

# Aprendizado de Máquina para a Previsão do Comportamento de Preços da Criptomoeda Ethereum

*Elaboration of a machine learning algorithm for predicting the price behavior of the Ethereum cryptocurrency*

**Lucas Penteado L. Silva<sup>1</sup>**

 [orcid.org/0000-0001-6228-633X](https://orcid.org/0000-0001-6228-633X)

**Luiz A. da Silva Junior<sup>1</sup>**

 [orcid.org/0000-0002-4712-2502](https://orcid.org/0000-0002-4712-2502)

**Josafat M. Falcão Neto<sup>1</sup>**

 [orcid.org/0000-0003-0264-8225](https://orcid.org/0000-0003-0264-8225)

**Geidson B. C. de Souza<sup>1</sup>**

 [orcid.org/0000-0003-4936-895X](https://orcid.org/0000-0003-4936-895X)

<sup>1</sup>Escola Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: [lucasplds@gmail.com](mailto:lucasplds@gmail.com)

**DOI: 10.25286/rep.v7i3.2445**

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Lucas Penteado L. Silva; Luiz A. da Silva Junior; Josafat M. Falcão Neto; Geidson B. C. de Souza. Aprendizado de Máquina para a Previsão do Comportamento de Preços da Criptomoeda Ethereum. Revista de Engenharia e Pesquisa Aplicada, Recife, v. 7, n. 3, p. 1-6.

## RESUMO

Elaboração de algoritmos de aprendizado de máquina para a previsão do comportamento de preços da criptomoeda Ethereum, utilizando-se uma base de dados pública (Kaggle). Os modelos elaborados foram do tipo linear (ARIMA, séries temporais) e não-linear (três modelos de redes neurais LSTM). Como melhor resultado, verificou-se que um dos modelos não lineares foi capaz de realizar previsões distantes em média de 4,32% dos preços reais.

**PALAVRAS-CHAVE:** Aprendizado de máquina, redes neurais, criptomoedas, Ethereum, Séries Temporais, ARIMA, LSTM.

## ABSTACT

Development of machine learning algorithms for the prediction of the behavior of prices of the cryptocurrency Ethereum using a public and widely available database (Kaggle). One linear model was developed (ARIMA, time series analysis), while three non-linear models were developed (deep learning models, using long term short memory framework). As a result, it was noticed that one of the deep learning models was able to predict prices that were, in average, only 4,32% far apart from real prices.

**KEY-WORDS:** Machine learning, time series analysis, cryptocurrencies, Ethereum, deep learning, ARIMA, LSTM.

# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

De acordo com o escritor norte americano Mark Twain, a história não se repete, mas sim rima com bastante frequência. Tal frase contém um saber empírico, relacionado com boa parte das ações humanas ao longo dos tempos. Não raro, determinados padrões de comportamento levam à predição de relações de causa e efeito entre fatos e dados, sendo os componentes básicos para a construção de informações, conhecimentos e experiências, respectivamente [1].

Por sua vez, o advento da transformação digital vem servindo como um importante catalisador desta construção de conhecimentos e experiências [2]. Através da conjunção de quatro tecnologias profundamente disruptivas – a computação em nuvem, a *internet* das coisas (IoT), o chamado *big data* e a inteligência artificial (IA) – muitas áreas vêm apresentando inovações radicais e inesperadas.

Isto posto, embora o Bitcoin seja a mais conhecida das criptomoedas a ponto de ser inclusive aceita como moeda oficial pelo governo de El Salvador [4], seu recente sucesso abriu espaço para outras soluções similares, como as criptomoedas *Ethereum* e *Cardano*.

## 1.2 DESCRIÇÃO DO PROBLEMA

Prognosticar preços de mercado para ativos econômicos é algo essencial no que se refere a investimentos financeiros, posto que permite aos investidores uma alocação mais otimizada de seu capital [5].

Logo, prever de modo aproximado a tendência de alta ou de baixa nos preços de mercado de criptomoedas pode resultar em ganhos financeiros ou, ao menos, na preservação do capital investido.

Todavia, a ideia de prever o comportamento de preços de mercado de ativos financeiros é fortemente rechaçada por adeptos da teoria dos mercados eficientes (*market efficiency hypothesis*), a qual postula que os preços de mercado refletem todas as informações já disponíveis ao público, impossibilitando, portanto, a previsão de seu comportamento. Do contrário, haveria a possibilidade de geração de lucros infinitos, resultantes de uma "máquina de gerar dinheiro", fato este impossível em uma economia com crescimento limitado [6].

## 1.3 OBJETIVOS

O objetivo principal deste artigo é o de, usando conhecimentos de ciência de dados (aprendizado de máquina e séries temporais, notadamente), elaborar algoritmos que realizem a predição do comportamento de preços da criptomoeda *Ethereum* por meio da análise do histórico de registros públicos de negociação a partir de 2018.

## 1.4 JUSTIFICATIVA

Dado o crescente interesse pelas criptomoedas nos últimos anos, a análise do comportamento dos preços de mercado de criptomoedas foi introduzida ao rol dos objetos analisados por meio de técnicas de ciência de dados, especialmente no que tange ao aprendizado de máquina.

Trata-se de uma área de estudo promissora, uma vez que se refere a um mercado recém criado (criptomoedas), cujo funcionamento ainda é suscetível a otimizações (*market efficiency*).

Tal constatação mostra-se ainda mais significativa quando considerado que as criptomoedas possuem capitalização de mercado de aproximadamente 2,3 trilhões de dólares e volume diário de negociação de 84 bilhões de dólares.

Logo, a utilização de ferramentas computacionais que sejam capazes de trazer insights na precificação de criptomoedas é um tópico que se mostra cada vez mais relevante, em especial se considerado o volume massivo de dados resultantes de transações envolvendo tal classe de ativos.

## 1.5 ESCOPO NEGATIVO

Este artigo possui natureza pedagógica, isto é, objetiva descrever de modo detalhado os passos e as etapas percorridas para a elaboração de um algoritmo de aprendizado de máquina a fim de prever o comportamento dos preços de mercado da criptomoeda *Ethereum*.

Sendo assim, não se tem como objetivo primário a utilização de informações de todas as criptomoedas existentes, mas sim a utilização de uma amostra, a qual contém registros coletados a partir de 2018. Ademais, também não será realizado o levantamento dos registros de negociação de forma independente, posto que a base de dados utilizada para análise é disponibilizada pelo portal *Kaggle*.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 CRIPTOMOEDAS E BLOCKCHAIN

Este trabalho se deterá na comercialização (trading) de criptomoedas, as quais usam uma tecnologia específica: o blockchain. Como o nome indica, blockchain é uma cadeia (do inglês "chain") de objetos (ou blocos) computacionais repletos de informações, em que cada bloco contém uma série de transações digitalmente assinadas, verificadas e armazenadas de forma distribuída na nuvem. Tudo isso de forma permanente, interdependente e inalterável, permitindo, assim, o acesso e a verificação posteriores do conteúdo de cada bloco pelos respectivos públicos de interesse [3].

Dentre as variadas aplicações do blockchain, cresce o interesse do grande público nas criptomoedas – sendo a mais conhecida delas o Bitcoin –, as quais nada mais são do que transações eletrônicas que seguem os princípios estabelecidos no blockchain, a saber:

1. Descentralização: as transações são validadas pelas partes envolvidas, dispensando terceiros como instituições financeiras e bancos centrais, de forma distribuída e simultânea no domínio da nuvem através de centros de armazenamento e de processamento de dados para transações (nós);

2. Cumulatividade: as transações são armazenadas de forma unívoca e continuada nos blocos, ou seja, todos os registros, uma vez realizados, permanecem armazenados (sendo assim uma cadeia de eventos) e acessíveis ao longo do tempo em caso de auditorias e verificações;

3. Proof-of-Work: Qualquer nova transação é replicada imediata e automaticamente a todos os novos blocos vinculados à rede, tornando virtualmente impossível a realização de fraudes. Assim, tendo em vista que as transações são descentralizadas e também cumulativas, caso haja a tentativa de alteração de uma dada transação de forma fraudulenta, seria necessário fazê-lo em todos os blocos vinculados à rede ao mesmo tempo e exatamente na mesma sequência; do contrário, esta transação será rejeitada.

4. Robustez: Presumindo que a maior parte do poder computacional de uma rede esteja em poder de nós que não estejam atuando para fraudá-la, a maior cadeia de blocos de informações (blockchain) será processada a um ritmo mais rápido do que a capacidade de processamento de eventuais fraudadores;

5. Flexibilidade: Os nós podem deixar a rede e retornar a qualquer momento, aceitando a cadeia mais longa disponível como proof-of-work do que tiver acontecido durante a sua ausência.

### 2.2 SÉRIES TEMPORAIS

Uma vez que os dados relativos às cotações dos ativos estão, via de regra, disponíveis regularmente na sequência em que acontecem, ou seja, em ordem cronológica, o conceito de séries temporais [8] vem naturalmente ao pensamento.

As cotações, devidamente coletadas em intervalos de tempo regulares, formam uma coleção de valores que frequentemente estão correlacionados com os valores imediatamente anteriores, razão pela qual é possível prever seu comportamento até certo ponto no tempo futuro. Essa capacidade é mensurada pela autocorrelação, que é a medida do grau de influência de quanto uma leitura de uma dada variável aleatória é capaz de atuar sobre suas leituras vizinhas [9].

Por sua vez, fazer inferências sobre o comportamento do modelo a partir de seus dados também envolve separá-los em dois domínios: treinamento, onde os dados servem de orientação para o modelo apreender o seu funcionamento, e validação, através do qual a acurácia das respostas produzidas pelo modelo pré-treinado é colocada à prova. Ambos os processos são fundamentais para a obtenção de respostas confiáveis, buscando evitar tanto que o modelo "decore" os dados de treinamento (overfitting) quanto que produza resultados que não fazem qualquer sentido com relação aos dados (underfitting).

Para este estudo, faremos uso de modelagem matemática em forma de série temporal para treinamento e validação dos dados, usando a métrica do erro médio quadrático (MSE) para averiguar a precisão do modelo nos resultados coletados.

### 2.3 APRENDIZADO DE MÁQUINA

Por se tratar de um mercado novo e em expansão, o mercado de criptomoedas é passível de constantes mudanças de comportamento ainda não mapeadas ou previsíveis – disso resulta que utilizar Machine Learning para acompanhamento e entendimento dessa grande quantidade de dados e variáveis (como os da Tabela 1), onde não são passíveis de emprego de fórmulas e equações de

modelagem, não é uma opção, mas uma necessidade intrínseca nesse recente mercado.

Diante do exposto, o Deep Learning (redes neurais), parte da área de Machine Learning, é a tecnologia responsável por apoiar na compreensão, classificação, reconhecimento e detecção de dados e informações.

A precificação de ativos empíricos (no caso em estudo, as criptomoedas) é um ramo importante da pesquisa financeira. Métodos de aprendizado de máquina têm sido aplicados cada vez mais neste domínio, devido à capacidade de selecionar com flexibilidade entre um número potencialmente grande de recursos e aprender relacionamentos complexos e de alta dimensão entre recursos e destinos [11].

Com o modelo treinado sobre as bases criadas das séries temporais, é construída a rede neural com quatro camadas de neurônios sobre a base de treinamento (fitting), transformando-a em base de dados diária e normalizando-a (para efeito de eficiência do modelo), criando-se em seguida as bases de treino da rede neural num período de 05 dias.

Removendo a normalização dos dados da base de treino e por meio do tabelamento dos dados obtidos (previsões), faz-se comparação com os valores tabelados disponíveis na base de dados real.

### 2.4 REDES NEURAIS RECORRENTES

As redes neurais recorrentes (RNN) são uma técnica conhecida no campo da inteligência artificial e ciência de dados e são usadas para realizar previsões sobre dados. Uma rede neural RNN é um tipo de rede onde os neurônios fornecem feedback uns aos outros por meio de um loop [12].

Esse tipo de rede tem tido um grande sucesso em diversos tipos de problemas: reconhecimento de fala, modelagem de idiomas tradução de textos, transcrição de áudio, geração de nomes, legendas de imagens, etc., e quase todos os resultados empolgantes baseados em redes neurais recorrentes são alcançados com LSTMs (DEEP LEARNINGBOOK, 2022).

As LSTMs são tipos de redes neurais recorrentes introduzida por Hochreiter e Schmidhuber [13], capazes de aprender sequências e processamento temporal e, portanto, podem ser aplicadas a diversos tipos de problemas de previsão tão diferentes quanto preços de criptomoedas (SIDDHARTH e KAUSHIK, 2022; BREMMER, 2018) e a geração de energia de usinas eólicas [14].

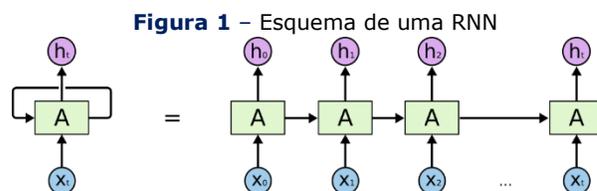


Figura 1 – Esquema de uma RNN

Fonte: DeepLearningBook.

### 2.5 TRABALHOS RELACIONADOS

Em artigo intitulado 'Forecasting and trading cryptocurrencies with machine learning under changing market conditions', Sebastião e Godinho [7] buscaram apresentar a previsibilidade do comportamento de preços de três criptomoedas (Bitcoin, Ethereum e Litecoin) por meio do uso de técnicas de aprendizado de máquina. Também foram avaliadas estratégias de investimento decorrentes dos resultados obtidos.

Primeiramente, foi utilizada uma base de dados para treinamento (learning) do algoritmo constante do histórico de negociação dos respectivos ativos entre 2015 e 2019.

Em sequência, foram elaborados 18 modelos individuais, os quais foram desenvolvidos a partir de composição de modelos (ensemble).

Como resultado, foram obtidos modelos que resultaram em estratégias de investimento com retornos líquidos anuais de até 9,62% para a criptomoeda Ethereum, o que dá suporte à afirmação de que seria possível a utilização de aprendizado de máquina para investimentos financeiros.

No estudo de Karakoyun e Cibikdiken [10] denominado Comparison of ARIMA Time Series Model and LSTM Deep Learning Algorithm for Bitcoin Price Forecasting, os modelos ARIMA e o LSTM foram comparados para estimar o preço futuro da criptomoeda Bitcoin.

Neste trabalho, os preços dos próximos 30 dias foram estimados com os modelos obtidos.

Como resultado, foram obtidos aproximadamente erros percentuais médios absolutos (mean absolute percentage error) de 11,86% com ARIMA e de 1,40% com LSTM. Em outros testes de acurácias os autores observaram que quando comparados, o modelo LSTM apresenta resultados melhores.

### 3 MATERIAIS E MÉTODOS

#### 3.1 DESCRIÇÃO DA BASE DE DADOS

Conforme já mencionado, a base de dados analisada foi disponibilizada pelo portal Kaggle, contendo 24.236.806 linhas e 10 colunas (atributos), abrangendo transações de criptomoedas entre os períodos de 01/01/2018 e 21/09/2021, conforme mostrado no Quadro 1.

**Quadro 1** - Atributos da Base de Dados

<b>timestamp</b>	Registro temporal da transação, medido no padrão Unix (segundos a partir de 01/01/1970), sempre a cada 60s
<b>Asset_ID</b>	Código para identificação das criptomoedas
<b>Count</b>	Número de negociações realizadas no minuto
<b>Open</b>	Preço de abertura do ativo no início do minuto (em dólar)
<b>High</b>	Maior preço do ativo no minuto (em dólar)
<b>Low</b>	Menor preço do ativo no minuto (em dólar)
<b>Close</b>	Preço de encerramento do ativo no fim do minuto (em dólar)
<b>Volume</b>	Volume comercializado por minuto
<b>VWAP</b>	Preço médio ponderado pelo volume durante o minuto (em dólar)
<b>Target</b>	Retorno do ativo nos últimos 15 minutos (ajustado para logaritmo)

Fonte: Kaggle.

#### 3.2 ANÁLISE DESCRITIVA

Primeiramente, ao se analisar a base de dados que contém os registros das transações com criptomoedas, verifica-se que todos os ativos possuem valores omissos na coluna 'Target' conforme mostrado na Tabela 1.

**Tabela 1** - Ativos da Base de Dados

Asset_ID	Ativo	Valores Omissos na coluna Target
0	Binance Coin	13.415
1	Bitcoin	304
2	Bitcoin Cash	4.861
3	Cardano	18.731
4	Dogecoin	144.974

5	EOS.IO	2.302
6	Ethereum	340
7	Ethereum Classic	9.326
8	IOTA	193.106
9	Litecoin	521
10	Maker	40.025
11	Monero	239.603
12	Stellar	61.816
13	TRON	21.014

Fonte: Kaggle.

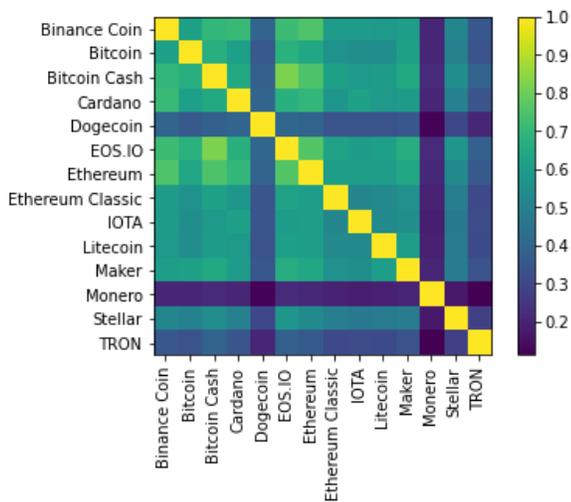
Também foi verificado que todas as criptomoedas possuem lacunas temporais nos registros de transação. Quanto a isso, não é despidendo mencionar que tal inconsistência, para fins de análise, foi sanada por meio do método *reindex*, utilizando-se o valor constante do último registro válido para situações em que haja lacuna de registros.

Finalmente, por meio da elaboração de uma matriz de correlação, foi verificado que nem todos os ativos analisados sofrem forte correlação com a variação de preços do *Bitcoin*, apesar de esta ser a criptomoeda de maior capitalização de mercado.

Isso posto, por meio da análise individual dos registros das criptomoedas *Bitcoin* e *Ethereum*, foi possível constatar que:

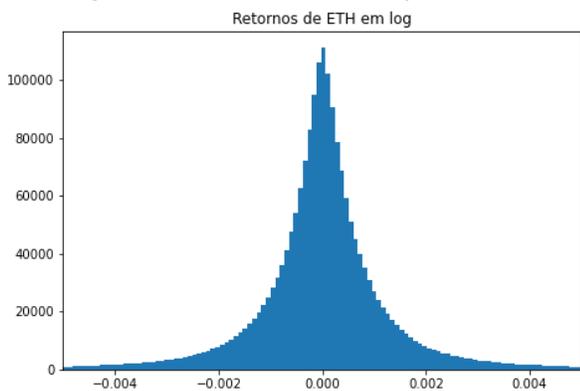
- As variações de preço (retornos) da criptomoeda *Ethereum* em cada intervalo de tempo registrado, após serem quantificadas em bases logarítmicas, apresentam distribuição que se assemelha à distribuição normal;
- O volume de criptomoedas negociadas em um determinado tempo não possui correlação significativa com a variação do preço do respectivo ativo. Todavia, o número de negociações (*trades*), por sua vez, apresenta correlação relevante com a variação do preço; e
- A variação dos preços das criptomoedas *Bitcoin* e *Ethereum* apresentam correlação alta, na medida em que representam os dois criptoativos mais negociados no mercado.

**Figura 2** – Correlação entre os ativos



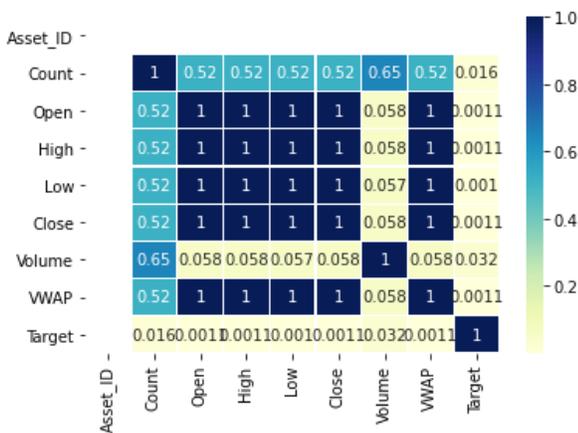
Fonte: Kaggle.

Figura 3 – Retornos - Distribuição normal



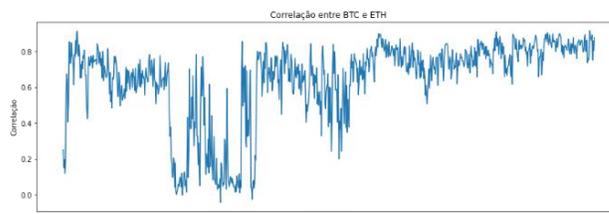
Fonte: Autoria própria.

Figura 4 – Correlação preço vs. nº de trades



Fonte: Autoria própria.

Figura 5 – Correlação entre BTC e ETH



Fonte: Autoria própria.

### 3.3 ELEMENTOS PÓS-TEXTUAIS

Inicialmente, no que se refere à área das séries temporais, é de relevante importância o conceito de autocorrelação.

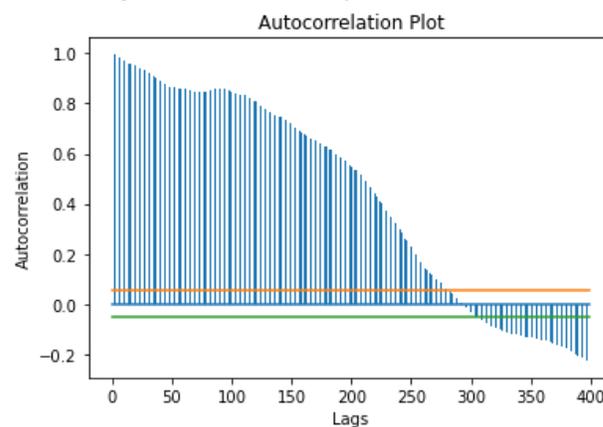
A autocorrelação almeja evidenciar a relação e a influência de registros passados de uma série de valores nos registros mais atuais.

Dessa forma, a autocorrelação pode ser calculada com valores que variam de -1 a 1, a depender do grau da correlação de um valor passado com um valor presente.

Um valor de autocorrelação próximo a 1 ou a -1 indica uma forte influência do respectivo registro passado no registro atual, ao passo que um valor próximo a 0 indica uma fraca influência.

In casu, na série temporal analisada, isto é, no comportamento dos preços da criptomoeda Ethereum, é possível verificar que os preços medidos diariamente demonstram autocorrelação positiva aproximadamente o 300º dia antecedente. A partir disso, os preços indicam uma autocorrelação negativa.

Figura 6 - Autocorrelação do Ethereum



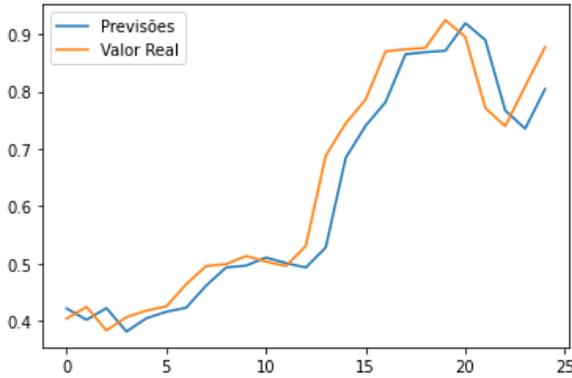
Fonte: Autoria própria.

Isto posto para fins de modelagem da previsão do comportamento dos preços do Ethereum, foi utilizado um modelo do tipo ARIMA (AutoRegressive Integrated Moving Average). Registre-se que um modelo ARIMA é baseado em álgebra linear, e de

modo a mostrar a relação de uma variável da série temporal  $X_t$  com os seus valores passados ( $X_{t-1}$ ,  $X_{t-2}$ ,  $X_{t-3}$  etc.), conforme a equação abaixo:

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_t + \varepsilon_t - B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \dots + B_p \varepsilon_t$$

Figura 7 - ARIMA vs. base de treinamento



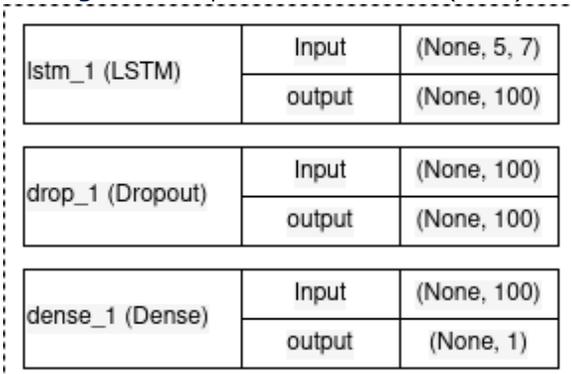
Fonte: Autoria própria.

### 3.4 MODELOS LSTM

Para a análise utilizando-se da LSTM foram desenvolvidos 3 modelos.

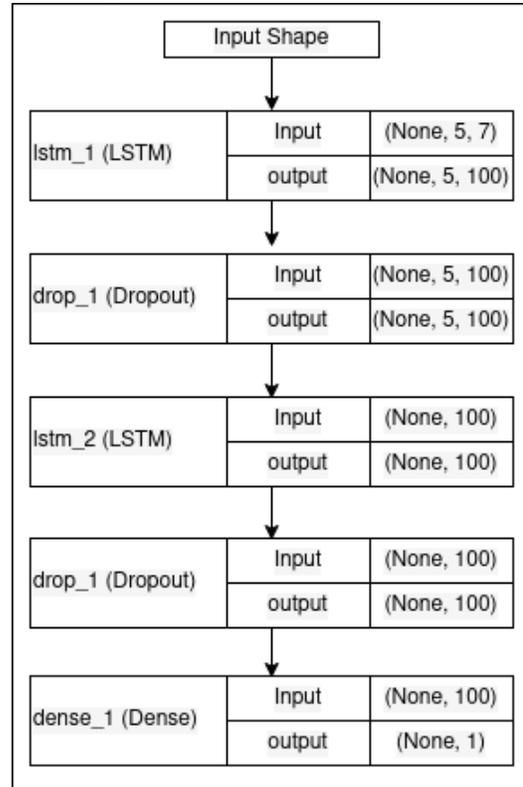
O primeiro modelo possui uma arquitetura com 3 camadas. A primeira camada oculta LSTM possui 100 neurônios de saída para a segunda camada de *dropout* (0.2), finalizando com uma camada *dense* com 1 neurônio com função de ativação linear. O otimizador *Adam* foi selecionado e utilizado para aprendizagem durante 120 épocas com lotes de tamanho 100, utilizando *loss function mse*.

Figura 8 - Arquitetura do modelo 1 (LSTM)



Fonte: Autoria própria.

Figura 9 - Arquitetura do modelo 2



Fonte: Autoria própria.

Por sua vez, quanto ao segundo modelo, esse possui uma arquitetura com 5 camadas, sendo 2 camadas LSTM de 100 neurônios separadas por camadas de *dropout* no valor 0.2 e a última camada *dense* com 1 neurônio com função de ativação linear. Foi selecionado como melhor otimizador de aprendizagem o *Adam* com valores padrão, executando por 400 épocas com lotes de tamanho 50.

Finalmente, o modelo 3 é possui uma arquitetura mais complexa e foi baseado no trabalho de Karakoyun e Cibikdiken [10], possuindo 3 camadas LSTM empilhadas totalmente conectadas. A primeira LSTM possui 64 neurônios, e a segunda e terceira camada possuem 128 camadas. Também possui 3 camadas *dense*, sendo a primeira com 64 neurônios, a segunda com 32 neurônios e a terceira com 1 neurônio. O otimizador utilizado foi *RMSprop* durante 600 épocas com lotes de tamanho 75.

### 3.5 OTIMIZAÇÃO COM HYPER-PARAMETER

Para cada modelo, foi executado uma grade de pesquisa utilizando a técnica de otimização de hiperparâmetros. Para tal, foi utilizada a classe *GridSearchSV* disponível no pacote *scikit-learn* da linguagem *Python*. As tarefas foram executadas em

uma *runtime Python 3 Google Compute Engine* do *Google Colab*.

O processo de execução da otimização foi feito individualmente para cada arquitetura de modelo LSTM, com diversas variações de parâmetros escolhidos arbitrariamente, como descrito a seguir:

### Modelo 1 (LSTM):

- epochs: [60, 90, 120, 200, 400, 500, 600]
- batch\_size: [50, 70, 75, 100, 128]
- cells (neurônios): [1, 2, 3, 5, 10, 20, 50, 70, 75, 100]
- optimizer: ['adam', 'RMSprop', 'adamax', 'nadam']
- parâmetros: *batch\_size: 100, cells: 100, epochs: 120, optimizer: 'RMSProp'*.

### Modelo 2 (LSTM):

- epochs: [60, 90, 120, 400, 600]
- batch\_size: [50, 75, 100, 128]
- cells (neurônios): [50, 75, 100]
- optimizer: ['adam', 'RMSprop', 'adamax', 'nadam']
- parâmetros: *batch\_size: 50, cells: 100, epochs: 400, optimizer: 'adam'*.

### Modelo 3 (LSTM):

- epochs: [60, 90, 120, 400, 500, 600, 800]
- batch\_size: [50, 75, 85, 100, 128]
- optimizer: ['adam', 'RMSprop', 'adamax', 'nadam']
- parâmetros encontrados foram: *batch\_size: 75, epochs: 600, optimizer: 'RMSprop'*.

Nesse modelo foi mantida a quantidade de neurônios nas camadas definida no estudo base dos autores. Assim sendo, não foi realizada grade de pesquisa de otimização para quantidade de neurônios.

## 4 RESULTADOS E CONCLUSÃO

**Tabela 2** - Preço Real vs. Previsto

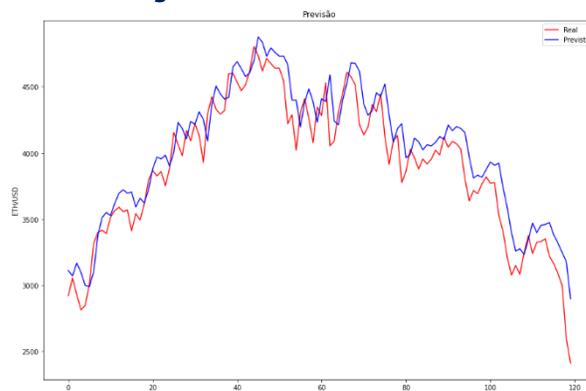
Modelo	Dif. Média	Desvio Padrão
ARIMA	4,575%	4,216%
LTSM1	4,325%	3,797%
LTSM2	15,464%	5,796%
LTSM3	5,675%	4,609%

**Fonte:** *Autoria própria.*

Após os testes dos quatro modelos descritos, verificou-se que o modelo com melhor desempenho foi o modelo de LSTM nº 1, uma vez que seus preços

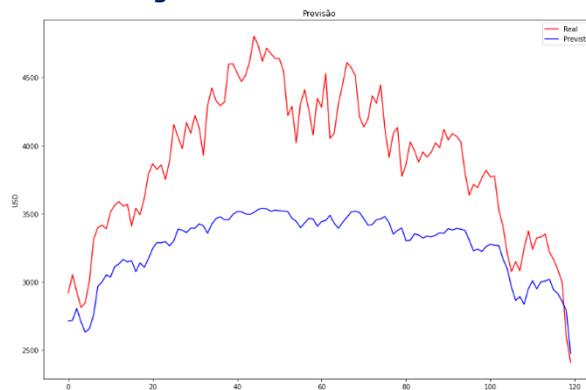
previstos (*outputs*) foram os que apresentaram menor diferença em relação aos preços reais, tanto no que se refere à média, quanto no que se refere à dispersão dos resultados (desvio padrão).

**Figura 10** - Previsão LTSM nº 1



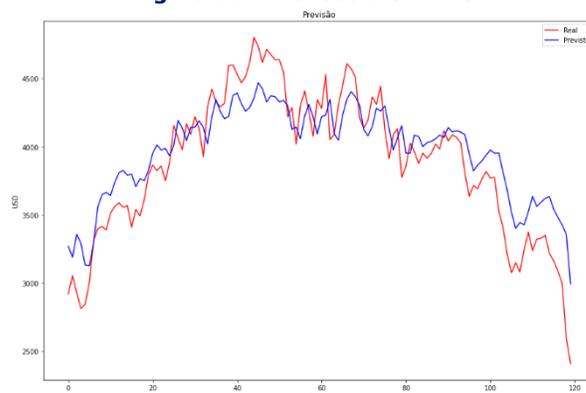
**Fonte:** *Autoria própria.*

**Figura 11** - Previsão LTSM nº 2



**Fonte:** *Autoria própria.*

**Figura 12** - Previsão LTSM nº 3



**Fonte:** *Autoria própria.*

Por sua vez, o modelo com pior desempenho foi o modelo de LSTM nº 2. Como resultado, verificou-se que os preços previstos por esse modelo são os que mais se distanciam dos valores reais, tanto no

que se refere à média, quanto no que se refere ao desvio padrão.

## REFERÊNCIAS

- [1] DAVENPORT, Thomas H. et al. **Working knowledge: How organizations manage what they know**. Harvard Business Press, 1998.
- [2] SIEBEL, Thomas M. **Digital transformation: survive and thrive in an era of mass extinction**. RosettaBooks, 2019.
- [3] SHRIMALI, Bela; PATEL, Hiren B. Blockchain state-of-the-art: architecture, use cases, consensus, challenges and opportunities. **Journal of King Saud University-Computer and Information Sciences**, 2021.
- [4] ESPOSITO, Anthony; RENTERIA, Nelson. El Salvador se torna o primeiro país a adotar bitcoins como moeda oficial. **CNN Brasil**, v. 7, 2021. Disponível em: <https://www.cnnbrasil.com.br/business/el-salvador-se-torna-o-primeiro-pais-a-adotar-bitcoins-como-moeda-oficial/> Acesso em 19/12/2021.
- [5] BUDIHARTO, Widodo. Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM). **Journal of big data**, v. 8, n. 1, p. 1-9, 2021.
- [6] TIMMERMANN, Allan; GRANGER, Clive WJ. Efficient market hypothesis and forecasting. **International Journal of forecasting**, v. 20, n. 1, p. 15-27, 2004.
- [7] SEBASTIÃO, Helder; GODINHO, Pedro. Forecasting and trading cryptocurrencies with machine learning under changing market conditions. **Financial Innovation**, v. 7, n. 1, p. 1-30, 2021.
- [8] COWPERTWAIT, Paul SP; METCALFE, Andrew V. **Introductory time series with R**. Springer Science & Business Media, 2009.
- [9] WIKIPEDIA CONTRIBUTORS. **Autocorrelation**. Disponível em: <<https://en.wikipedia.org/w/index.php?title=Autocorrelation&oldid=1090345069>>. Acesso em: 20 jun. 2022.
- [10] KARAKOYUN, Ebru Seyma; CIBIKDIKEN, A. O. Comparison of arima time series model and lstm deep learning algorithm for bitcoin price forecasting. In: **The 13th multidisciplinary academic conference in Prague**. 2018. p. 171-180.
- [11] GU, Shihao; KELLY, Bryan; XIU, Dacheng. Empirical asset pricing via machine learning. **The Review of Financial Studies**, v. 33, n. 5, p. 2223-2273, 2020.
- [12] BREMMER, Gerlof. **Predicting tomorrow's cryptocurrency price using a LSTM model, historical prices and Reddit comments**. 2018. Tese de Doutorado. Master thesis, Tilburg University, Tilburg, the Netherlands.
- [13] HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural computation**, v. 9, n. 8, p. 1735-1780, 1997.
- [14] SHAHID, Farah; ZAMEER, Aneela; MUNEEB, Muhammad. A novel genetic LSTM model for wind power forecast. **Energy**, v. 223, p. 120069, 2021., Disponível em: <https://doi.org/10.1016/j.energy.2021.120069>.