ESTIMATING FUTURE FACTS ABOUT DIGITAL ASSETS BASED ON BLOCKCHAIN ACTIVITY DATA

Zvezdin Besarabov
National High School of Mathematics and Natural Sciences, Sofia, Bulgaria, me@zvezd.in
Supervisor: Todor Kolev, Chairman of the Comrade Cooperative, Comrade, todor@comradecoop.com

1. Introduction

Most stock market predictions are based solely on previous price points. However, such prediction strategies perform very poorly in the stochastic and volatile cryptocurrency markets [1]. To combat that we can analyze the public blockchain data, in addition to financial data, in an attempt to predict future blockchain facts.

2. Research Methods

In our paper we explore how modern Deep Learning techniques can be applied to predict future facts about the Ethereum blockchain [2]. Specifically, we are interested if blockchain's public raw data, such as the number of transactions and the account balance distribution, can be used to predict other measures like the number of new accounts created and the market price per ETH token.

We have chosen the Ethereum blockchain for our experiments. Due to its various use cases, its stochastic environment will best show the true accuracy of our predictions. We extract features from the Ethereum blockchain, which include simple values, like the asset price or the total computational power of the network. Other more complex features include spatial 2D distributions of the number of accounts in the network based on 2 account properties:

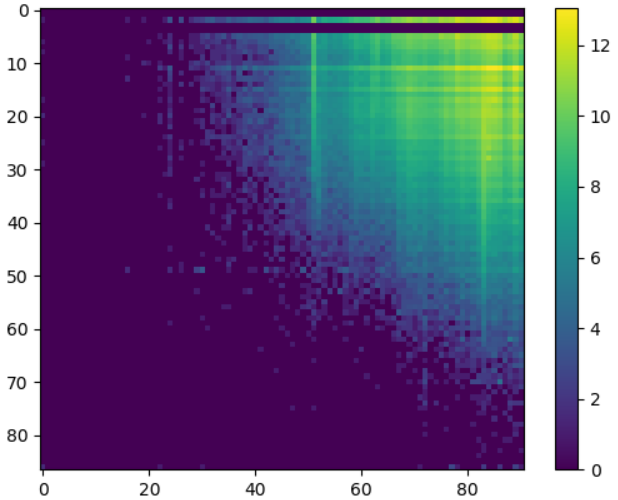

Figure – Distribution of all existing accounts based on their balance (Y) and seconds since last activity (X)

A set of chosen features is shaped to a dataset using spatial stacked layers modeling for neural network training.

We train our long short-term memory (LSTM) and convolutional neural models on making predictions on the basis of the dataset.

Moreover, we have developed a reusable framework providing data gathering, processing, and storing functionality for performing Deep Learning experiments over blockchain data. We are currently working on automating neural network architecture and meta-parameter optimization tasks through training controller Machine Learning models on these tasks. We are developing a genetic engine that can generate a neural architecture for a given task by evolving neural models, based on the principles of natural evolution.

3. Results

During a series of experiments on value predictions, we achieved 330% lower error scores with blockchain data than an LSTM approach with trade volume data. By utilizing blockchain account distribution histograms, stacked dataset modeling, and a convolutional architecture we reduced the error further by 35%. We have also obtained accurate predictions on the amount of new cryptocurrency users.

In our research project, we have evaluated different Deep Learning methods at their ability to find patterns in the Ethereum blockchain data and provide estimations about future measurements of the blockchain assets, including their market price. To facilitate our experiments, we have implemented a data gathering, processing, and storing framework that enabled us to deal with the enormous amount of data on the blockchain (more than 300GB) and efficiently produce datasets for training.

4. Conclusion

To the best of our knowledge, this is the first research in the field that predicts by analyzing blockchain data in depth. We believe that these results can assist users by lowering the risk of using the unstable crypto markets. Our predictions can also be used as a basis for new trading algorithms. Thanks to the flexible framework created for our experiments, we can migrate the same predictive strategies to less noisy blockchains, like Bitcoin, expecting higher accuracy values. Our genetic engine can also be used to find a solution to any machine learning problem, possibly surpassing any human-made state-of-the-art model, as other research has shown [3]. The engine requires a lot of computational power. Since the Ethereum network already utilizes a great amount of GPU processing power, which might become obsolete due to the adoption of a Proof-of-Stake algorithm, we believe that 2018 presents a unique opportunity for achieving an exascale decentralized supercomputer, dedicated to training AI.

5. Literature

[1] Isaac Madan, Shaurya Saluja, Automated Bitcoin Trading via Machine Learning Algorithms. https://goo.gl/SutJsG

[2] Vitalik Buterin, Ethereum Whitepaper. https://goo.gl/9vwEj9

[3] Quoc Le, Barret Zoph, Using Machine Learning to Explore Neural Network Architecture, https://goo.gl/82i9Vo