

Crypto-Currency Price Forecasting

Swapnil Jadhav
23rd Dec 2017

Proposal

Create a Deep Learning based model to forecast prices for one or more Crypto-Currencies in one or more currency exchanges.

Domain Background

Cryptocurrency is an electronic money created with technology controlling its creation and protection, while hiding the identities of its users. Crypto is short for "cryptography", and cryptography is computer technology used for security, hiding information, identities and more. Currency simply means "money currently in use".

Cryptocurrencies are a digital cash designed to be quicker, cheaper and more reliable than our regular government issued money. Instead of trusting a government to create your money and banks to store, send and receive it, users transact directly with each other and store their money themselves. Because people can send money directly without a middleman, transactions are usually very affordable and fast.

To prevent fraud and manipulation, every user of a cryptocurrency can simultaneously record and verify their own transactions and the transactions of everyone else. The digital transaction recordings are known as a "ledger" and this ledger is publicly available to anyone. With this public ledger, transactions become efficient, permanent, secure and transparent. With public records, cryptocurrencies don't require you trust a bank to hold your money. They don't require you trust the person you are doing business with to actually pay you. Instead, you can actually see the money being sent, received, verified, and recorded by thousands of people. This system requires no trust. This unique positive quality is known as "trustless".

The first cryptocurrency was bitcoin.

Other terminologies and easy explanation could be found here -
<https://medium.com/decryptionary/introduction-to-cryptocurrency-2f182c577d7c>

Problem Statement

Mining with cryptocurrency has become difficult and return on investment is very less. Depending on usability/utility of cryptocurrency it can also be traded like any other commodity. Recently there has been surge of currency exchanges which allows to buy and sell of cryptocurrencies (or simply coins). Though looks very similar to stock market, there are very few studies done on price variations and forecasting of the same.

Some of the famous cryptocurrency exchanges are -

- <http://coinbase.com/>
- <http://cex.io/>

In this capstone project following problem statement will be targeted to solve using Deep Learning methods.

Forecast the average price of a particular crypto-currency for the next N days given history of price variations for the same.

Few other mini problem statements can also be solved -

- Forecast price variation/volatility for each day
- Forecast price by hour instead of day-wise

Datasets and Input

There is no standard dataset available to solve this problem. Instead dataset can be prepared by either using price-history APIs or web crawling the website.

Major data source identified is <https://coinmarketcap.com/> which has APIs and also historical data can be easily scrapped.

Initial experiment will be done on only few crypto-currencies like Bitcoin (BTC), Ether (ETH), Litecoin (LTC), Ripple (XRP) and Bitcoin-Cash (BCH) in that particular order. Also, we can take data for more than one cryptocurrency exchange and predict for all. But for initial experiments only one of the cryptocurrency exchange will be chosen and methodology will be tested.

Output will be forecast for Coin-X for next 7 days.

Major input data-points for any coin will be -

- Historical price for given Coin-X for exchange 1 to M
- Historical price for rival coins for exchange 1 to M

Other input points which could be very useful are -

- Major events/dates for the Coin-X or other rival coins
- Social media or Market sentiments about the coins each day

Solution Statement

Creating the features for unknown behaviour in crypto-currency is very difficult and using Neural Network would be ideal to capture the same through hidden layers.

As it is time series model, following basic models will also be generated for baseline comparison.

- Regression Models - XGBoost regressor, LinerSVR, LineraRegression
- Time Series Based Models - AR, MA, ARIMA

Following deep-learning techniques can be used

- Shallow neural network
- Multi-layer neural network
- Recurrent neural networks (RNN)
- Long Short Term Memory networks (LSTM)
- Gated recurrent Unit (GRU)

Some other techniques which might be helpful individually or with any one of the above methods in combinations are -

- Bidirectional RNN
- Attention based RNN
- Convolution Neural Network (CNN)

Benchmark Model

As it is a completely new domain of forecasting, benchmark results are not available. But we can compare with Stock Market standards for algorithm and human experts.

Few of the studies of stock market forecasting and benchmarking can be found here and can be compared with results of

- On **cxoadvisory** website human expert accuracy percentages are given and max accuracy is 68.2% and min is 20.8% (<https://www.cxoadvisory.com/gurus/>)
- According recent study by **financial-math** 78.7% was the highest human-expert accuracy found with 16% having >60% accuracy and minimum was 17% with 27% having <40% accuracy. (<http://www.financial-math.org/blog/2017/03/how-accurate-are-market-forecasters/>)
- Comprehensive algorithms performance can be found in this literature paper review published in 2017. (<http://ccsenet.org/journal/index.php/ijef/article/view/70421/38874>)

Evaluation Metrics

Performance of the algorithm can be checked with various matrices

- Mean Squared Error (MSE)
- Mean Absolute Error (MAE)
- If we convert problem to find only increase and decrease in next day as Classification ... then accuracy (Precision/Recall)
- Assuming that very small transaction will not affect market, actual profit/loss that could have occurred in validation period can be measured.

Project Design

Project will have 5 major modules.

- Data generation
- Data preparation
- Models generation and evaluation (kind of simulation)
- Result analysis and comparisons
- Simple Web API to get forecast for next 7 days and currency and exchange market can be chosen.

Preferred programming language: Python

Preferred python libraries: Numpy, Scipy, Pandas, Scikit-Learn

Preferred Neural network libraries: Keras, Tensorflow

Server if possible will be hosted with public domain and on amazon-aws.

Otherwise local flask server to check the results locally.