Data Requirements
We don't have much requirements besides the data (market price and indicators) being accurate.

The raw price data isn't suitable for learning. The magnitude of the price movements will be different for each cryptocurrency pair. In order for the learner to be able to generalize across multiple market conditions (such as low or high volatility and different cryptocurrency pairs) we will need to provide the learner with normalized technical indicators. The TradeBot project already provides some of the most commonly used indicators such as Relative Strength Index, Moving Average Convergence Divergence, and Double Bollinger Bands. By calculating the values of these indicators over a certain period of time we can provide the learner with the necessary data. Other technical indicators might also prove prudent but are out of scope right now.

- RSI - Double value (0 - 100). Time required 0.
- MACD - Double value. Need to find a way to normalize it.
- DBB - Integer (0 - 5).

The availability of the data is not a concern.
We don't need to invest any time into gathering the data because we already have it on hand. We have gathered and processed all of the aggregated price data for the BTCUSDT pair from 01.01.2019 to 01.01.2020. The total size of the raw price data csv file is 2 GB.

We have Bitcoin (BTC) and Tether (USDT) price data spanning exactly over one year. The data was gathered as an Object Oriented Programming course project.

If we find that we need more indicators (for the models to work better), then we can always implement it and re-generate data.

We don't have any criteria besides accuracy, and since we have generated and automatically checked the data ourselves, we know that the data is indeed accurate.

The quality is as good as possible, there are some inconsistencies because we have generated the data using https://www.binance.com/en and its API so when there has been maintenance we have a gap, but this will most likely not affect our project. The data is and has to be in chronological order.

Once we are done generating the data, there is a method to check whether everything is in chronological order and that the gaps are only caused by Binance maintenance. There cannot be any other gaps.