

Weekly Paper Summary (25 points total)

Paper Title	JFDSC - Machine learning for cryptocurrency market prediction and trading
Authors	Patrick Jaquart, Sven Köpke, Christof Weinhardt
Student Name	Julia Troni
Student ID	Student ID: 109280095

1. What do you think the paper is about in layman's terms? What did the research focus on, what did the authors find and what are the main conclusions (if any) [5 points]

The paper aims to highlight the application of different machine learning models in cryptocurrency market prediction and trading. Specifically for generating statistical arbitrage in the cryptocurrency market. Statistical arbitrage is a type of trading strategy that assigns stocks a desirability ranking and then constructs a portfolio to reduce risk as much as possible. This rigorous approach to investing relies on computer models and in depth analysis. (Source: [Statistical Arbitrage: Definition, How It Works, and Example](#))

The researchers used six different machine learning models to predict the relative daily performance of the 100 largest cryptocurrencies by market value. Furthermore, they tested a long-short trading strategy, buying the cryptocurrencies predicted to do well and selling others predicted to do poorly. Then they analyzed the daily closing price and market capitalization data from five different 800-day periods to see how well the models and trading strategy worked.

All models are trained on the binary classification problem of predicting whether a single coin will outperform the cross-sectional median of returns on the subsequent day, based on solely on price information of the previous 90 days. They test and compare recurrent neural networks, convolutional neural networks, tree-based ensemble methods, and the logistic regression (LR) model as a simple and efficiently computed benchmark. They concluded that all employed models make statistically viable predictions with an average accuracy of around 54%. and significantly outperformed a random classifier. Especially, recurrent neural networks (RNN) and tree-based ensembles are particularly effective in classifying the daily performance of cryptocurrencies.

They also found that when they used a long-short portfolio strategy based on the predictions of the models, they were able to achieve better returns than a buy-and-hold strategy. Comparing the long and short leg predictions indicates that short legs are more predictable, as there was a slightly higher accuracy for short leg predictions. A potential explanation for this might be that investors process financial gains and losses differently which may lead to investor behavior being more predictable during a market downturn due to loss aversion. Behavioral biases might be more pronounced for the cryptocurrency market, as cryptocurrencies do not exhibit a fundamental value in

the traditional sense.

Two types of RNN models, the GRU and LSTM models were especially accurate for the long short term trading strategy, since these models predictions yield the highest risk-adjusted performance.

While the study also noted that there were limitations to their research, such as the assumption that they could buy and sell at mid-price and short-sell, and that the results may not be generalizable to other market environments, overall the study suggest high level of predictability in cryptocurrency market.

How would you extend the research paper – what new area(s) would you focus the paper on? [5 points]

Based on the results of this study, there are several pathways for future research. First, I would explore the use of more transparent models. RNN, particularly LSTM GRU and LSTM are quite complex and models such as neural networks particularly LSTM are extremely difficult to understand the decisions and predictions of the models. As explored and discussed extensively in Weekly Paper 1, model interpretability is desired particularly for applications with high stakes. Since cryptocurrency prices can have significant financial effect on investors, models with higher human interpretability would be of benefit. Thus, I would extend the research to developing more explainable models for cryptocurrency market prediction to enhance model transparency.

Second, I would like to focus further research into developing models to predict cryptocurrency prices for coin based vs. token based cryptocurrency. Coin based and token-based cryptocurrencies are created, distributed, and used in very different ways. Thus, the factors affecting the prices of each are different and would thereby affect price prediction models.

2. Discuss at least two real-world applications (not mentioned in the paper) that would benefit from the focus of / applications mentioned in the paper and why [15 points]

The prices of cryptocurrencies depend on a lot of factors like technological progress, internal competition, pressure on the markets to deliver, economic problems, security issues, political factor etc. Their high volatility leads to the great potential of high profit if intelligent investing strategies are taken. Unfortunately, due to their lack of indexes, cryptocurrencies are relatively unpredictable compared to traditional financial predictions like stock market prediction.

We see in this article that recurrent neural networks (RNN) and tree-based ensembles were especially effective in classifying and predicting the daily performance of cryptocurrencies. Particularly GRU and LSTM models performed with high accuracy. Given that Cryptocurrency markets are so challenging to predict, these models can be applied in similarly volatile and unpredictable markets.

For example, Water demand depends on many factors, namely weather, prices, time of day and human behavior. Thus, water demand and consumption is very affected by both the short and long term trends, and Jaquart and Wenheart show that LSTM model performs well in with these conditions. The LSTM model can be useful to predict water demand. This is because LSTM is highly efficient for learning long-term dependencies, especially in sequence prediction problems.

Another way such a model can be applied is for market predictions in the crop and agriculture market. Here, weather, other crops and animals, disease, and supply/demand are large factors that influence agriculture costs. Predicting future price is of upmost importance to allow for farmers, consumers, and governments to adequately respond and prepare to appropriately manage food supply. Recently we are seeing the impact of the avian flu, and a RNN model could be effective in predicting the cost of eggs and other chicken products. The research in this paper provides evidence to support this hypothesis, however future research would need to be extended for RNN in the case of agriculture prices.