Business Understanding

✓ Identifying your business goals
The goal of the project is to train a machine learning model to know when to buy and when to sell cryptocurrency when given the values of a couple of technical indicators.. The end-goal of the project would be to have the bot earn more than the base market performance. To do this our model will need to adequately predict when the price is about to start rising so it can send buy signals for buying the currency and sell signals before the price starts dropping. The signal can be either in the form of a confidence rating (between 0 and 1) or a discrete value (-1 for sell, 0 for hold, 1 for buy).

Assessing your situation

Resources
- The TradeBot codebase, which can be used to collect raw aggregated trade price data from Binance and to calculate the values of the included technical indicators based on the collected data.

- A workstation PC with a Ryzen 3900X, 32 GB of RAM, and GTX 1080TI.

- 3 school laptops.

- 500/500 Mbps internet connection

- Weights and Biases developer tools.

- Publically available examples of previous attempts. This includes learners for both stocks and cryptocurrency prices.

Requirements, assumptions, and constraints
- Needs to be completed by the course project deadline.

- No legal or security obligations.

The project can be considered successful if the produced model is able to generate accurate enough buy and sell signals based on real-time market price changes to generate more profit than the market performance baseline. We should be able to hook the trained model binary to the TradeBot codebase and use it with a live Binance API key to trade actual cryptocurrency and generate real profit. The network or ensemble should be sparse/simple enough to be calculated in real-time to respond to live market price updates in a timely manner.

Risks and contingencies
In the probable scenario that our model does not meet the success criteria we will use this project to further analyze why machine learning algorithms are ill-suited for analyzing the price movements in currency and/or commodity markets.

Terminology

- Technical indicator - A statistical or mathematical tool used to analyze trends and patterns in price movements. These are usually calculated based on candlestick data (such as closing price). We are using 5-minute candles for this.
- RSI - Relative Strength Index. Intended to chart the current and historical strength or weakness of a stock or market based on the closing prices of a recent trading period. (https://en.wikipedia.org/wiki/Relative_strength_index)
- MACD - Moving Average Convergence Divergence. (https://en.wikipedia.org/wiki/MACD)
- DBB - Double Bollinger Bands. A type of statistical chart characterizing the prices and volatility over time of a financial instrument or commodity, using a formulaic method propounded by John Bollinger in the 1980s. Double Bollinger Bands means that we use two Bollinger Bands. One calculated using one times standard deviation, the second one using two times the standard deviation. (https://en.wikipedia.org/wiki/Bollinger_Band)
- Market performance - Given a time period, the market performance for that period will be the cryptocurrency price at the end of the period minus the price at the beginning of the period. In other terms, it is the average price change over the period.
- Bot performance - Given a time period, the bot performance will be the total account cryptocurrency wallet value (in USDT) at the end of the period minus what it was at the beginning of the period. In other terms, it is the total profit of the bot over the time period.
- Fiat cryptocurrency or stablecoins - The relevant example in this project is Tether (USDT). These are low-volatility cryptocurrencies commonly used to trade other cryptocurrencies as opposed to buying them with physical currencies like USD. All physical currency transactions would have to go through a payment provider and are often subject to transaction fees. This is (usually) not the case for trading cryptocurrencies against fiat currencies.
- Cryptocurrency market symbols - Each cryptocurrency has an abbreviated market symbol. This is also used to identify it in API calls and transactions. Examples: Bitcoin - BTC, Tether - USDT, Ripple - XRP, Ethereum - ETH. Cryptocurrencies are exchanged in pairs (just like physical currency exchanges). This is why the market symbols also commonly show up in pairs to denote the exchange rate between two cryptocurrencies. Examples: BTCUSDT, BTCETH, XRPUSDT.
- Aggregated trades - This is the ticker criteria used to update the live price of a currency pair on Binance. An aggregated trade is aggregated to a single taker order. This means that if N makers are offering a cryptocurrency at X price, an aggregated trade event is pushed when one taker orders the cryptocurrency at X price from a subset of the N makers to fill the amount they want to buy. How frequently this updates depends on how many users are trading the cryptocurrency pair at any moment. For our one-year BTCUSDT price data an aggregated trade price change occurs on average every 350 ms, and every 10-30 ms during high volume trading periods.

Costs and benefits

The only costs that will be sunk into the project are our time and maybe the use of Weights and Biases to help improve the development process. The benefits if the project is successful would be a bot that is able to trade on the Binance cryptocurrency exchange and

generate profit. Given enough time or enough of a starting budget, the bot would be able to provide much more benefit than our meager costs.

✓ Defining your data-mining goals
We already have everything we need in terms of data and we can always generate more. The TradeBot multithreaded collector is able to download raw price data from the Binance API at the internet speed cap (500 Mbps for one member), and our workstation is able to process 2 GB of this data into indicator values in less than 2 minutes so getting more data is trivial.

✓Data-mining success criteria
The collected market price data and indicator values must be accurate. This means that they must correspond to what is shown on Binance. This is especially true for the indicators, since we calculate them ourselves in the TradeBot code. There must be no other gaps in the data beside Binances maintenance and everything has to be in chronological order.

✓ Producing your project plan

List of tasks
- Learning about different models that could possibly work and how the parameters work - 6 hours per person

- Discussing within the group about which models to use - Planning to have a 2-4 hour meetup where we will discuss who tries what model

- Coding and testing the different models - on average 20 hours per person

- Implementing and training the final leaner (maybe ensemble) - 15 hours per person

- Validating against the entire 2019 and more recent datasets - 2 hours per person

✓Methods and tools we are planning to use
We are planning to initially write the project in Java. Assuming that it works and we don't have to resolve to using python we can use an agile development practice so that each team member can try at least one model. We will start researching different models as our first step.

We are also planning to try Weights and Biases.