Week 1 (01/03/2021 - 05/03/2021)

When I first started my internship, I immediately received a warm welcome and some more information about my assignment. The first week consisted mainly of general introductions to both Brainjar and other students.

Furthermore, during the week I worked mainly around feature engineering. I plan to go through the necessary documentation this weekend, in order to start programming my model.

Week 2 (08/03/2021 - 12/03/2021)

During my second week at Brainjar, I mainly considered which models or techniques would be most suitable for my assignment. I thought it would be a good idea to start developing an algorithm for binary classification first, so it may serve as a foundation for further solutions. I have compared just about every model from scikit-learn to get a better understanding of the most interesting algorithm suited for my problem.

After going through a series of data transformations, I currently have a machine learning algorithm with a precision of 75%. By next week I plan to add more data and make a possible switch to artificial neural networks, in order to achieve better statistics.

Week 3 (15/03/2021 - 19/03/2021)

My third week of internship was mostly about deep learning. I tried to make a switch from machine learning to deep learning. As the week ended, I concluded that I do not have enough data to experience the benefit of artificial intelligence. After this I decided, together with Brainjar, to try to realize the same thing, but with different data.

Before this, my data came mainly from my previous analyses, now my data comes from the exchange itself (Binance) or through Alpha Vantage API.

Week 4 (22/03/2021 - 26/03/2021)

The fourth week of my internship was a very productive one. My long-term memory model is starting to gradually take shape. Currently, my model is able to predict the sentiment of the next session by making an estimation of ohlc.

Toward next week, I'm going to try to predict both flow and sentiment. However, this is a trickier story because you always have to take the previous prediction as input for the next one. Next week promises to be an interesting one!

Week 5 (29/03/2021 - 02/04/2021)

During my 5th week, I made some changes regarding the input/output of my model. I no longer predict the ohlc per candle, but instead I have now focused on the closing price. The advantage of this is that it allows my model to achieve better overall precision.

Further through the week, my presumptions were confirmed that it is not a good idea to make a prediction on an earlier one. Instead of doing that, I transformed my model's output to 7. The coming weekend will be mainly devoted to additional validation, in order to confirm the quality of the model.

Week 6 (05/04/2021 - 09/04/2021)

Week 6 was mainly about debugging. We could not really figure out why our training loss converged to zero almost immediately. It was suggested that leakage would be present. After many hours of debugging and rewriting my model, I have not been able to find any signs of leakage.

Later on, we realized that it had nothing to do with leakage but rather with the data being processed. It is simply because of the price volatility that goes along with cryptocurrencies.

Week 7 (12/04/2021 - 16/04/2021)

During the seventh week of my internship, I encountered some unique problems. The problem was mainly in the way I built up my training and test set (80/20). In this way, my model had no knowledge regarding the current volatility of the market. The way price fluctuated in the past had absolutely no value in predicting current conditions. I'm going to try to solve this issue using walk forward cross validation.

When my model treats a less volatile instrument (SPY), there is less difference on the fluctuations included in the training and test set. This therefore immediately results in more accurate predictions.

Week 8 (19/04/2021 - 23/04/2021)

During my eighth week, I learned more about last week's problems. I have come to the conclusion that there is nothing wrong with the way I validate/split my set (80/20). It turned out to be another data problem, as there are not enough characteristic features in my data. With the help of a few libraries, I was able to attach more characteristic features to my data.

Week 9 (26/04/2021 - 30/04/2021)

During week nine, I again tried to maximize my output several times. I switched to a new method towards the end of the week; predicting session extremes rather than closing prices. That way, I'm already letting time out of the equation.

Week 10 (03/05/2021 - 07/05/2021)

After a series of disappointing backtests, I was again forced to look for another output form. The results again differed greatly, so I spent the rest of week trying to come up with new ideas.

Week 11 (10/05/2021 - 14/05/2021)

After a down week, I finally managed to come up with some ideas. Predicting price at specific times is virtually impossible. This is precisely why I rework my model to a binary classifier. In this way, I hope to achieve better overall quality based on signals.

Week 12 (17/05/2021 - 21/05/2021)

In my second to last week, I successfully completed a series of backtests. The model finally seems to be becoming usable. After a series of simulations, the project appears to have succeeded in a big way after all.

Week 13 (24/05/2021 - 28/05/2021)

The last week was mainly about presenting in front of the company, doing job interviews and creating documentation.