BitCoin, Machine Learning and Using Them together.

Bitcoins-Magic Internet Money

Bitcoin is a purely

- digital
- Open Source
- P2P
- secure
- Currency with no central regulating authority.

Short Overview of how it works

1. Each person in the bitcoin network has an account with a public id and a private password like an email id.
2. When one id wants to send money (bitcoins) to another id it broadcasts its intent to all ids in the network .
3. This way all bitcoin nodes(id’s) know the position of all bitcoins ever created so there can be no fraudulent transactions.(absence of central authority)

A Deeper Dive

Why Cryptocurrency?

Because the integrity of the entire bitcoin network depends on a few basic cryptographic principles.

As will be discussed in the following slides.

How the intent is broadcasted

1. The intent of transferring bitcoins is broadcasted into the bitcoin network by the sender of the coins.
2. The sender’s computer contains the bitcoin software that creates an encrypted message using the sender’s secret password .
3. This message can be verified to be associated with the sender using his/her public verification key (generated by the bitcoin software).

Confirmation of transaction

1. There are special bitcoin nodes called miners (any node can become a miner if it so wishes).
2. These miner nodes accumulate all broadcasted but unhashed(more on this later) transactions and add a reward transaction where they pay themselves for their effort.
3. The miners then add a **random string** and the **previous hash** at the end of the message so created and hash the block of transactions with the SHA-256 algorithm.

Hashing

Process of converting a message of arbitrary length into a fixed length string of seemingly random alpha-numeric characters.

- Prime hashing example.
- 31662.

Calculation of nonce

Nonce is the random string that miners attach to the end of the transaction message to create a desirable hash.

- Link the prime number to the nonce.

Remember introduction?

Note that the only way the bitcoins are introduced in the network is by the miners rewarding themselves bitcoins for their effort of hashing the transactions.

- We want to make sure that the introduction of bitcoins is not random and occurs in a controlled manner.Why?
- The system has a tool for this called **difficulty.**

Difficulty

- The hashes created by the miners can not be random.
- The system dictates that the nodes only create hashes that are less that a challenge.
- The challenge is auto generated by the system looking at how quickly the previous transactions were hashed.
- The system maintains hashing time at 10 min per block .

Remember Nonce?

- To meet the challenge of the system the miners add an appropriate nonce so that the hash created by them is less than the challenge.
- Finding this nonce is not simple it involves trial and error of around 300,000,000,000,000(300 trillion) nonces.
- This is because no one knows how to back predict the SHA-256 hashing algorithm.
- If you do you become a billionaire just by mining faster than any man alive.

The Block chain-20GB of transactions

- The Ledger that has all transaction on it since the beginning of bitcoin time.
- Once hashed and accepted by nodes in the network, the block of transactions along with its nonce and hash is added to the Blockchain and the new Blockchain is synced up with all nodes in the network.

The security of the network

- Any node in the network can potentially change the block chain. So how is this ever secure.
- Well there is this assumption that Good always Beats Evil.
- Remember the nodes when presented with 2 blockchains always choose the longer and more complex block-chain.

Good beats Evil

- The evil node that wants to change a past transaction in the block-chain would have to do the following .
- First he would have to modify the transaction then create a hash based on the difficulty of that time.
- Then he would have to recompute the hashes of all subsequent blocks as they include the hash of the previous block.
- The hope is the good nodes have enough computing power to always stay ahead in the block-chain and never give evil a chance.

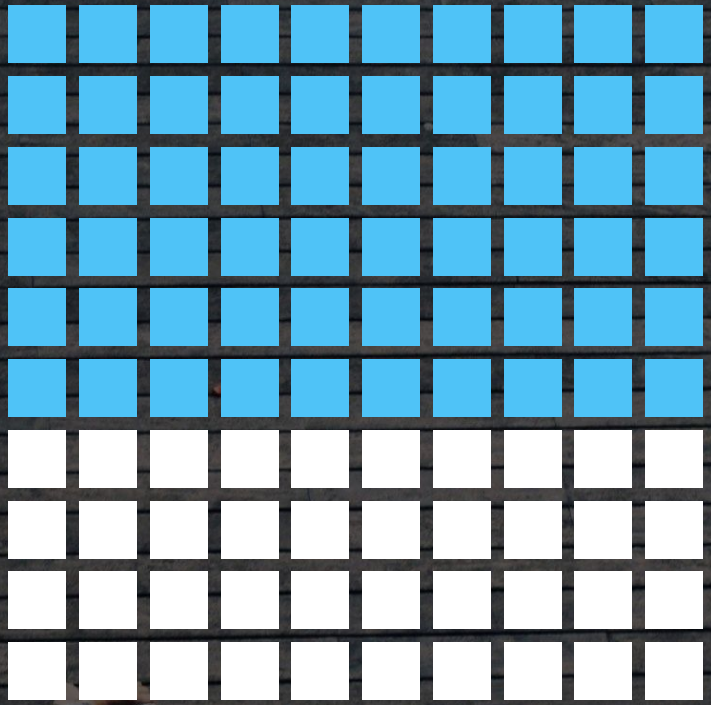
Empty Block problem

- The problem is that miners can keep creating blocks with no transactions except their rewards this is detrimental to the system.
- This is countered by producing only a finite number of bitcoins ever and the reward for mining is also reduced to half after every 210,000 bitcoins mined starting from 50 bitcoins for the first block(Genesis block).
- Finally only 21million bitcoins can be produced.
- After which miners live on transaction fees.

Predicting BitCoin Exchange Rates

Why BitCoins?

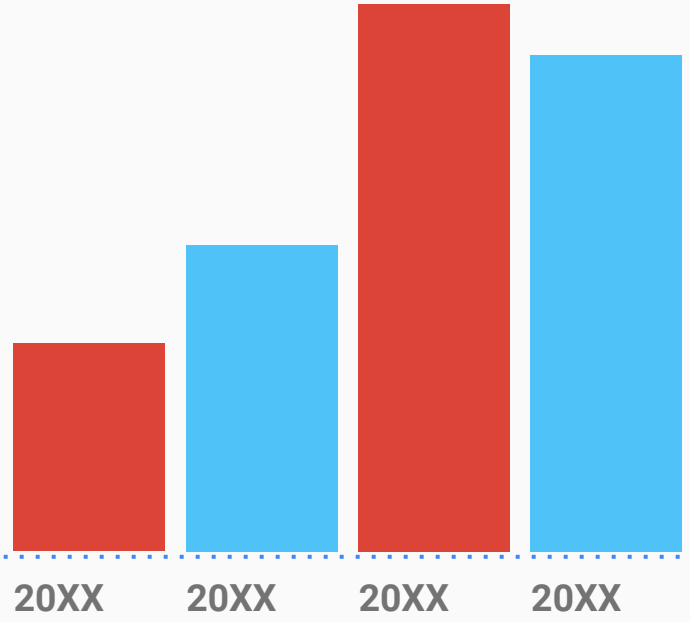
Bitcoins have come into prominence because of their novelty and openness . This has led to highly **fluctuating exchange rates , and high return on investment**. This motivated us to choose bitcoins for our project.

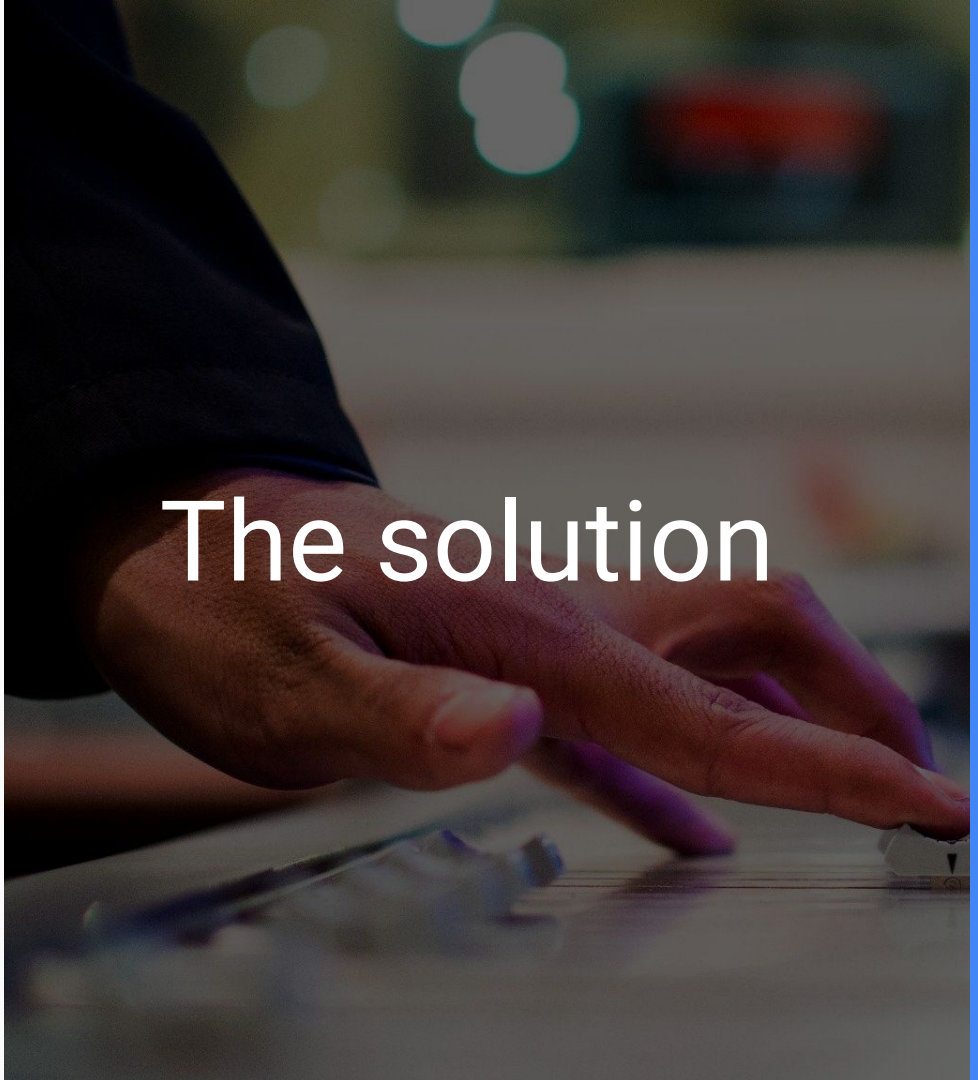


The Problem

Due to the absence of any regulatory mechanisms Bit Coin exchange rates are extremely volatile.

We set out to solve the problem of finding the best model and features to fit the Bit Coin exchange data.





The solution

We went through a variety of training models that we trained on data and looked for min. Test error(Root mean square)

The models we tested

SVM

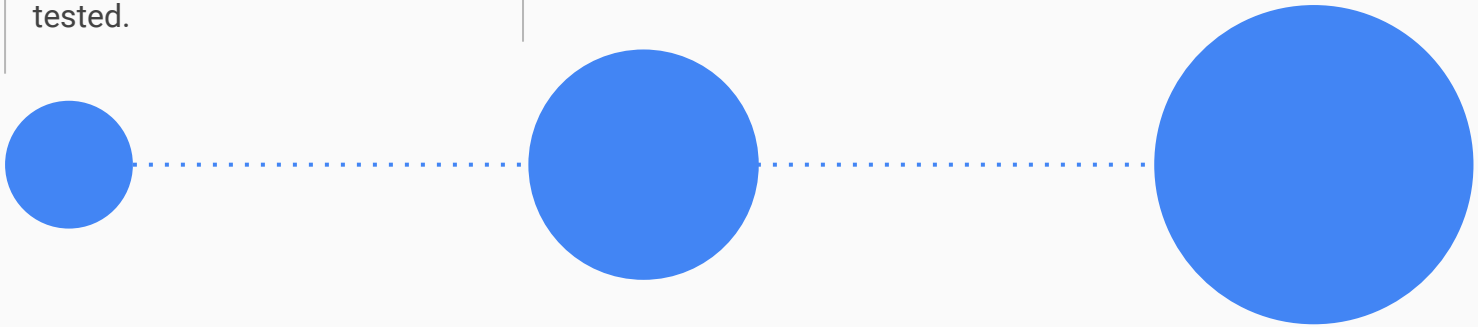
Lowest error for all time marching models tested.

NN

The NN regression from R was used it gave a less accurate result.

Random Forest

The random forest gave the highest rms error.



Features Used

The features we used to train our models include:

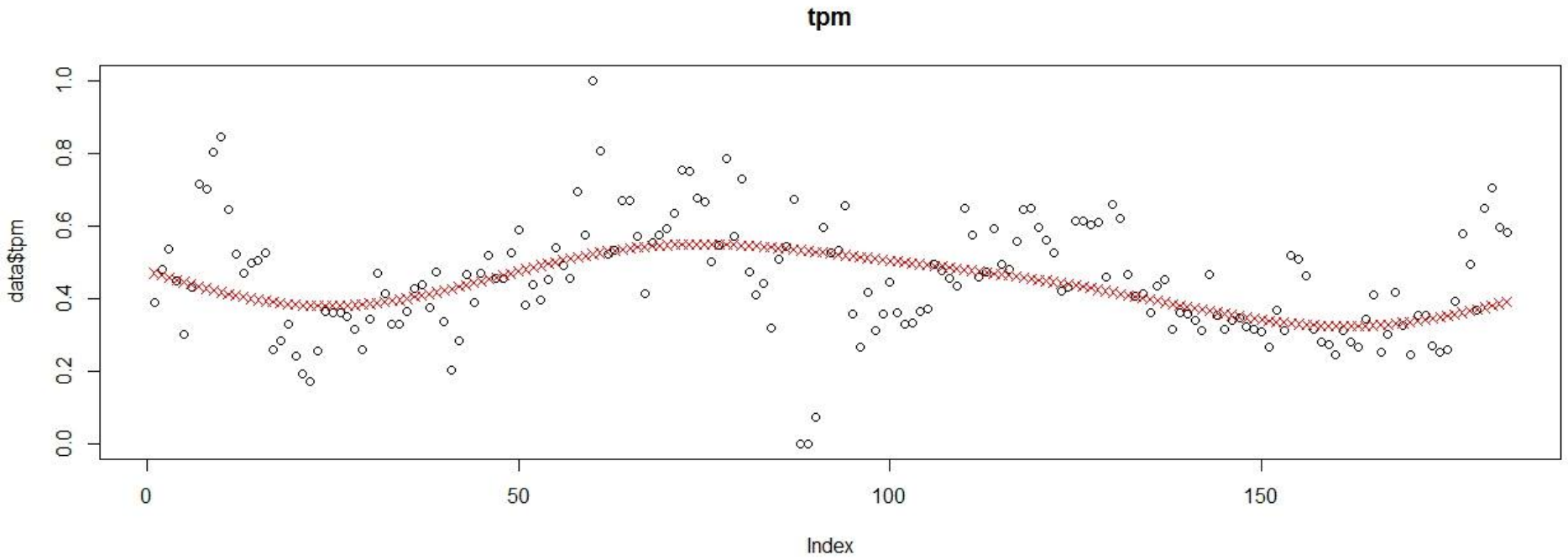
- The mining difficulty of the bitcoins.
- The number of trades per minute.
- The volume of transactions occurring in past days.
- The number of transactions in past days.

Method Used to train the model

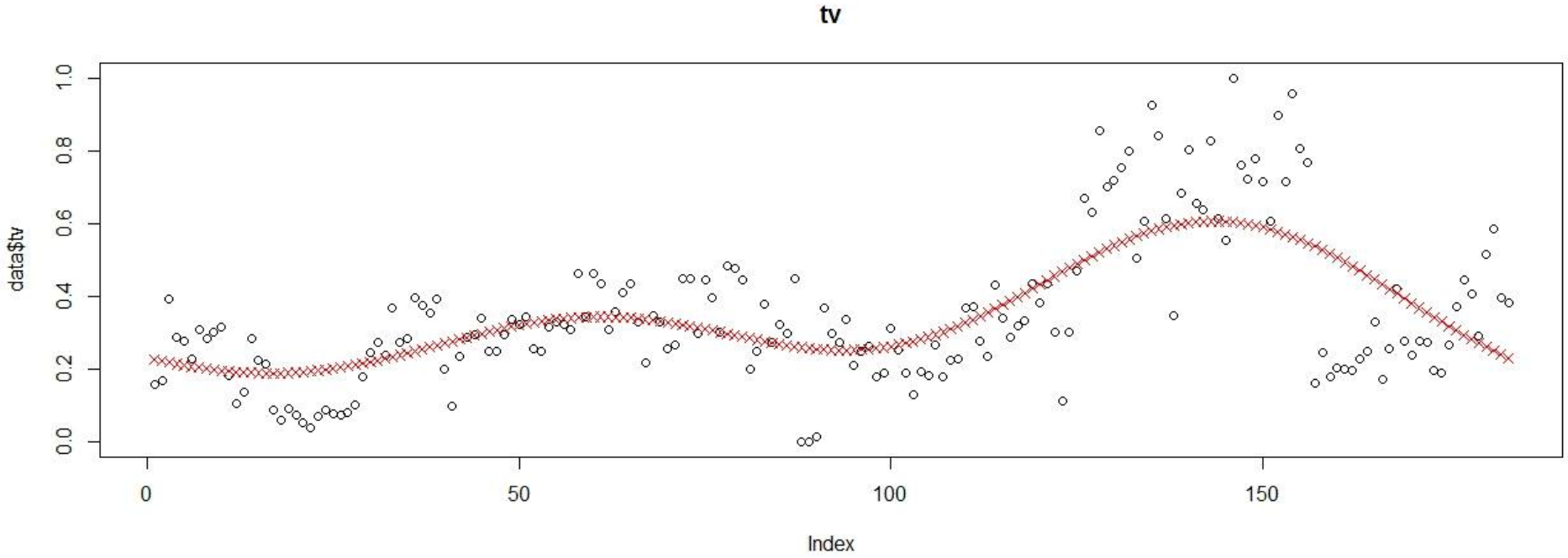
The features that we used to train all three models

1. First all the required features were obtained from the data sources mentioned at the end of the presentation.
2. The relationship between price and the features considered is found by using the using the existing data as a training set.
3. The parameters were tuned using the cross validation set.
4. Now we predicted the features at a particular day by using a model that we fit on the feature vs. time graph.

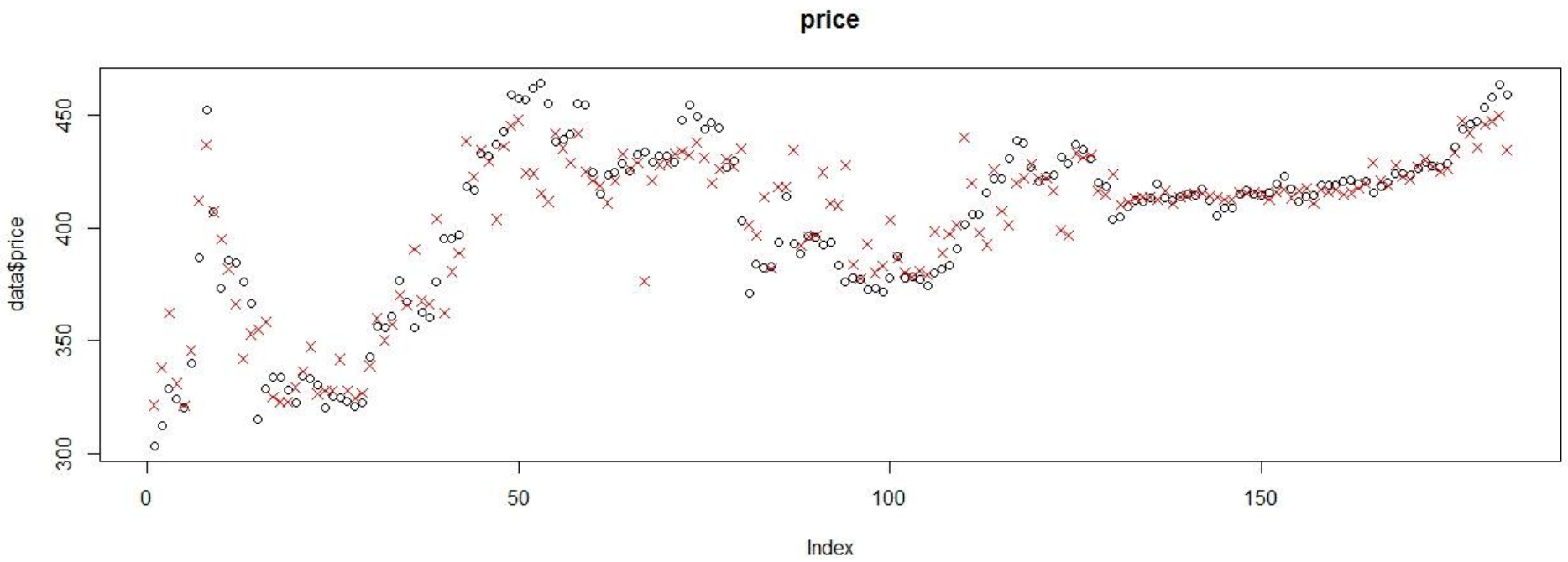
Transactions per minute model



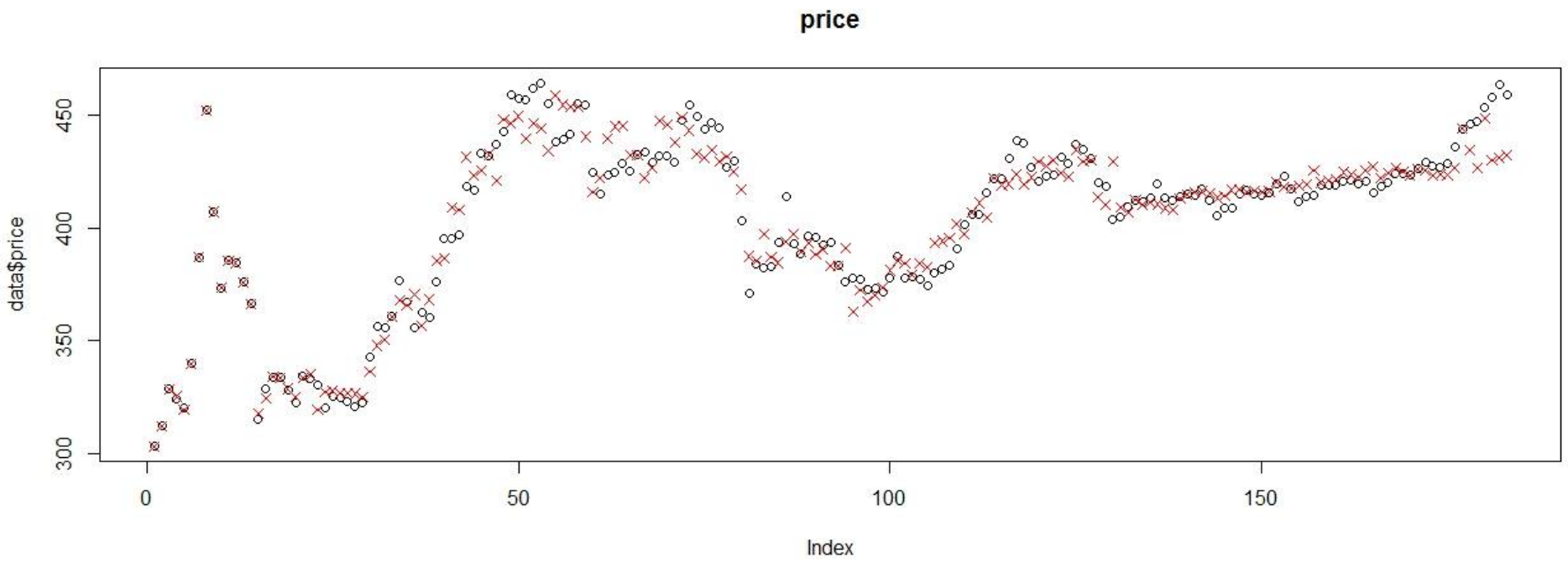
Trade volume with model



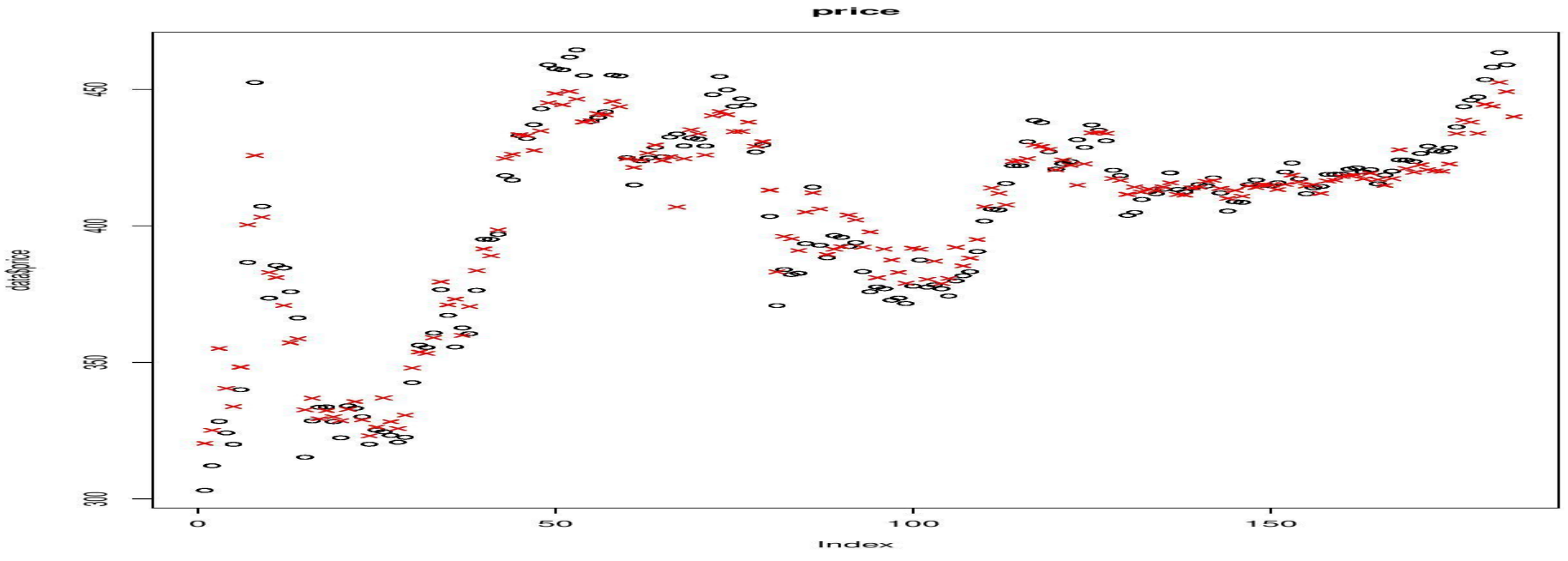
Price model using SVM



Price model using NN



Price Model using Random Forest method



Appendix

Data Sources:

- ***Bitcoinity.org***
- ***Wikipedia for data on time series analysis***
- ***R tutorials for help regarding application of R for our project.***

The team

Srinivas

130010033

Abhilash

130010006

Siddharth

130010038

Abhijeet

130010030


Thank You
For the opportunity given to us.