set. Eventually, we hope to run the algorithm on live Coinbase data, using real-time transaction data as a test set.

Performance metrics P: We will measure the mean squared error to evaluate our error when predicting the future price of a cryptocurrency. We may also use metrics like accuracy and R^2 score to run additional evaluations based on whether or not a positive/negative price change is predicted accurately. We will use the HODL (always predicting BUY interval) strategy as a benchmark to analyze the algorithms performance for each respective cryptocurrency.

ML Methods: We plan to approach this problem through a variety of methods, evaluating the results at each stage and going further with more successful methods. Methods we plan to use include regular regression algorithms (SVMs, Linear Regression, Decision Trees), as well as Deep Learning algorithms like Neural Networks, LSTM Networks and 1-D Convolutional Neural Networks. Using this variety of methods essentially allows us to experiment with different types of models, from fairly linear ones to multi-layered ones to time-series ones. We also plan to use many ML-affiliated techniques in the realms of data preprocessing, cross-validation and testing.

Why we care about this: Brendan previously created technical analysis strategies (non-ML) for crypto trading, and is long-time bullish on BTC. Karan has an interest in using quantitative/ML methods to predict market behavior, and is intrigued to investigate ML capabilities on blockchain networks. Raph is an active crypto trader, and is excited about the idea of allowing an ML algorithm to automatically execute his trades.

Expected Challenges: Given the nature of the data - it is in the form of a time series, hence making it important for us to dive into novel approaches that don't utilise standalone data, as it must take a degree of sequence into account. As the cryptocurrency world is also fairly new, a lack of training data may also be a foreseeable issue, as there doesn't exist decades of data.

Work Plan over the next ~5 weeks: First, we will study relevant literature on similar projects, and collaboratively make the decision of whether pulling Coinbase api data is appropriate to get a large training dataset, or if other historical datasets need to be explored. Once the data is procured, we will try out several preprocessing methods. This includes both removing malformed and null entries in the dataset, and iteratively deciding which features will be used in our classification algorithm. Finally, we will need to decide on what proportion of the dataset will

be used for training, and what amount of it will be used as testing data. Once these points are decided, we need to devise a method of logistic regression that allows us to classify elements of our dataset as either buy or sell moments. Finally, we will decide what statistical analyses will be appropriate to decipher the accuracy of our model and how it compares to the control group of a steady holding of a cryptocurrency position.

Prior Work / References / Projects that inspired us (this part can be over the 1-pg limit):

1. Laura Alessandretti, Abeer ElBahrawy, Luca Maria Aiello, et al. "Anticipating Cryptocurrency Prices Using Machine Learning". Link: <https://www.hindawi.com/journals/complexity/2018/8983590/>
2. Mohil Maheshkumar Patel, Sudeep Takwar, et al. "A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions". Link: <https://www.sciencedirect.com/science/article/abs/pii/S2214212620307535>
3. Wharton startup, [Alphagora](#), has an autonomous crypto-trading platform.