Cryptocurrencies: A Time Series Analysis Through Logistic Regression, Prophet, Cointegration, Clustering, and Geometric Brownian Motion

Team Members: Santino Luppino, Zheng Li, Patrick Besser, Luke Ditton, Yitao Huang

Advisor: Professor Khaldoun Khashanah

Abstract

- The cryptocurrency market has experienced a high level of returns and a high level of volatility through the past few years
- Bitcoin's market capital grew from 1 billion in 2013 to 1200 billion in 2021
- Prophet seems to be a great tool for fitting and predicting cryptocurrencies prices
- UUP, SPY, USO, TNX, VIX are the most influential side factors on the price movement in the sample of cryptocurrencies.

Goals

- Determine the predictability of cryptocurrencies
- Find which method works best
- Use these findings to allow investors to be more confident in adding cryptos to their portfolio

Data

- Bitcoin, Ethereum, Binance Coin: XRP USD, Litecoin USD, & DogeCoin (DOGE-USD)
- Side Factors:SPY, VIX, IEF, DWAS, VV, UUP, GLD, IAU, USO,
- From 2018-01-01 to 2021-12-31 due to the availability of the data

Methods of Classification

Logistic Regression

BTC	0	1	ETH	0	1
0	54	54	0	128	149
1	305	315	1	225	226
BNB	0	1	DOGE	0	1
0	251	89	0	248	206
1	107	281	1	140	134
XRP	0	1	LTC	0	1
0	245	201	0	236	240
1	138	144	1	118	134

Cointegration

Used to determine if there is a correlation between several time series in the long term

- All 6 currencies are cointegrated
- Introducing DOGE has caused the model to skew due to extreme fluctuations in value
- Johansen Trace test: Allows for multiple cointegration relationships

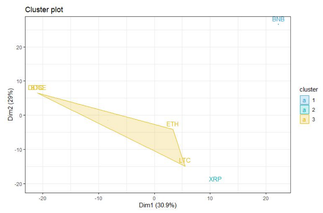
Influence of Cryptocurrencies on each other

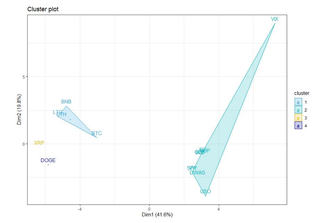
	BTC	BNB	ETH	XRP	LTC
BTC	1.00	1.00	1.00	1.00	1.00
BNB	-0.04	-0.61	0.41	0.37	-1.20
ETH	-0.65	1.50	-0.28	-0.70	0.80
XRP	-0.13	-0.31	0.43	0.14	0.03
LTC	-0.22	-1.98	-1.12	-0.82	0.34

Influence rank	Without DOGE	With DOGE
	1. BTC	1. ETH
	2. LTC	2. BNB
	3. ETH	3. XRP
	4. BNB	4. LTC
	5. XRP	5. BTC
		6. DOGE

Clustering

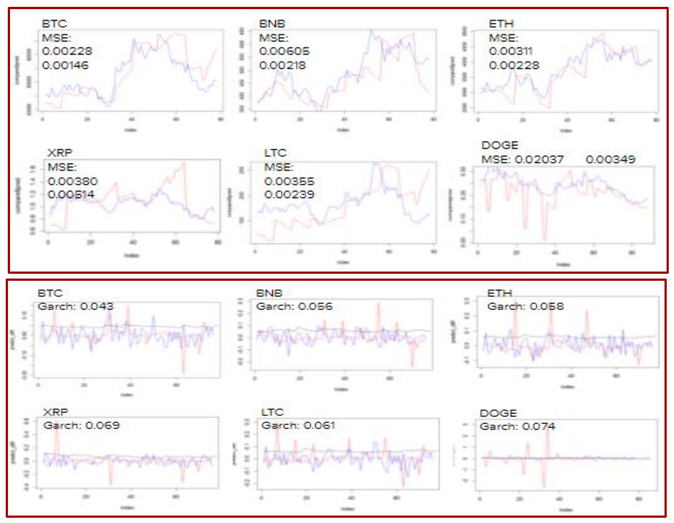
- Forms subgroups based on similarities
- Used K-means clustering, grouped on mean log daily returns
- Two cluster plots, one without side factors and one with them





Methods of Prediction

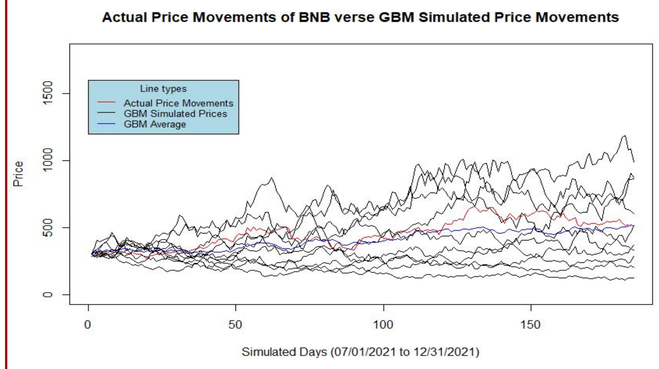
Prophet: $y(t) = g(t) + s(t) + h(t) + e(t)$



Geometric Brownian Motion

$$F_t = F_0 * e^{\left(\mu_p - \frac{\sigma_p^2}{2}\right) * t + \sigma_p Z_t}$$

$$\sigma = \sqrt{365} * s \text{ and } \mu = 365 * m + \frac{\sigma^2}{2}$$



Analysis

- **Prophet:**
  - Average of **0.0028 MSE** with side factors.
  - Side factors improve the MSE by 0.0006
- **GBM:**
  - Average of **0.0151 MSE**
- **Logistic Regression:**
  - Average accuracy of 54.86%
- **Clustering**
  - All side factors separate from crypto.
- **Cointegration**
  - the six cryptocurrencies cointegrated and each possible cryptocurrency pair is also cointegrated.

Conclusion

- Prophet is the better model for cryptocurrency prediction
  - GBM MSE is 5x larger than Prophet's
- UUP, SPY, USO are great indicators for cryptocurrency fluctuation
  - the smaller price the crypto is, the less the prediction power
- Fat tail happens more frequently in cryptos

References

- Kwon, et al. "Time Series Classification of Cryptocurrency Price Trend Based on a Recurrent LSTM Neural Network."
- Göttfert, J. (2019). Cointegration among cryptocurrencies : A cointegration analysis of Bitcoin, Bitcoin Cash, EOS, Ethereum, Litecoin and Ripple (Dissertation).
- Pichl, Lukáš, and Taisei Kaizoji (2017). "Volatility Analysis of Bitcoin Price Time Series."
- Yenidogan, et al(2018). "Bitcoin Forecasting Using ARIMA and PROPHET"