

Data Science Final Project

比特幣的價格走勢與其他幣種價格之間的關係

DS Final Project概述

研究目的：

想瞭解比特幣的價格走勢與其他幣種價格之間的關係

研究方法：

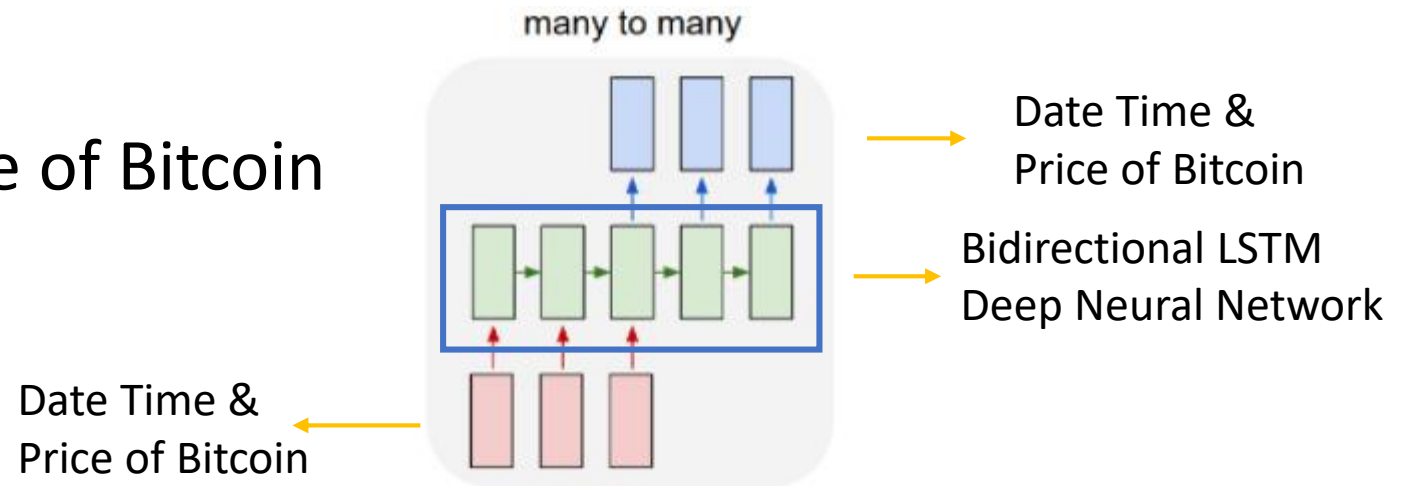
先用深度學習模型預測比特幣的價格，對資料集調整後再用相關性分析探討幣種價格之間的關係，再針對比特幣近年發生重大的事件進行更深入的討論。

研究步驟：

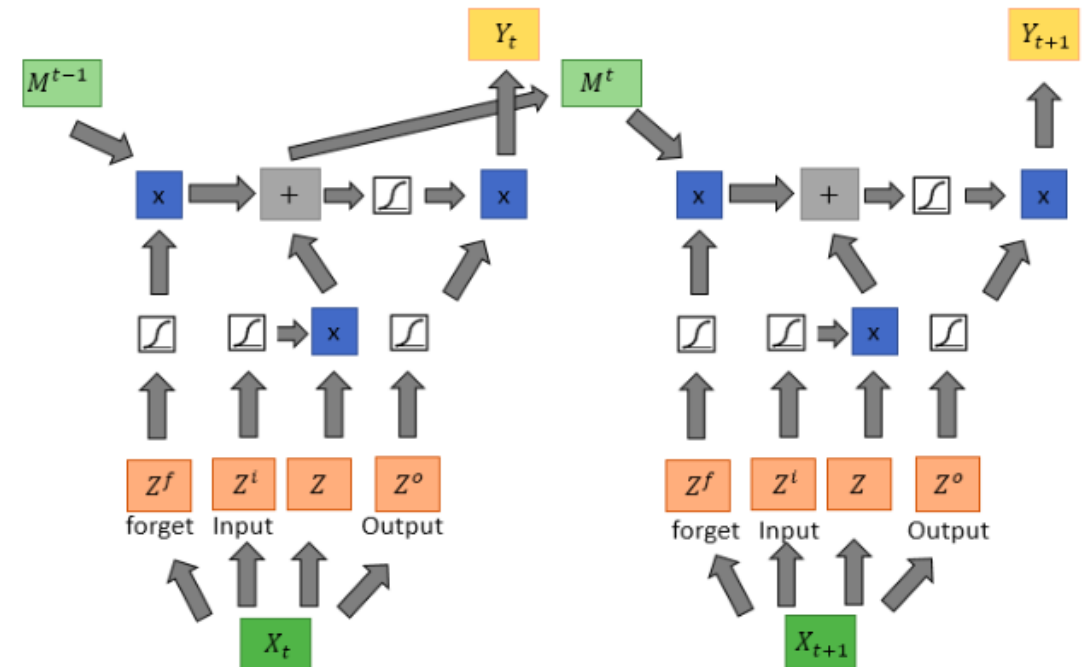
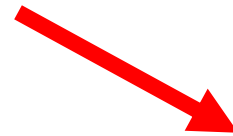
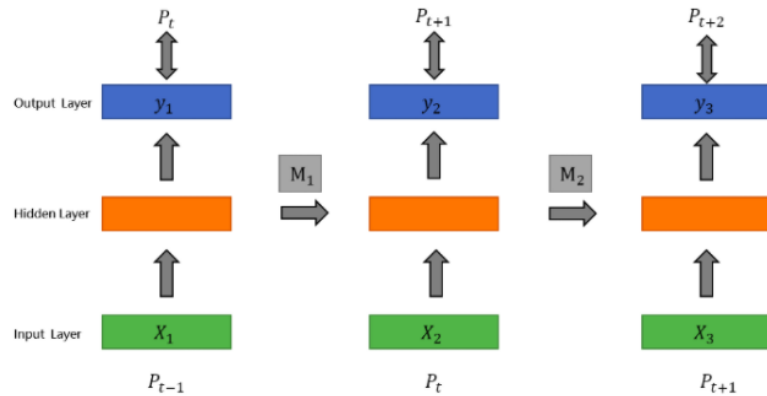
1. 透過幣安API蒐集資料
2. LSTM模型預測比特幣價格
3. 資料相關性分析
4. 與時事結合選定三個比特幣價格波動的事件
5. 相關性分析針對特定事件深入探討

Deep Learning – Predict BTC Price

- Goal : Predict the price of Bitcoin as precise as possible
- Tools : Google Colab GPU
- Deep Learning FrameWork : TensorFlow
- Deeplearning Model : Bidirectional LSTM Deep Neural Network
- Dataset : Binance API
- Input : Date Time & Price of Bitcoin



Deep Learning – Model LSTM Introduction



1. A Type of Recurrent Neural Network
2. Deals with Time Series Model
3. Composed of 4 Gates :
 - Input Gate
 - Output Gate
 - Memory Cell
 - Forget Gate

Deep Learning Model – Data Preprocessing

- `MinMaxScaler()` : Normalize the data between (0,1)
- `df. df.BTCUSDT.values.reshape(-1, 1)` : reshape the BTC value between (-1,1) to better normalization

```
scaler = MinMaxScaler()

BTCUSDT_price = df.BTCUSDT.values.reshape(-1, 1)

scaled_BTCUSDT_price = scaler.fit_transform(BTCUSDT_price)

scaled_BTCUSDT_price.shape

scaled_BTCUSDT_price = scaled_BTCUSDT_price[~np.isnan(scaled_BTCUSDT_price)]

scaled_BTCUSDT_price = scaled_BTCUSDT_price.reshape(-1, 1)

np.isnan(scaled_BTCUSDT_price).any()
# print(scaled_BTCUSDT_price)
```

Deep Learning Model – Build LSTM Model

- 3 layer LSTM Recurrent Neural Network
- Bidirectional : to train on the sequence data in forward and backward (reversed) direction
- Dropout Rate : 20 %

```
DROPOUT = 0.2
WINDOW_SIZE = SEQ_LEN - 1

model = keras.Sequential()

model.add(Bidirectional(CuDNNLSTM(WINDOW_SIZE, return_sequences=True),
                           input_shape=(WINDOW_SIZE, X_train.shape[-1])))
model.add(Dropout(rate=DROPOUT))

model.add(Bidirectional(CuDNNLSTM((WINDOW_SIZE * 2), return_sequences=True)))
model.add(Dropout(rate=DROPOUT))

model.add(Bidirectional(CuDNNLSTM(WINDOW_SIZE, return_sequences=False)))

model.add(Dense(units=1))

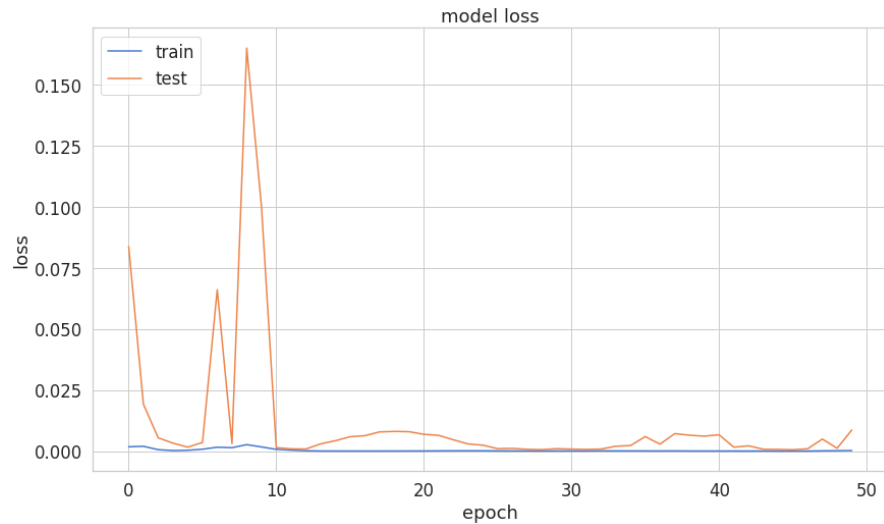
model.add(Activation('linear'))
```

Deep Learning Model - Training

- loss function : Mean Squared Error
- optimizer : Adam
- epochs : 50

```
model.compile(  
    loss='mean_squared_error',  
    optimizer='adam'  
)  
BATCH_SIZE = 64  
  
history = model.fit(  
    X_train,  
    y_train,  
    epochs=50,  
    batch_size=BATCH_SIZE,  
    shuffle=False,  
    validation_split=0.1  
)
```

Deep Learning Model – Visualization&Evaluation



```
model.evaluate(X_test, y_test)
```

```
plt.plot(history.history['loss'])  
plt.plot(history.history['val_loss'])  
plt.title('model loss')  
plt.ylabel('loss')  
plt.xlabel('epoch')  
plt.legend(['train', 'test'], loc='upper left')  
plt.show()
```

```
y_hat = model.predict(X_test)
```

```
y_test_inverse = scaler.inverse_transform(y_test)  
y_hat_inverse = scaler.inverse_transform(y_hat)
```

```
plt.plot(y_test_inverse, label="Actual Price", color='green')  
plt.plot(y_hat_inverse, label="Predicted Price", color='red')
```

```
plt.title('Bitcoin price prediction')  
plt.xlabel('Time : hours')  
plt.ylabel('Price')  
plt.legend(loc='best')
```

```
plt.show();
```



敏感度分析

以比特幣為標準，對其他56種幣做相關性的分析

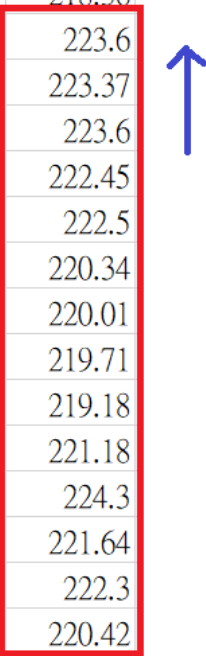
做法:

對其他56種幣的資料做一次次的shift，並拿這些資料與比特幣原本的資料做相關性分析，探討這些幣對於比特幣的關係會不會隨著時間變化，找出這些幣對於比特幣的敏感度。

執行方式如圖：

BTCUSD	ETHUSD
8547.25	218.36
8639.28	223.6
8630.86	223.37
8620.36	223.6
8580.73	222.45
8586.83	222.5
8543.48	220.34
8541.16	220.01
8532	219.71
8539.63	219.18
8570.43	221.18
8648.37	224.3
8575	221.64
8609.34	222.3
8565	220.42

BTCUSD	ETHUSD
8547.25	218.36
8639.28	223.6
8630.86	223.37
8620.36	223.6
8580.73	222.45
8586.83	222.5
8543.48	220.34
8541.16	220.01
8532	219.71
8539.63	219.18
8570.43	221.18
8648.37	224.3
8575	221.64
8609.34	222.3
8565	220.42



也就是說，如果我們今天拿A幣與比特幣做相關性分析，如果他們兩個的關係在shift較少次數後，達到最高點，就表示A幣對比特幣的敏感度高，反之則表示對比特幣的敏感度高。

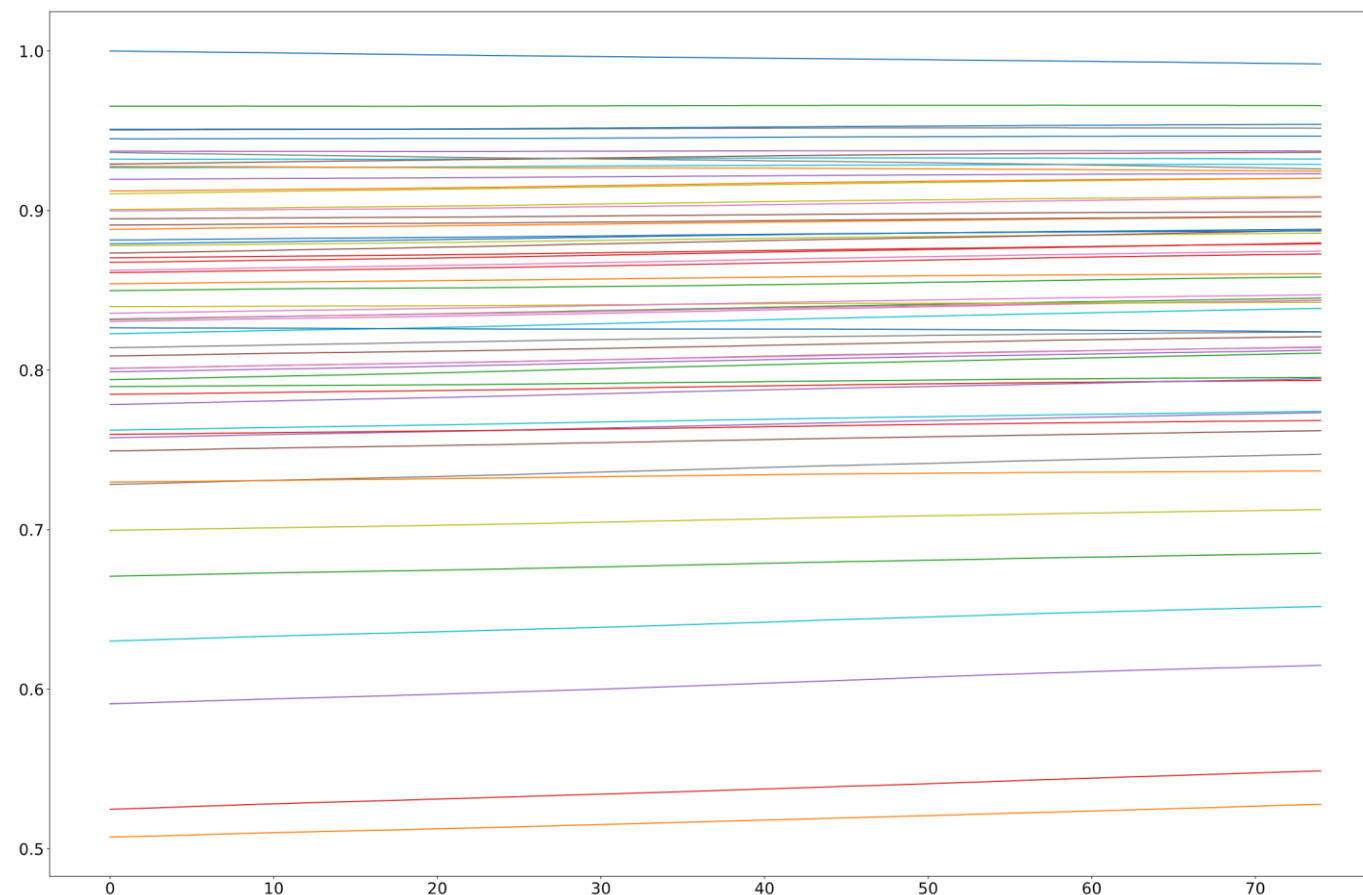
結果:

可以看到每條線幾乎都是水平發展，所以我們得出一個結論：

這些幣與比特幣的關係有沒有shift都沒差，也就是說，如果A幣與比特幣有高相關性，那他的高相關性會一直維持著，反之也是

x軸為shift的次數，
y軸為相關係數

我們x軸取到75小時，
也就是大約3天之後，
如果超過3天我們視為無關



2.不同時間區間的相關性

對比特幣與其他56種幣
進行各種相同的時間區
間的相關性分析。

做法:

指定時間區間，對此區間內的比特幣與其他幣進行相關性分析，並且分析不同的時間區間內，相關性有何變化。

執行方式如圖：

BTCUSD	ETHUSD
8547.25	218.36
8639.28	223.6
8630.86	223.37
8620.36	223.6
8580.73	222.45
8586.83	222.5
8543.48	220.34
8541.16	220.01
8532	219.71
8539.63	219.18
8570.43	221.18
8648.37	224.3
8575	221.64
8609.34	222.3
8565	220.42

BTCUSD	ETHUSD
8547.25	218.36
8639.28	223.6
8630.86	223.37
8620.36	223.6
8580.73	222.45
8586.83	222.5
8543.48	220.34
8541.16	220.01
8532	219.71
8539.63	219.18
8570.43	221.18
8648.37	224.3
8575	221.64
8609.34	222.3
8565	220.42



圖中以**10**格為例，而我們實際是以一個月為一個區間單位，而每一筆資料是相隔一個小時，所以我們是以**720**格為一個區間單位進行分析。

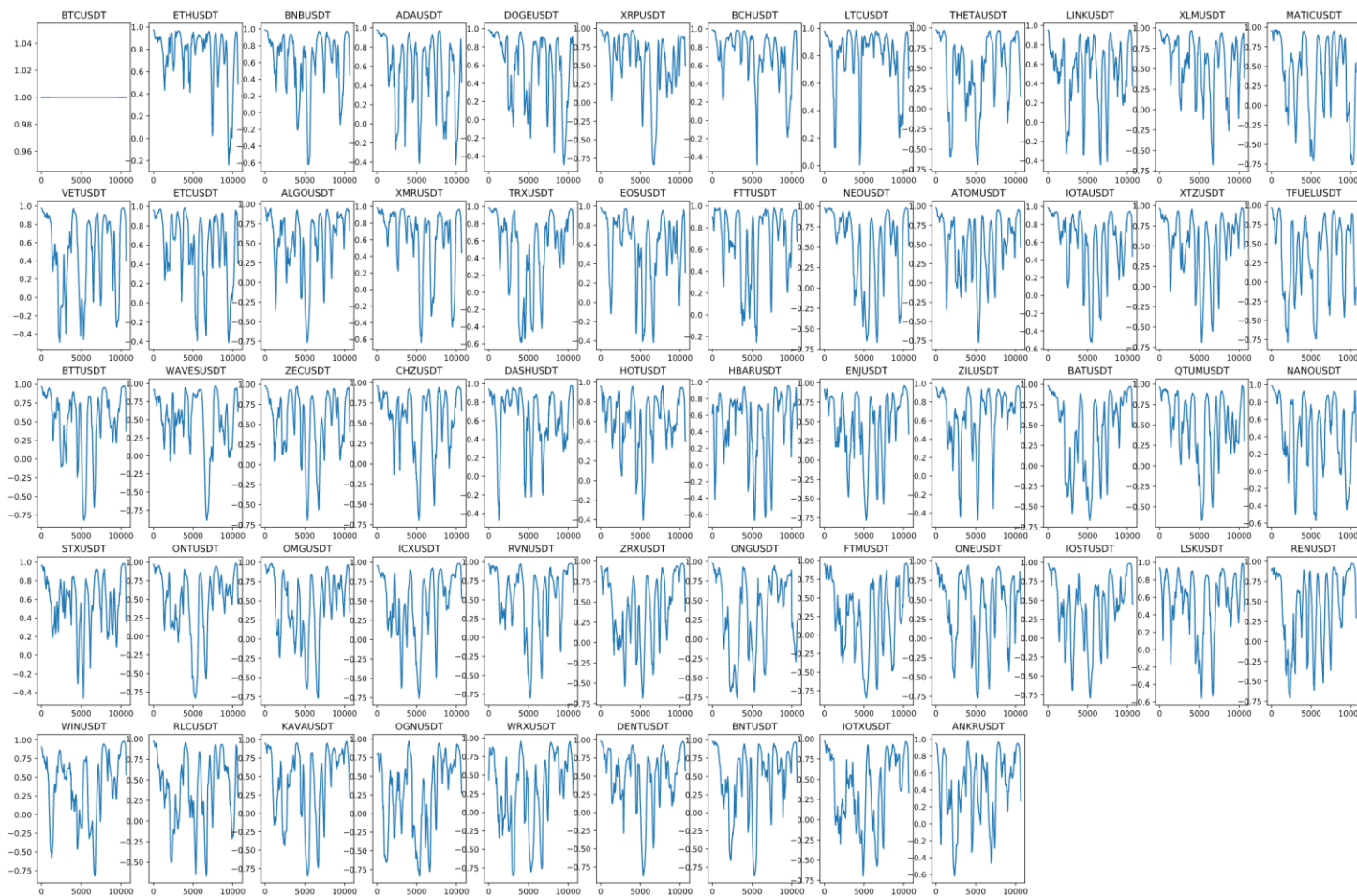
結果:

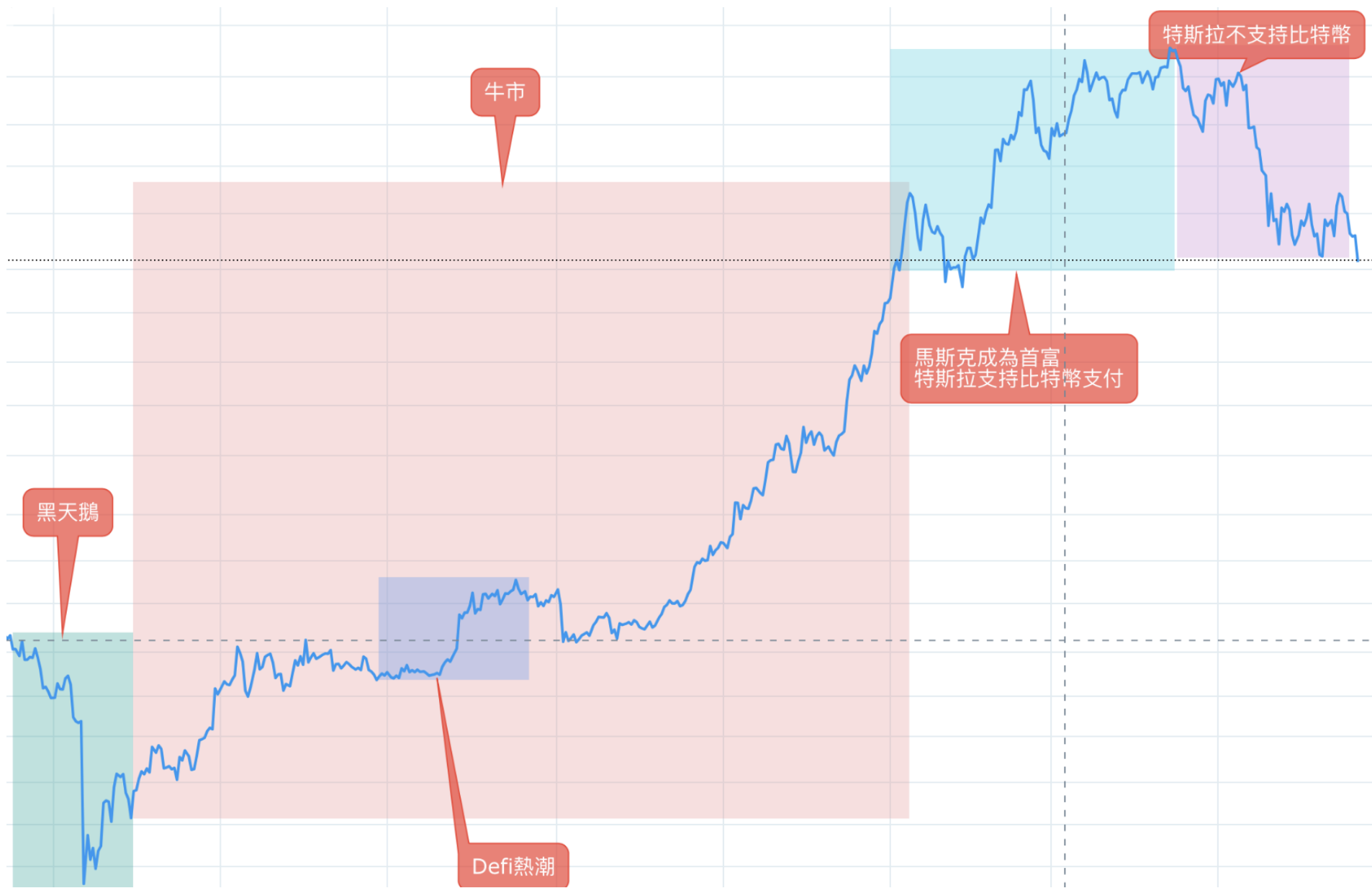
從圖中可以看出，每一種幣與比特幣的相關係數在不同時間區間有很大的變化，也就是說，如果A幣與比特幣上一個月呈現高度正相關，但是下一個月可能就變成高度負相關，這些幣與比特幣之間的關係變動很劇烈。

X軸為時間區間的頭為第幾格

(如果X=5，表示[5,5+719]是我們分析的時間區間)

Y軸為相關係數



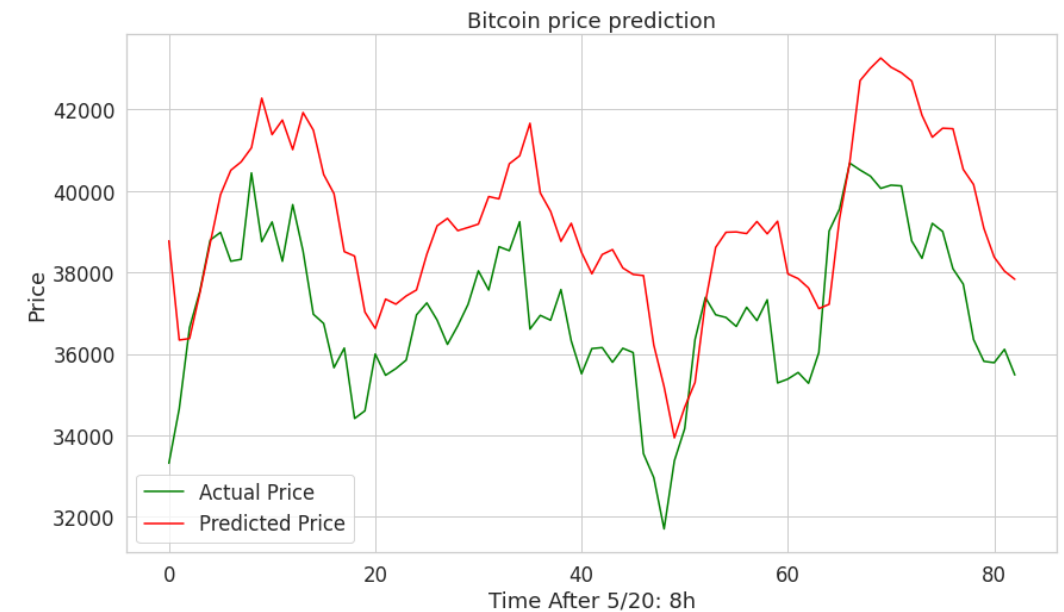


BitCoin Price Overall

Predict : 2021.5.20~2021.6.20 Bitcoin Price

Train : 2017.8.17~2021.5.19 KLINE_INTERVAL_8HOURS

Val : 2021.5.20~2021.6.20

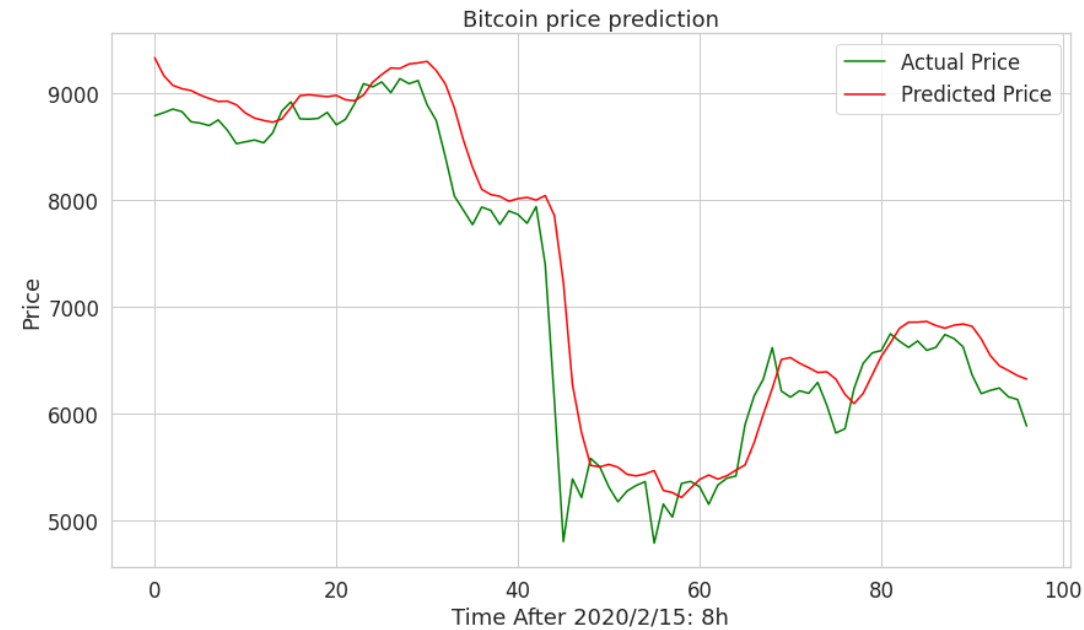


2020黑天鵝時期

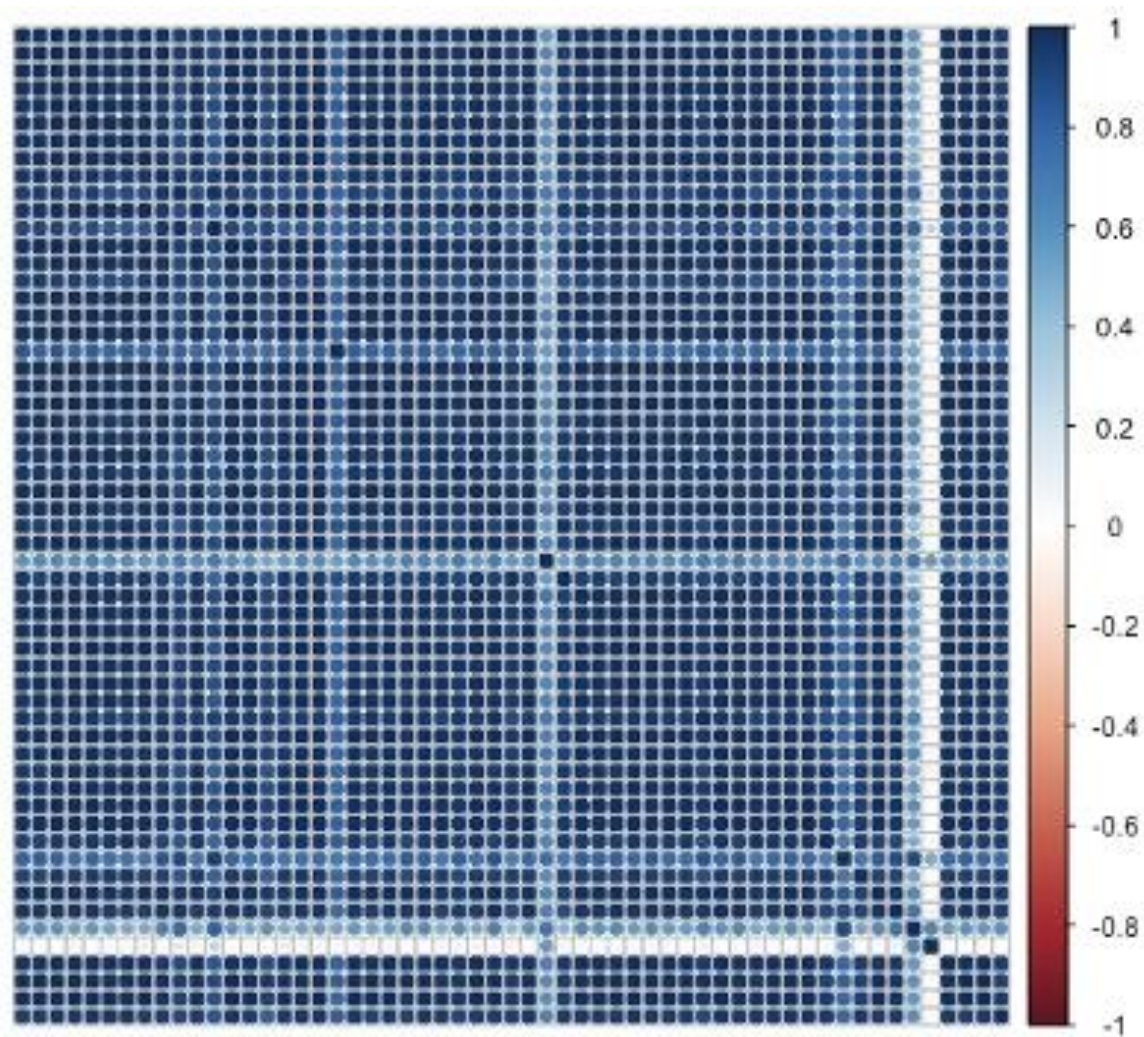
Predict : 2020.2.15~2020.3.30 Bitcoin Price

Train : 2017.8.17~2020.2.14 KLINE_INTERVAL_8HOURS

Val : 2020.2.15~2020.3.30



2020黑天鵝時期



HBAR, OGN

跟BTC的correlation都是0.588

相對其他幣已經算小了

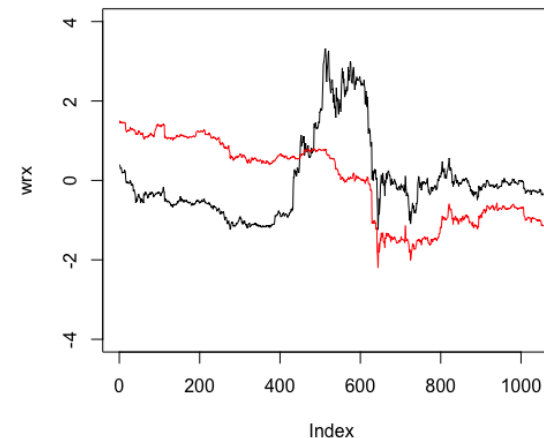
而WRX的correlation

[1] -0.0294147

2.5%

97.5%

-0.067272341 0.008545093

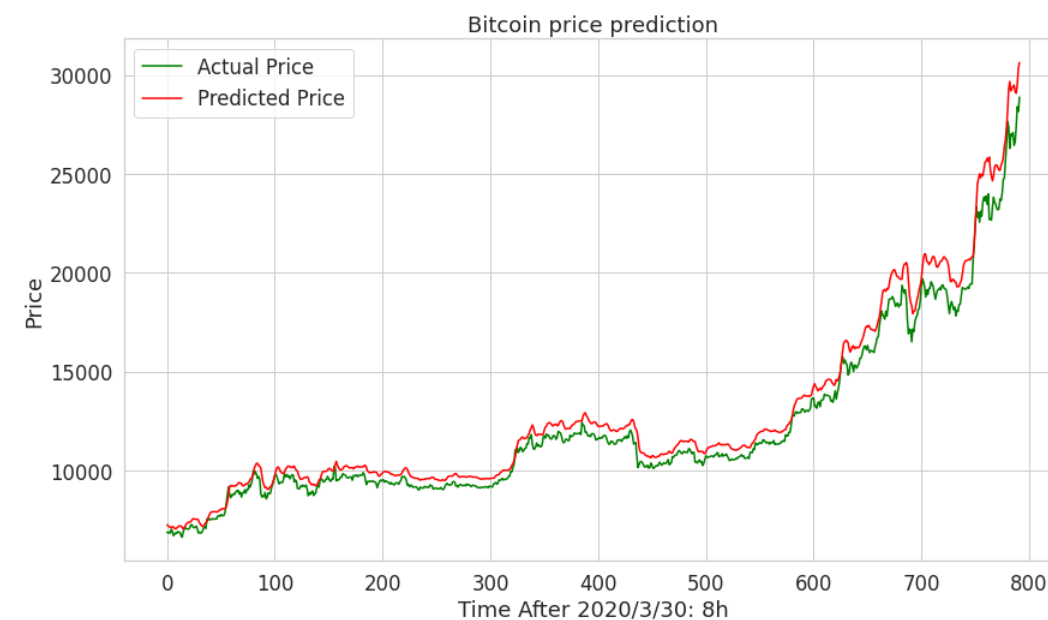


2020牛市

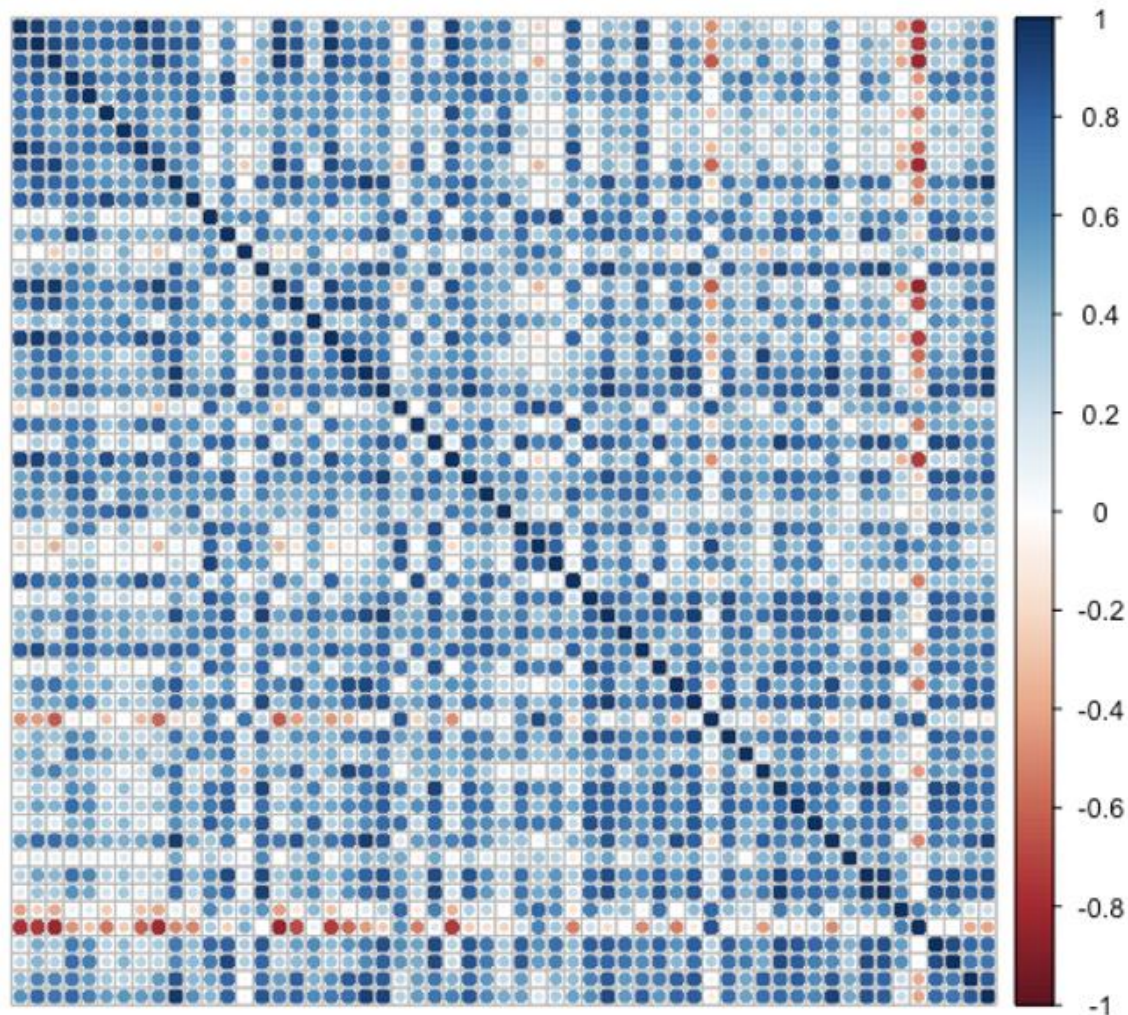
Predict : 2020.3.30~2020.12.31 Bitcoin Price

Train : 2017.8.17~2020.3.29 KLINE_INTERVAL_8HOURS

Val : 2020.3.30~2020.12.31



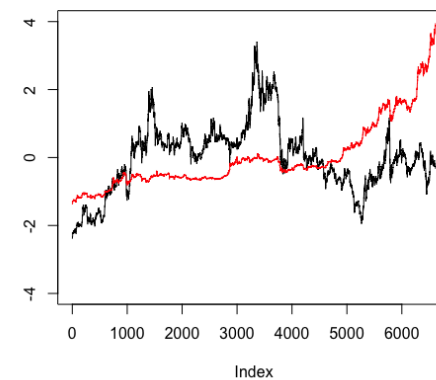
2020牛市



correlation的絕對值 <0.15 ：8種幣
correlation <-0.4 ：3種幣

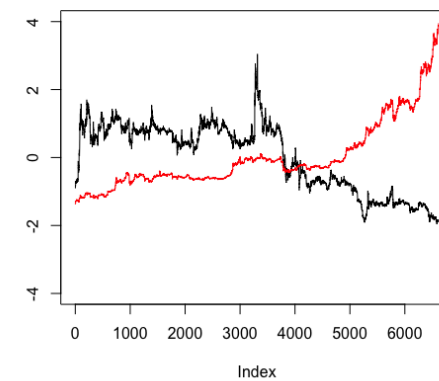
MATIC

[1] 0.05744087
2.5% 97.5%
0.008934104 0.104829321



WRX

[1] -0.7600062
2.5% 97.5%
-0.7826623 -0.7351016

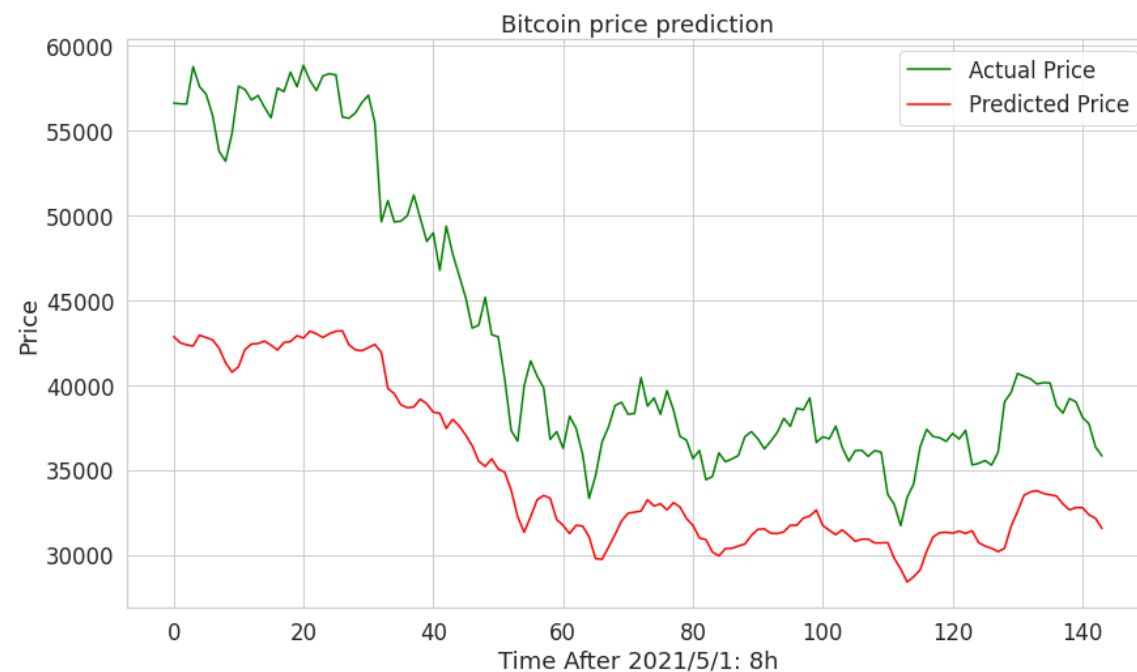


2021特斯拉

Predict : 2021.5.1~2021.6.19 Bitcoin Price

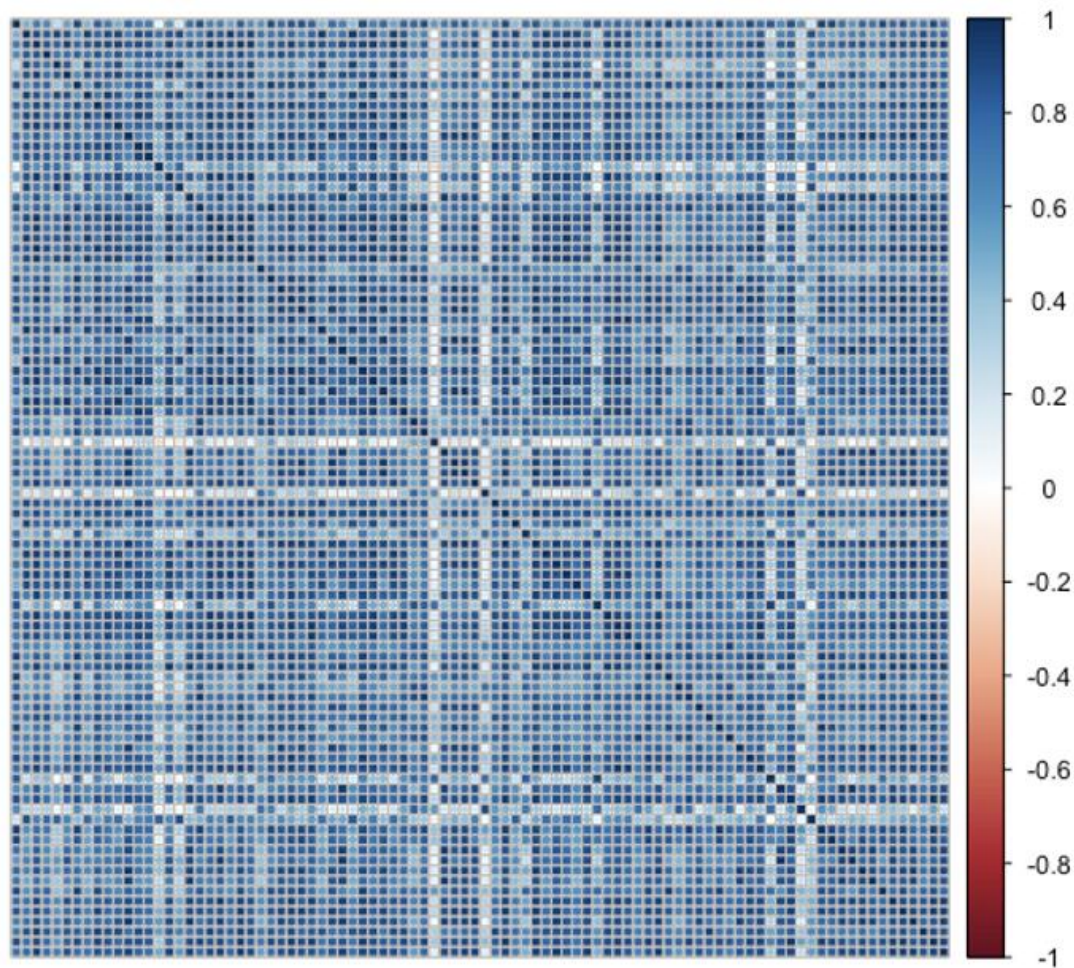
Train : 2017.8.17~2021.4.30 KLINE_INTERVAL_8HOURS

Val : 2021.5.1~2021.6.19



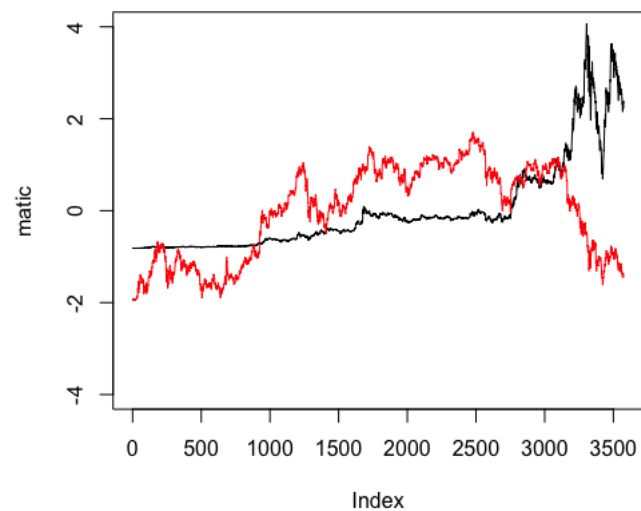
特斯拉買比特幣

correlation的絕對值<0.1： MATIC

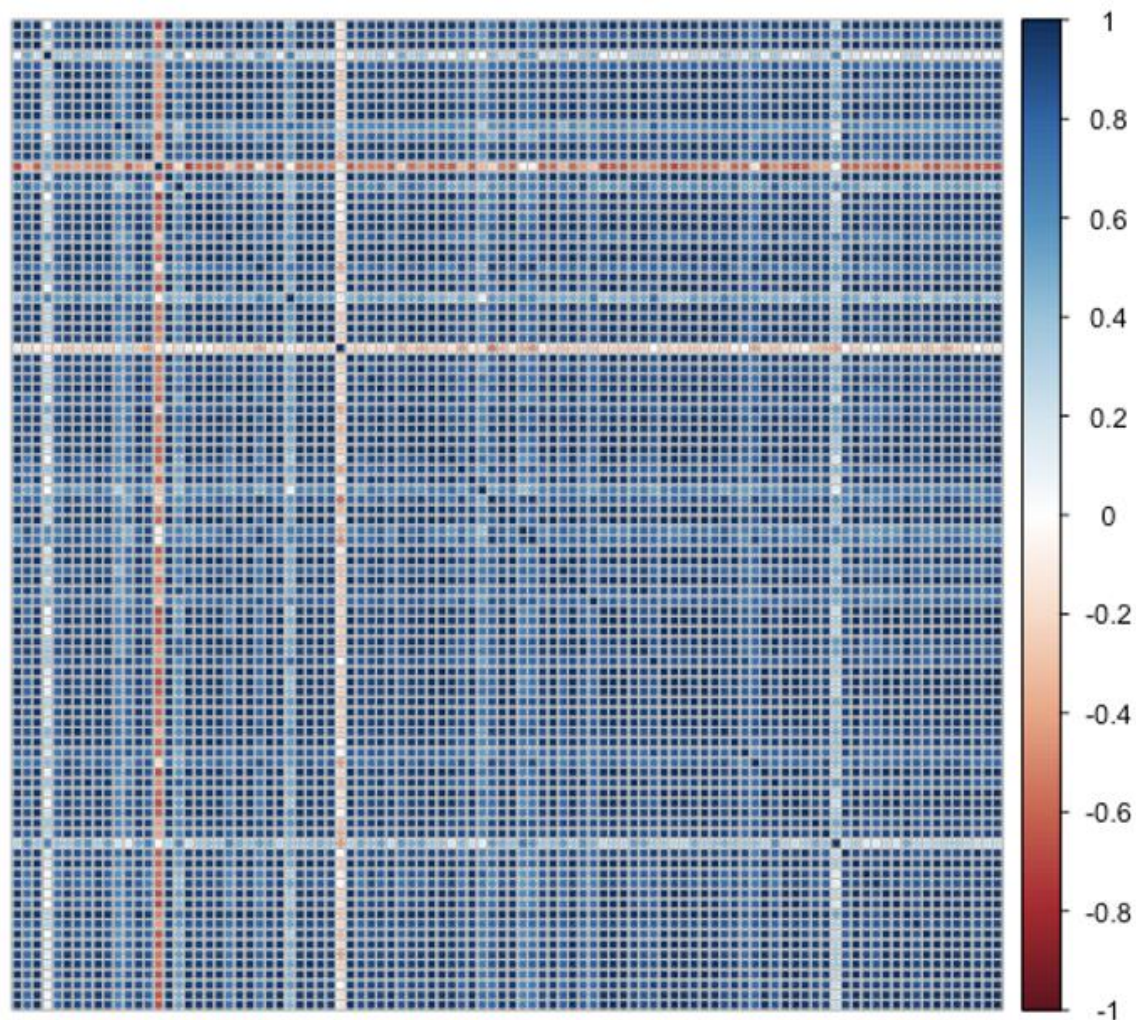


MATIC

```
[1] 0.0907682
      2.5%      97.5%
0.02984026 0.15402220
```



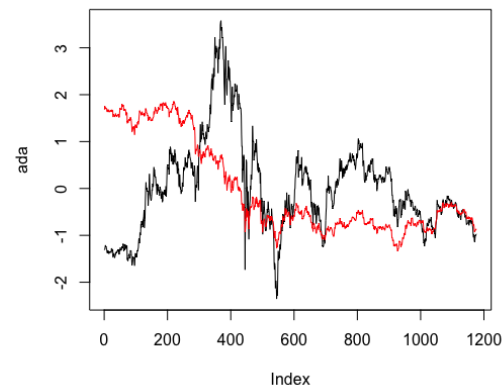
特斯拉不支持比特幣



MATIC相關性為負，而且跟其他幣也是負相關
ADA表現的跟BTC沒啥關係

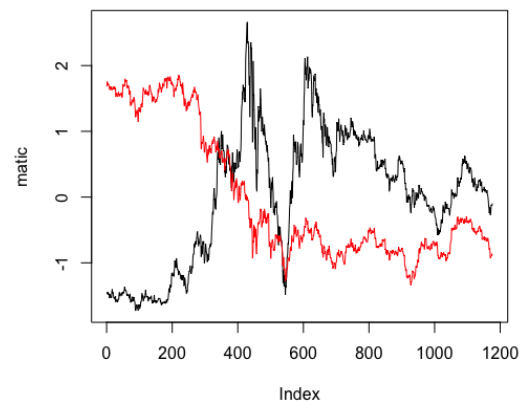
ADA

[1] 0.07338558
2.5% 97.5%
0.01612523 0.13003924

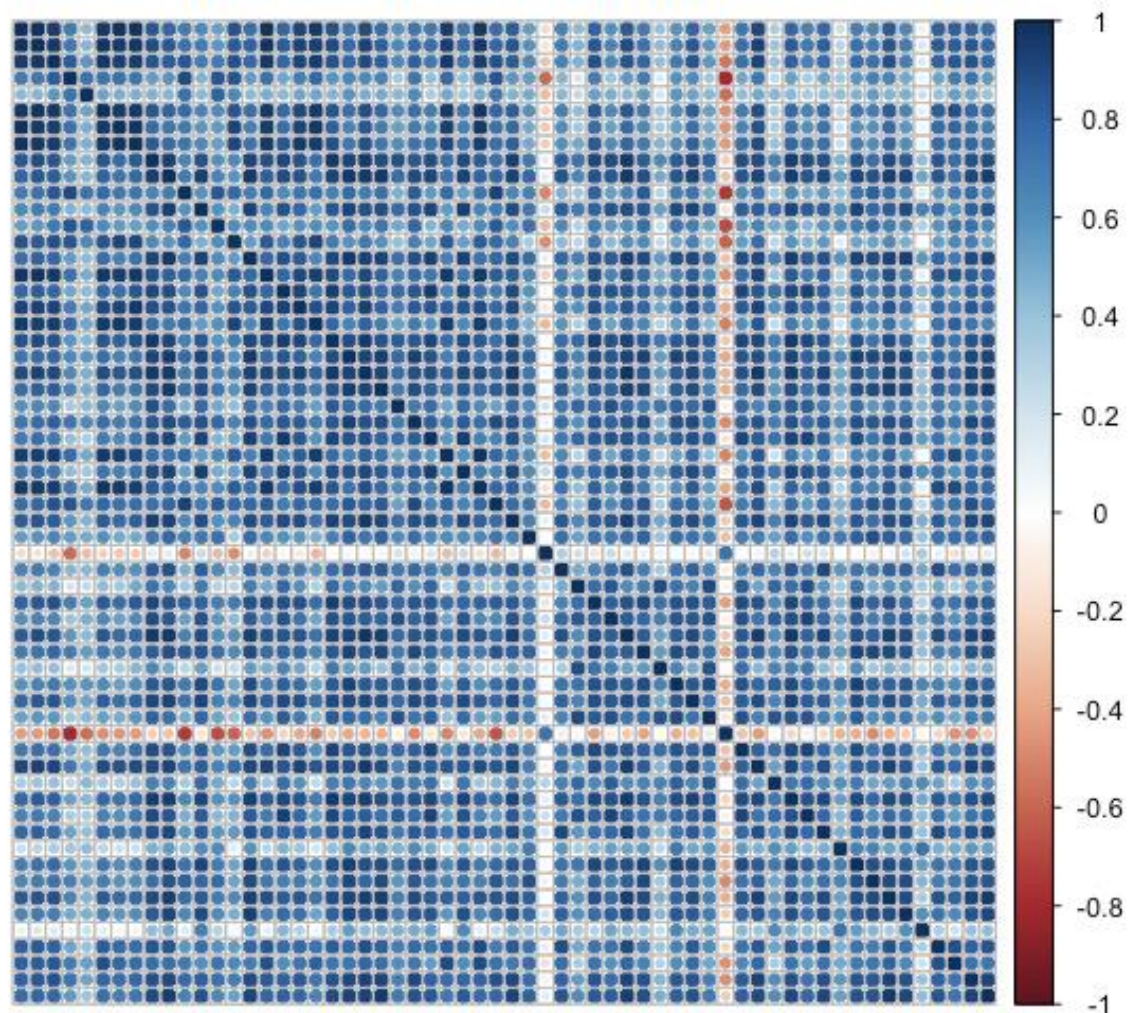


MATIC

[1] -0.696116
2.5% 97.5%
-0.7293116 -0.6603955



Defi熱潮

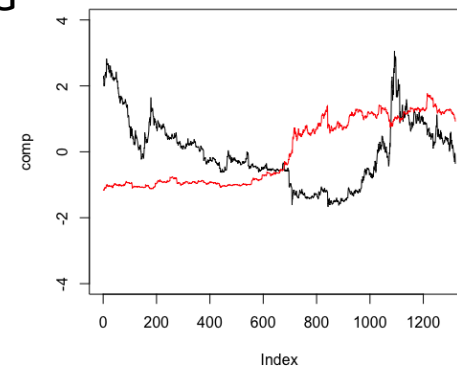


correlation的絕對值 <0.15 : WRX

correlation <0 : COMP, ONG

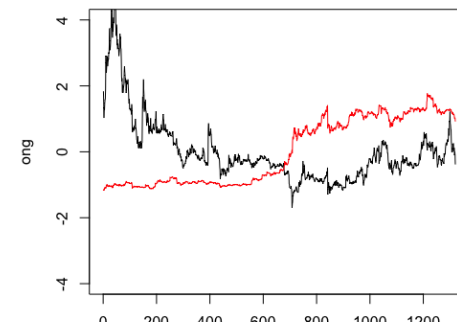
COMP

[1] -0.2001225
2.5% 97.5%
-0.2564786 -0.1416453



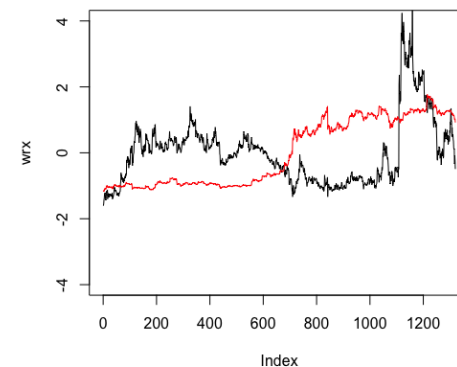
ONG

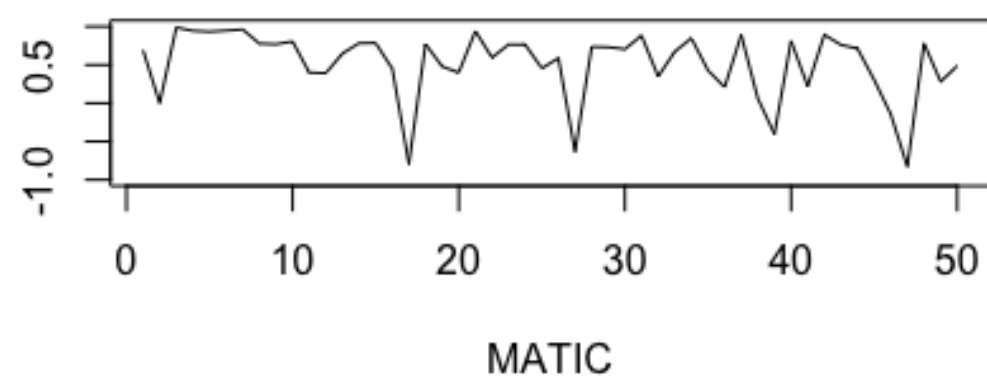
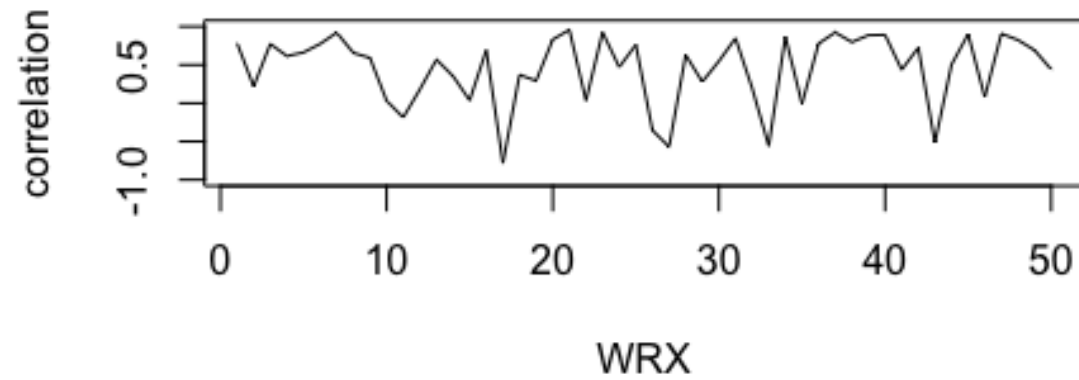
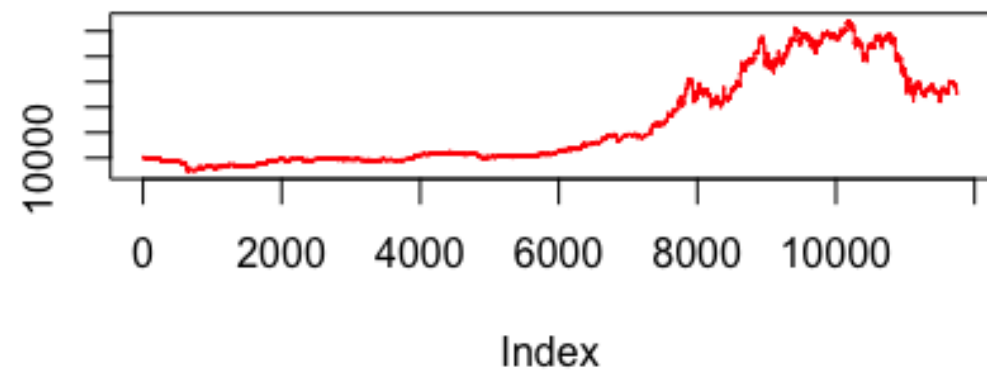
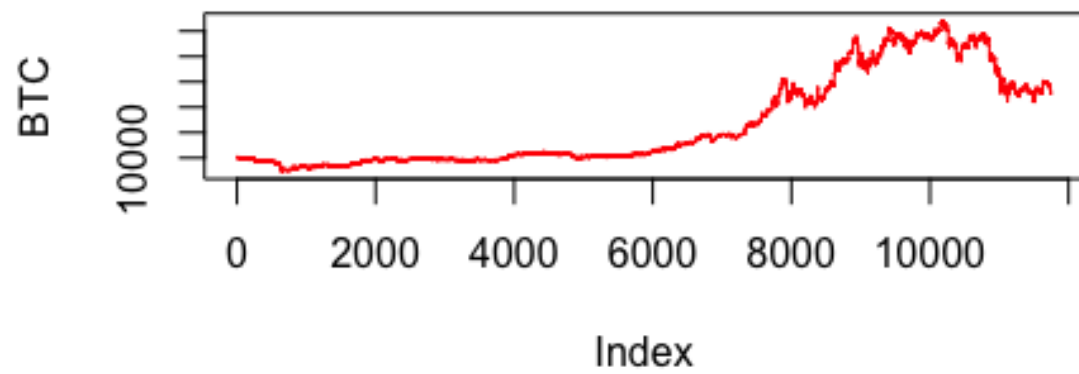
[1] -0.4213776
2.5% 97.5%
-0.4524088 -0.3901144

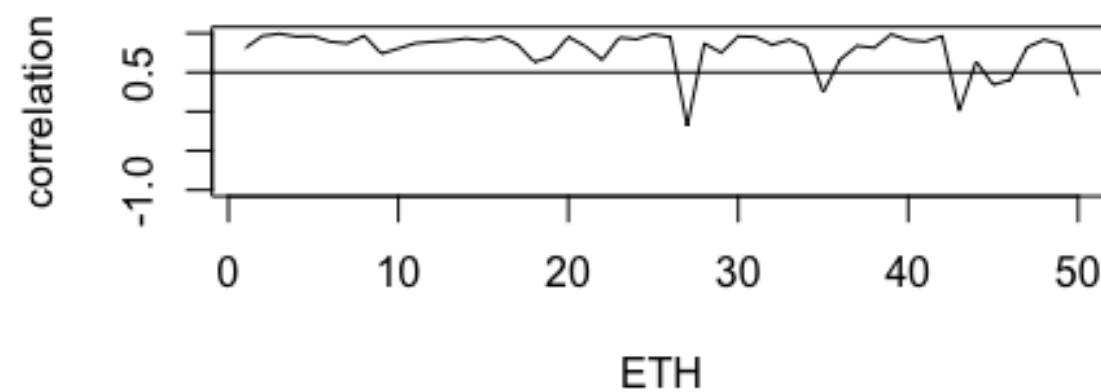
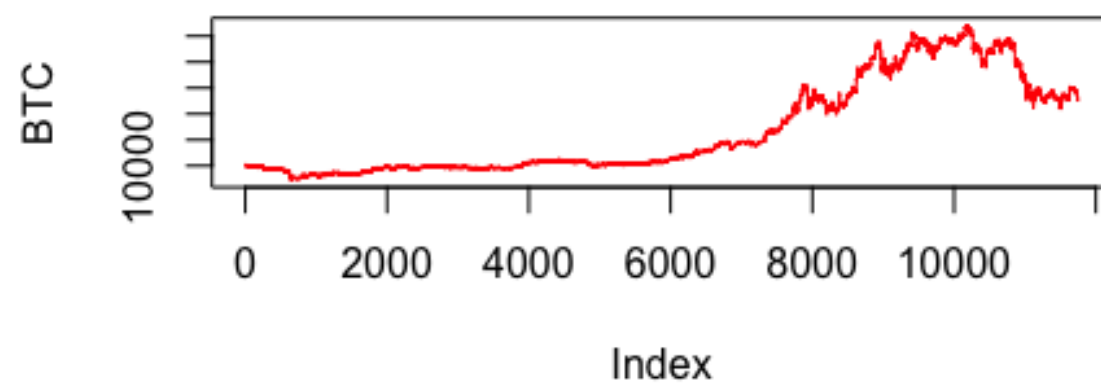


WRX

[1] 0.1030387
2.5% 97.5%
0.0363729 0.1660138







結論：

從以上兩個分析可以得出

1.長期來看，這些幣與比特幣的關係大致不變

但是，

2.短期來看，這些幣與比特幣的關係其實變動很大

所以，

假設今天我們挑選一個與比特幣長期來看呈現高度相關的A幣，我們同時投資比特幣和A幣，長時間來說，如果比特幣賺錢，以期望值來說A幣也會賺錢，

但是，如果是短期投資，A幣與比特幣的漲跌情況就很難用比特幣去預測，因為他們之間的關係變化劇烈，也就是說，他們長期看似相關，但是短期更偏向獨立。