

Executive Summary

In 2008, Nakamoto (2008) published his white paper on a peer-to-peer electronic cash system called Bitcoin (BTC). Until then, online payments relied mostly on intermediaries that served as trusted parties. By using the consensus mechanism *proof-of-work*, he solved the so-called *double spending problem*, the possibility of spending digital money twice if there is no intermediary checking all transactions for their validity. Bitcoin uses cryptographic techniques to bundle transactions into blocks, and then links blocks together into one chain called *blockchain*. The blockchain serves as a transaction log and an ultimate authority regarding the ownership of Bitcoins.

Since then, many cryptocurrencies have emerged, some of which turned out to be a scam. The exchange of coins or tokens can be done directly between two parties or facilitated through a centralized exchange. Large exchanges for various cryptocurrencies have emerged in the last few years due to the increasing popularity of cryptocurrencies. While the market cap of all cryptocurrencies was around \$10 billion at the beginning of 2011, it has ballooned to around \$2 trillion by March 2022, and then declined again to around \$860 billion by November 2022. Equally, academic interest in crypto markets has increased tremendously over the past years.

As crypto markets mature, more attention was and will be paid to them from regulators (e.g. European Commission (2020)) and institutional investors. There is a vast amount of high-quality research about cryptocurrencies available and ongoing, mainly from a technical point of view. From a finance perspective, despite the growing interest, there is still a lack of knowledge about the structure and behavior, and the predictability of cryptocurrencies markets.

To add to the sparse literature, this thesis investigates the predictability of Bitcoin prices over various time intervals. Using long short-term memory networks (LSTM), a type of recurrent neural network (RNN), binary forecasts are made about the development of Bitcoin prices. That is, using historical data, predictions are made whether the price of Bitcoin will increase or decrease.

A unique and proprietary dataset provided by the company “Entropy Research” lays the foundation for the analysis. The dataset ranges from 1 July 2021 to 28 March 2022 and contains every trade during that period and periodic limit order book snapshots with a depth of 10 per side. The data is cleaned and aggregated into time intervals of 1 second, 1 minute, and 5 minutes. Various features are

created based on the original and the aggregated data. Then, for all newly created datasets, multiple recurrent neural networks with different parameter specifications are trained and evaluated.

Although some models show a higher test accuracy than the benchmark value, the differences are very small and inconsistent. Furthermore, the models predict mostly, sometimes exclusively, one of the two classes. Additionally, the models are highly sensitive to even small changes in the parameter values, indicating that the data contain no strong signals.

Based on promising results presented in section 1.2, the datasets could be enhanced with network-based data, that is, data regarding information about the Bitcoin blockchain, or with data regarding Bitcoin's investor attractiveness, or other features. Furthermore, a more systematic approach could be used to find a suitable network architecture, for example a grid search. This is, however, somewhat complicated by the long training times of some of the networks, especially when using short time intervals.