Blockchain and Cryptocurrencies Final Project
Research of correlation between "big accounts" transactions and BTC price

Submitted by: Bradley Feitsvaig 311183073.

Introduction

It seems that whenever a "big BTC user" buys bitcoin then the BTC price raises. This project, Analysis the correlation between BTC "big accounts" transactions and BTC price. The project implemented in NodeJS and uses SVM from "ml-svm" which is a machine learning package.

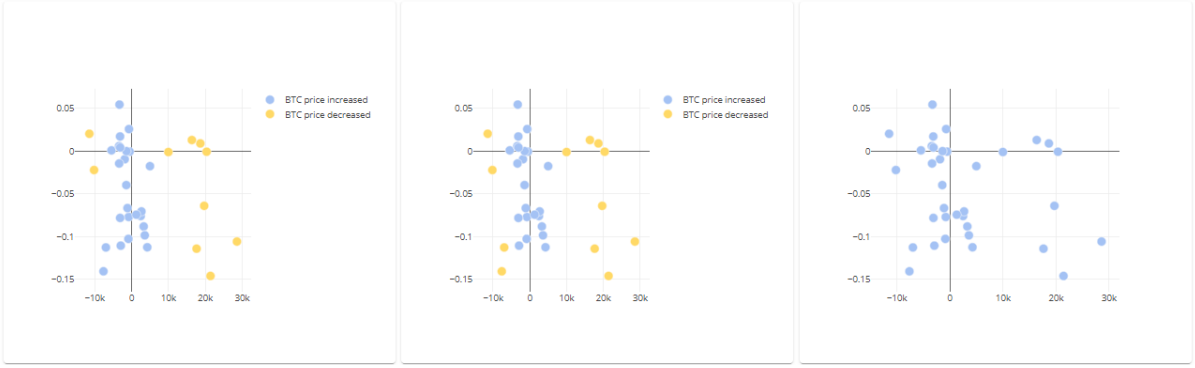
Project flow

The project contains 3 SVM models. The first model is trained and predict Bitcoin price increases or declines the day after the transaction. The second one is for the third day after the transaction. The last is for a week after the transaction.
Both datasets consist of 6 columns [Date, BTC, LastWeek, NextDay, NextThreeDays, NextWeek].
Date- date of the transaction.
BTC- amount of Bitcoin in the transaction (-: if sell, +: if buy).
LastWeek- Percentage change in the past week.
NextDay– value is 1 If there was an increase in BTC price and is -1 If there was a decrease in BTC price in the next day after the transaction.
NextThreeDays, NextWeek – the same as for "NextDay" but for the third and the seventh day accordingly.
The training data set is for Bitcoin Address 1P5ZEDWTKTFGxQjZphgWPQUpe554WKDfHQ, with Balance: 125,956.57335776 BTC, 4,683,289,625.09 USD.
The testing data set is for Bitcoin Address bc1qm34lsc65zpw79lxes69zkqmk6ee3ewf0j77s3h, with Balance: 40,470.16856242 BTC, 1,504,993,438.1 USD.
The output of the program is 3 scatter plots, one for each model. And output to console the accuracy of the models.
The scatter plot should display points for each transaction from the testing dataset. X-axis is for BTC purchase amount, and Y-axis is the Changes (in fraction) of BTC in last week before the transaction. If the point on the scatter plot is blue, it means that the relevant SVM model predict that after the transaction the BTC price should increase (1,3 or 7, days after the transaction- according to the model). Otherwise, if the prediction is decreasing, the point will be yellow.

Obtained results

"One day" model accuracy is: 54.285714285714285
"Three days" model accuracy is: 45.714285714285715
"Week" model accuracy is: 34.285714285714285

"One day" model | "Three days" model | "Week" model



As it stands from the result, the model accuracy is not high enough and it's affecting the prediction of the model. The medium accuracy of the models could be derived from a low quantity of transactions in the training dataset.

Conclusions

From observation at the train and test datasets it seems that there is a correlation between "big accounts" transactions and BTC price. The accuracy of the models in this project was medium as mentioned in obtained result part, Accordingly, for better result, there is a need for larger training dataset and maybe use of another machine learning model or modification of the of SVM's parameters. My conclusion is that this method could be a good way to predict changes in BTC prices, but it must be strengthened by some other method to obtain better result such as tweet analysis or use of RNN with BTC prices history.

How to run the program

REQUIRED PACAGES

Nodeplotlib for plots : npm i nodeplotlib
ml-svm for SVM: npm i ml-svm
fs for file reading: npm i fs

Run

node index.js

Run will take a few seconds, in the end it will outputs the models accuracy in the console and open a browser with the scatter plots with the predicted values for testing data set with the 3 SVM models.