

Trabalho Prático - IA

Eric Azevedo de Oliveira
Belo Horizonte, Brasil
eric.azevedo@sga.pucminas.br

Felipe Nepomuceno Coelho
Belo Horizonte, Brasil
felipe.coelho.1265277@sga.pucminas.br

Iyan Lucas Duarte Marques
Belo Horizonte, Brasil
ildmarques@sga.pucminas.br

ACM Reference Format:

Eric Azevedo de Oliveira, Felipe Nepomuceno Coelho, and Iyan Lucas Duarte Marques. 2018. Trabalho Prático - IA. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 3 pages. <https://doi.org/XXXXXX.XXXXXX>



Figure 1. Bitcoin físico. (<https://medium.com/@JaskySingh/what-if-we-had-exams-everyday-well-everyone-would-be-better-off-f97919edeac>).

1 Introduction

A base a ser utilizada será a "Bitcoin Historical Data" que pode ser encontrada em <https://www.kaggle.com/datasets/mczielski/bitcoin-historical-data>. A base é composta por 8 atributos, sendo eles nominais, binários e numéricos:

- Timestamp - A data da instância (UNIX Timestamp)
- Open - Preço da abertura do Bitcoin no início do pregão
- High - Teto do preço do Bitcoin do pregão
- Low - Piso do preço do Bitcoin do pregão
- Close - Preço do fechamento do Bitcoin no término do pregão

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXX.XXXXXX>

- Volume_(BTC) - Volume de Bitcoins transacionado no pregão
- Volume_(Currency) - Volume de outras moedas transacionado no pregão¹
- Weighted_Price - VWAP- *Volume Weighted Average Price*

A base "Bitcoin Historical Data" possui 4857378 instâncias e é um problema de predição e apresenta o histórico de preço da moeda, desde 2012 a janeiro de 2021.

2 Related Works

Moedas outras formas de pagamento são utilizadas desde 2200 AC, agora na era da informação, estamos digitalizando esta forma milenar de relações interpessoais. Desta forma, abaixo estão apresentados trabalhos a respeito deste tema, bem como os conceitos das criptomoeda e o bitcoin, além das suas implicações nos governos e cotidiano:

As criptomoedas e as suas tecnologias como blockchain contrastaram com os métodos tradicionais, as moedas fiduciárias². como dito no artigo[3], os autores discorrem sobre as diferenças entre as criptomoedas e as moedas *fiat*. Uma das diferenças mostradas por eles são por exemplo: o modelo as quais ambas foram arquitetadas. No qual o portador da moeda fiat, através um dos seus vetores o cartões de credito, realiza uma transação que é rastreada da sua origem para o seu destino. Já as criptomoedas, são projetadas com a finalidade do anonimato e privacidade, de forma a não ser possível identificar a origem e nem os rastreo das mesmas.

Já no artigo [2], os autores iniciam a discussão contextualizando como surgiu a moeda, referenciando seu contexto histórico e fazendo comparações com o passado. Logo em seguida, é especificado que as moedas virtuais vem crescendo nos últimos tempos por vários motivos como: a não regularização da moeda, diminuindo assim o monopólio governamental, o sigilo por parte dos negociantes, o valor da moeda ser flutuante, entre outras vantagens. Neste contexto os autores ampliam a discussão mostrando que a descentralização do Bitcoin é um passo fundamental e um catalisador de um novo modelo econômico. Modelo esse que não necessita de uma figura estatal como controladora e mediadora para se obter um sistema econômico que permita as transações.

¹Moedas convertidas para BTC, como Real para BTC

²Moeda fiduciária é qualquer título não-conversível, ou seja, não é lastreado a nenhum metal (ouro, prata) e não tem nenhum valor intrínseco.[4]

Se aprofundando mais no quesito da flutuabilidade, o maior atrativo para investidores e a movimentação de capital em pregões, de acordo com [5], uma das maiores diferenças entre os tipos de moeda, as criptomoedas são bem mais voláteis em relação às tradicionais. Não só isso, mas o impacto de eventos como o coronavírus e as coberturas midiáticas afetam a moeda virtual de forma bem mais incisiva. Concluindo com as comparações das flutuações cambiais no período da pandemia, e um pequeno briefing de períodos bem notáveis.

O artigo [1], tem como objetivo: fazer uma análise sobre as influências do cenário e perspectiva da criptomoeda Bitcoin no contexto econômico e social brasileiro. Criptomoedas são códigos virtuais valorados por seu custo físico e sua flutuação dentro de um cenário financeiro próprio e descentralizado, mas que possuem a possibilidade de venda e troca por produtos tanto digitais quanto moedas fiduciárias. Argumentando ser externa a cotação de moedas fortes e já existentes como o dólar, o autor fala de como o bitcoin se tornou uma ótima alternativa de investimento, ainda mais em contextos de crises globais. Apresentando a ideia de sua criação, a ausência de um órgão regulador e a rede Blockchain, eles apresentam um potencial democrático de investimento na moeda, dando acesso a um valor de compra de, por exemplo, 1\$ e ainda obter lucros consideráveis.

No [6] o autor mostra 2 modelos de predição sobre o preço da criptomoeda bitcoin. O primeiro modelo realizava uma predição, utilizando GLM/Random Forest, direta ao valor da moeda, o mesmo que teve uma taxa de erro e desvio alta. O segundo Modelo trouxe uma estratégia um pouco diferente, sua predição é baseada na tendência de subida/queda da moeda, mediante a utilização de uma rede neural artificial, dessa forma esse modelo foi capaz de atingir 79% de acurácia se mostrando efetivo. Em uma terceira análise, são comparados Multi-Layer Perceptron (MLP) e Non-linear auto-regressive exogenous (NARX), sendo concluído que o MLP pode ser utilizado para predição, entretanto sem superar o NARX. O autor também ressalta o impacto da normalização dos dados utilizando 5 técnicas, sendo elas Log Normalization, In built MATLAB method, Standard deviation normalization, Z score normalization, Boxcox normalization. Embora os resultados sejam individuais, todas as técnicas levaram a uma melhora significativa à acurácia.

2.1 ARIMA

Para esse projeto escolheu o ARIMA, que baseia-se em métodos de predição idealizados por G. E. P. Box e G. M. Jenkins nos anos 1970. Essa técnica geralmente tem como etapas, as seguintes:

1. Seleção e identificação do modelo
2. Estimativa de parâmetros de auto regressão (AR), integração ou diferenciação (I) e média móvel (MA)
3. Verificação do modelo

O funcionamento do modelo se baseia na regressão de uma variável sob valores passados, sendo assim possível a captura de relações complexas, considerando erros e a observação de "lagged terms". Sendo assim o modelo utiliza de valores antigos gerados na série temporal para prever os novos resultados, essa aplicação dos valores passados é realizada por meio da alocação de pesos para tais valores, que são relativos a quanto tempo foram gerados em relação ao novo valor a ser calculado.

Quanto as etapas, nesse projeto primeiramente executamos um algoritmo, com a passagem de alguns pré-parâmetros, para avaliar o modelo como demonstrado abaixo:

parameters	aic
19 (1, 0, 0, 1)	173.616374
21 (1, 0, 1, 1)	174.766384
25 (1, 1, 0, 1)	175.547360
37 (2, 0, 0, 1)	175.554048
7 (0, 1, 0, 1)	175.590037

SARIMAX Results						
Dep. Variable:	Weighted_Price_box	No. Observations:	112			
Model:	SARIMAX(1, 1, 0)x(0, 1, [1], 12)	Log Likelihood:	-83.808			
Date:	Thu, 28 Apr 2022	AIC	173.616			
Time:	13:38:55	BIC	181.402			
Sample:	12-31-2011	HQIC	176.766			
	- 03-31-2021					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.4253	0.085	5.030	0.000	0.260	0.591
ma.S.L12	-0.9943	4.568	-0.218	0.828	-9.948	7.960
sigma2	0.2440	1.100	0.222	0.825	-1.913	2.401
Ljung-Box (L1) (Q):	0.02	Jarque-Bera (JB):	3.59			
Prob(Q):	0.89	Prob(JB):	0.17			
Heteroskedasticity (H):	1.16	Skew:	0.29			
...						

Figure 2. Trecho de log de resultado da análise para identificação do modelo.

Após a análise gerada pelo algoritmo, foi possível determinar os hiperparâmetros (p,d,q) OU ((AR),(I),(MA)) a serem utilizados no ARIMA, a fim de obter os melhores resultados possíveis, sendo definidos como (1,1,0).

Após os ajustes de hiperparâmetros, foi possível realizar as predições com o modelo ARIMA.

3 Metodologia

3.1 Pré-processamento

No pré-processamento da base, inicialmente foram feitos alguns passos antes da seleção do modelo e de outras ações. No processo da obtenção dos dados, foram gravados a data e a hora em formato UNIX, que para melhor manuseio da mesma, foi prontamente convertida para o formato padrão UTC. Além disso, foi feito o resample baseado em certas frequências de tempo, sendo as mesmas, diária, mensal, trimestral e anual.

Após esse tratamento inicial, foi-se então analisado as tendências da criptomoeda nas frequências de tempo descritas acima, as quais podem ser observadas aqui:

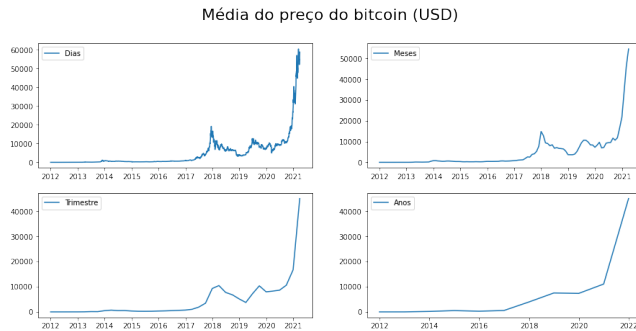


Figure 3. frequência do bitcoin nos períodos (da esquerda para a direita) diária, mensal, trimestral e anual .

3.2 Stationary check

No stationary check, que está totalmente interligada a plotagem dos dados, verificação de tendência e componentes, com esses dados iremos determinar uma série temporal em que foi gerada por um processo Stationary. Observando apenas seu plot nos dados.

3.3 STL Decomposition

Sobre o Stl Decomposition, o utilizamos para decompor uma série temporal em dois componentes de tendência relacionados a predição da média por mês de nossos dados. Utilizamos a Seasonal pela sua duração da tendência ser mais suave e nos dar em torno de 150 da temporariedade, e utilizamos o modelo da Regular para termos também uma duração mais sazonal.

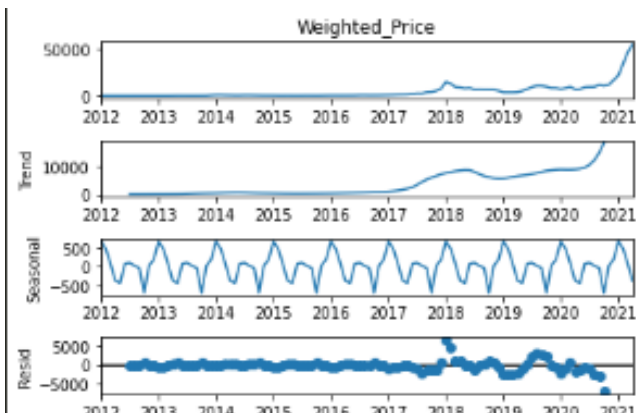


Figure 4. Stationary and Stl Decomposition

3.4 Seasonal differentiation

A Seasonal differentiation foi utilizado, pois, supondo que temos dados mensais, nas quais ha 12 períodos de uma estação em si. Desta forma, utilizamos a diferença sazonal de Y no período de t , que é $Y_t - Y_{t-12}$. Insto onde a diferença sazonal de Y é idêntica e sem autocorrelação, com isso o Y é descrito por um modelo distribuído aleatoriamente.

3.5 Regular differentiation

Regular Differentiation, foi utilizado, pois, supondo que temos dados mensais, nas quais ha 12 períodos de uma estação. Com isso, utilizamos a diferença sazonal de Y no período de t , que é $Y_t - Y_{t-12}$. Onde, se a diferença sazonal de Y for idêntica e sem autocorrelação, com isso o Y é descrito por um modelo de distribuído aleatoriamente. Para utilizá-lo pensamos em uma forma mais eficaz para ajustar o Índice e subtraí-lo por cada Regular Differentiation que foi observado na mesma estação, gerando assim características mais brutas da sazonalidade de uma série em questão.

References

- [1] Genisson dos Santos Pinheiro, Alvani Bomfim de Sousa Júnior, and Josivan dos Santos Moura. 2021. Cenário e perspectiva da criptomoeda bitcoin no Brasil. *Entrepreneurship* 5, 1 (2021), 1–8.
- [2] Sandro GIBRAN e Sandra Mara LIMA e Sérgio Itamar ALVES JÚNIOR e Roberto José KOSOP. 2017. O BITCOIN E AS CRIPTOMOEDAS: RE-FLEXOS JURÍDICOS EM UM COMÉRCIO GLOBALIZADO. *Administração de Empresas em Revista* 1, 12 (2017).
- [3] Mohammad Rabiul Islam, Rizal Mohd Nor, Imad Fakhri Al-Shaikhli, and Kabir Sardar Mohammad. 2018. Cryptocurrency vs. Fiat Currency: Architecture, Algorithm, Cashflow amp; Ledger Technology on Emerging Economy: The Influential Facts of Cryptocurrency and Fiat Currency. In *2018 International Conference on Information and Communication Technology for the Muslim World (ICT4M)*. 69–73. <https://doi.org/10.1109/ICT4M.2018.00022>
- [4] F Lara. 2008. *Um estudo sobre moeda, juros e distribuição*. Ph.D. Dissertation. Tese de doutorado apresentada ao PPGE/UFRJ.
- [5] Zaghun Umar, Francisco Jareño, and María de la O González. 2021. The impact of COVID-19-related media coverage on the return and volatility connectedness of cryptocurrencies and fiat currencies. *Technological Forecasting and Social Change* 172 (2021), 121025. <https://doi.org/10.1016/j.techfore.2021.121025>
- [6] Siddhi Velankar, Sakshi Valecha, and Shreya Maji. 2018. Bitcoin price prediction using machine learning. In *2018 20th International Conference on Advanced Communication Technology (ICACT)*. 144–147. <https://doi.org/10.23919/ICACT.2018.8323676>