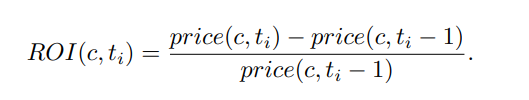
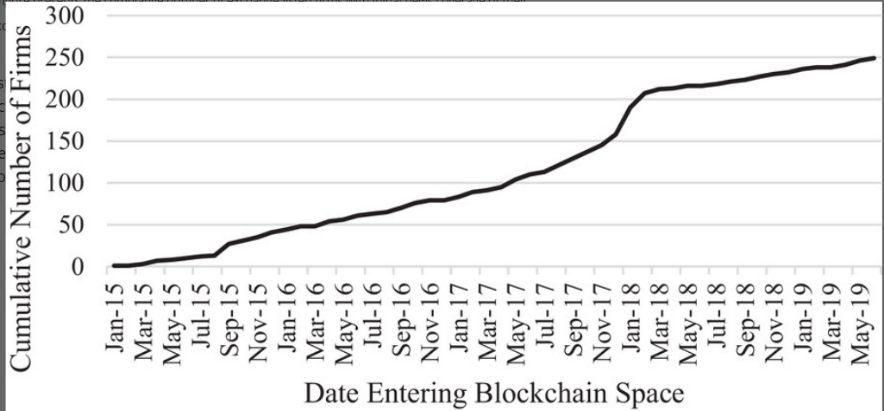
Research:

* Anticipating Cryptocurrency Prices Using Machine Learning:
  + exploiting inefficiency in crypto market to gain profits
  + crypto skyrocketed in 2017 due to several months of exponential growing market capitalization.
  + DEFINITION: market capitalization, the total dollar value of a company’s outstanding shares. (share priced x total outstanding shares)
  + DEFINITION: total outstanding shares, the number of issued shares minus the number of shares held in the companies treasury. (for crypto this will be the amount of the currency minted)
  + Long term properties of crypto marked have remained stable between 2013 and 2017
  + Best ML predictor for the price fluctuation of BTC was neural network.
  + Deep reinforcement learning was showed to beat the buy and hold method in predicting the prices of 12 cryptocurrencies over one year period.
  + DEFINITION: trading volume, number of coins exchanged in a day
  + Coin market cap daily prices is computed as the volume of weighted average of all prices reported at each market
  + Coin Market Cap data information (not really relevant to me at the moment) The website lists cryptocurrencies traded on public exchange markets that have existed for more than 30 days and for which an API as well as a public URL showing the total mined supply are available. Information on the market capitalization of cryptocurrencies that are not traded in the 6 hours preceding the weekly release of data is not included on the website. Cryptocurrencies inactive for 7 days are not included in the list released. These measures imply that some cryptocurrencies can disappear from the list to reappear later on. In this case, we consider the price to be the same as before disappearing. However, this choice does not affect results since only in 28 cases the currency has volume higher than 105 USD right before disappearing (note that there are 124,328 entries in the dataset with volume larger than 105 USD).
  + Profitability of the currency is quantified using return on investment (ROI), article considers return on investment after 1 day (short-term trading)
  + DEFINITION: Return on investment (ROI), ration between net income and investment
  + For volatile numbers, the geometric average provides a far more accurate measurement of the true return by taking into account year-over-year compounding that smooths the average
  + Methods performed better when prices were based on btc as opposed to usd (interesting)
* Blockchain speculation or value creation? Evidence from corporate investments
  + When a new institution publicly announced its crypto holdings (or any new tech they will be involved with in crypto) the stock price reaction grew 13% on average and then reversed in 3 months
  + Credible investments into the crypto space are defined as such:
    - First, credible investments are those at an advanced stage in which the firm is currently using or will imminently use an existing or new blockchain for a commercial purpose (i.e., blockchain is or will imminently be integrated into the firm's business operations), as opposed to firms at a preliminary stage such as those currently studying potential applications, involved in initial joint collaborations with other firms, or joining research consortiums to study the technology's potential uses.
    - Second, credible investments are those in which the firm specifically mentions the term blockchain in its quarterly or annual financial statements (10-Q or 10-K) at the end of the announcement quarter or in the next three quarters. This second measure is ex post, as these reports become public on average 4 months after the initial blockchain announcement and therefore do not directly influence the announcement reaction. These measures of credibility reasonably reflect the firm's level of commitment to its blockchain investment.
  + Institutional investment spikes in July to September 2015 and in November to March 2018 (If we look at this chart along side the btc price it looks like institutions will buy the dip or get involved during the dip.
* Cryptocurrencies As an Asset Class? An Empirical Assessment
  + short term market activity is highly correlated to investors sentiment
  + There is only a mild, and not significant, correlation between returns on cryptocurrencies and returns on traditional asset classes on a daily basis. • Past returns significantly drive trading volume, consistent with the idea that short-term market activity is primarily driven by sentiment. • Macroeconomic factors such as the term structure of interest rates and inf lation expecta-tions do not seem to affect market activity in either the short or the long term.
* What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis
  + data used:
    - btc price
    - blockchain data (btc) data, it might be useful to try and see if there is a lag in the btc data to the eth data found on blockchain.info
    - google trends
    - financial stress index (global overall value)
    - gold prices
  + In economic theory price level of a currency is standardly driven by its use in transactions, its supply and its price level.
  + this study uses wavelet analysis to find correlation between btc price and different fundamental data points.
  + As a measure of transaction use (demand for the currency) they use the ratio between trade and exchange transaction volumes, Trade-Exchange. Usage should be leading the price.
* Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero
  + market beta (measure of volatility), trading volume and volatility are significant determinants for the currencies price
  + DEFINITION: market beta, a measure of individual stock risk relative to the overall volatility of the stock market, it is calculated using Capital Assets Pricing Model.
  + Diagram

    Description automatically generatedCrypto vs fiat, not backed by government, not centralized, price of crypto is dependant on what the investor is willing to pay for it. All member are equal with btc no central server tells everyone what to do. Difficulty of mining is due to the length of the hash?
  + weak correlation between the S&P500 and the crypto market
* Bitcoin Volatility Forecasting with a Glimpse into Buy and Sell Orders
  + DEFINITION: order book, An order book is the list of orders that a trading venue uses to record the interest of buyers and sellers in a particular financial instrument.
  + uses volatility history and order book data to predict btc price
  + In data mining and machine learning models areas, [16], [17] used the historical price time series for price prediction and trading strategies. [15], [18] utilized social information like the sentiment, comments, and replies on forums to forecast price fluctuations. [14] explored the predictive ability of Blockchain information for Bitcoin price. As for volatility prediction, [19] evaluated the performance of GARCH models on Bitcoin. However, volatility forecasting using order book information of Bitcoin is still under-researched. (Go through all this research to demonstrate reasoning behind the choice of data for our agent to consume).

• *spread* is the difference between the highest price that a

buyer is willing to pay for a BTC (bid) and the lowest

price that a seller is willing to accept (ask).

• *Ask/Bid Depth* is the number of orders on the bid or ask

side.

• *Depth difference* is the difference between ask and bid

depth.

• *Ask/Bid Volume* is the number of BTCs on the bid or ask

side.

• *Volume difference* is the difference between ask and bid

volume.

• *Weighted spread* is the difference between cumulative

price over 10% of bid depth and the cumulative price

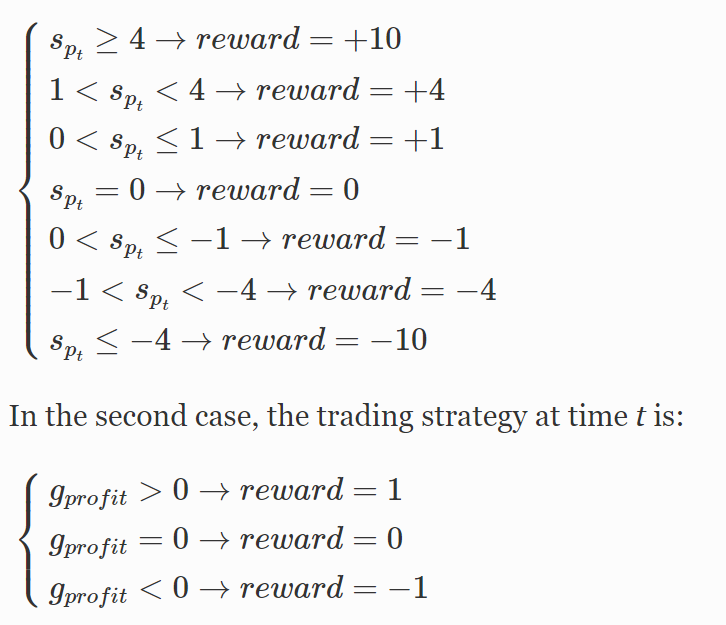
over 10% of ask depth.

• *Ask/bid Slope* is estimated as the volume until \_ price

offset from the current traded price, where \_ is estimated

by the ask (or bid) price at the order that has at least 10%

of orders with the higher ask (or bid) price.

* Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis
  + Tweet volume along with google trends data about the asset is a better price predictor than tweet sentiment.
  + Google trends and tweet volume will increase and decrease with the cryptocurrencies price
  + Used Tweepy to filter out hashtags followed by “bitcoin” and “Ethereum” (this was just to get the actual tweets for sentiment analysis and not for tweet volume)
  + Google trends provides a search volume index which is a number scaled between 0 to 100 on a search term’s proportion to all searches on all topics. When trends data is queried for a period of longer than 90 days the SVI returned are aggregated at a weekly level. (IS THIS A REAL ISSUE THAT NEEDS TO BE SOLVED?)
  + Site with more data to potentially include in the trading bot (<https://bitinfocharts.com/>)
  + Pearson R to look for correlation in crypto prices and google trend data. Found high correlation (0.817 with pvalue of 0.000)
  + sentiment of tweets were useless
  + tweet volume had a strong correlation with the price of btc (0.841 pvalue 0.000)
* Sentiment Analysis of News for Effective Cryptocurrency Price Prediction
  + showed good results
* Recommending Cryptocurrency Trading Points with Deep Reinforcement Learning Approach
  + DRL agent that acts properly in the market and maximizes the trader profits. Agent starts by taking a random action a1 (buys) and then the next move is to sell (a2) (include the trade fee in the calculation). The reward is estimated after selling by subtracting the buy price by the sell price. If the result is positive the agent gets a positive reward. If the result is negative, it gets a negative reward.
  + They only consider hourly stock prices of the crypto but we want to add much more information so that the agent has the best chance to learn off the same information any trader would be able to have access to.
  + Need to divide the data into teaching data and testing data to see if the agent knows how to play the game properly. They didn’t have a sell all option. We will be selling all of our holdings and buying with all of our money (this simplifies things but might need to be reconsidered if the bot does not work as planned)
  + 41% profit trading eth but just using the historical price data
* A Deep Reinforcement Learning Approach for Automated Cryptocurrency Trading
  + Double Deep Q Network (target network)
  + Deep Q Network
  + Dueling Deep Q Network
  + Uses stop loss (-5%) and take profit at (12%) (they don’t mention why they used these numbers. Seems like a too defensive approach to trading with the bot)
  + transaction cost of 0.3% is used
  + rewards using Sharpe Ratio (Spt) and the nominal return (4 models total were created)
  + DDQN uses CNN and ADAM for weight optimization. MSE for the loss function. Discount of 0.98 and RELu activation function.
  + The simple DQN has the worst results over all the testing periods (going to need to use a more complex model for the algorithm)
  + *Sharpe*D-DQN has demonstrated to be the best Q-learning trading system
  + They don’t test the method against the buy and hold method
  + a parameter optimization should be done to improve the performance of the learning techniques.
* **Deep Reinforcement Learning in Cryptocurrency Market Making (USE PIECES OF THIS PAPER TO EXPLAIN THE ORDER BOOK DATA THAT WILL BE USED IF ACCESS TO THIS KIND OF DATA CAN BE FOUND)**
  + Market making and limit order books with reinforcement learning (policy gradient method)
  + forward feed neural network
  + DEFINITION: market making, provide liquidity to a marketplace. They make money from the spread of bid and ask price
  + DEFINITION: limit order books, the price a trader is willing to pay for an asset
  + DEFINITION: order flow imbalances, the number and frequency of orders received by the limit order books
  + DEFINITION: trade flow imbalances, to the difference between market buy and sell orders during a given period
  + Partially Observable Market Decision Process (POMDP) A POMDP models an agent decision process in which it is assumed that the system dynamics are determined by an MDP, but the agent cannot directly observe the underlying state.
  + uses the actor critic method
  + policy gradient method is more effective at learning the dynamics of trading than Q-learning, a value-based method [1, 5, 8, 11, 12].
  + Other findings show that using risk-based metrics, such as the Sharpe Ratio (or other variations like Differential Sharpe Ratio), for the reward function (i.e., feedback signal) outperform algorithms that use the more intuitive profit-and-loss reward function [5,8,12]. These papers only used price returns though not order book data.
  + Proximal Policy Optimization Algorithms: the surrogate function, not the policy, is optimized directly. This surrogate function removes the incentive for a new policy to depart from the old policy, thereby increasing learning stability. Kind of like the target network in DDQN.
  + Actor-Critic with Experience Replay (ACER) was originally in scope for this paper, but was de-scoped due to poor performance after demonstrating instability and slow learning in comparison to A2C and PPO.
* Neural Networks with Online Sequential Learning Ability for a Reinforcement Learning Algorithm
  + mRan (a potential algorithm to use and compare with the other ones)
* An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information
  + data used
    - prices
    - volume (done)
    - trading volume (done)
    - avg. block size (done)
    - transactions (done)
    - median confirm time
    - hash rate (done)
    - difficulty (done)
    - cost of transaction (done)
    - miners revenue
    - confirmed transaction
    - total number of unique btc addresses (done)
    - Macro economics data: S&P500, Eurostoxx, DOW30, NASDAQ, Crude oil, SSE, Gold, VIX, Nikkei225, FTSE100 (done)
    - global currency ratio: GBP, JPY, CHF, CNY, EUR
    - These results provide empirical evidence for the fact that the recent volatility in Bitcoin prices stems mostly from the Blockchain information directly involved in supply and demand of Bitcoin and not from other macro-financial markets.
    - Several statistical problems are caused from the multicollinearity which is the situation that some regressors have a linear relationship with other regressors. It can cause undesirable regression analysis: very high R2 for some coefficients that are not statistically significant and their t-statistics sensitive to data variation [33]. One of the prescriptions for dealing with multicollinearity is to do a linear regression except for variables with large VIF values, which is a sort of measure of the linear relationship between variables [33]. (THIS WILL NEED TO BE DONE FOR MY CORRELATION ANALYSIS)
* Reinforcement Learning with Self-Attention Networks for Cryptocurrency Trading
  + self-attention network
  + Thus, an episode in the environment consists of a specific number of days, in which the agent is allowed to buy and sell assets at every 24-h time interval.
* Application of deep reinforcement learning in stock trading strategies and stock forecasting (2019)
  + using DDQN with replay memory (batch processing)
  + Dueling Double Deep Q-Network (DQN using an advantage function)
* Adaptive stock trading strategies with deep reinforcement learning methods (2020)
  + Gated Recurrent Unit to extract feature information from raw financial data
  + Reward functions with risk-adjusted ratio for trading strategies for stable returns in volatile conditions
  + Gated Deep Q-learning trading strategy
  + Gated Deterministic Policy Gradient trading strategy
  + These trading strategies had more stable returns in volatile markets for stocks (might be worth looking into since crypto is very volatile
  + GDPG with an actor critic framework had the best results
  + uses Sortino Ratio for the reward function
  + they use technical analysis and stock prices for the predictive algorithm (seems kind of stupid because ta draws on prices so they would all be somewhat correlated in my opinion)
  + has experience replay memory
* Modern Reinforcement Learning Algorithms

Reinforcement Learning in Stock Trading

Predicting fluctuations in cryptocurrency transactions based on user comments and replies

Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model

Volatility estimation for bitcoin: A comparison of GARCH models

calculating market beta for Ethereum: If we apply the concept of the beta coefficient to

cryptocurrency markets, Bitcoin could be used as the benchmark. So one could calculate the beta for BNB or other altcoins in relation to Bitcoin’s price and volatility. Alternatively, Bitcoin’s volatility could be measured against gold or stock markets. The resulting beta would give insights into the correlation between Bitcoin and traditional financial markets.