



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO, DE CIÊNCIAS EXATAS E EDUCAÇÃO  
DEPARTAMENTO DE ENG. DE CONTROLE, AUTOMAÇÃO E COMPUTAÇÃO  
CURSO DE GRADUAÇÃO EM ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Aryane de Paula Cezario

**Inteligência Artificial Para a Predição de Ações no Mercado  
Financeiro: Aplicação no Mercado de Bitcoin**

Blumenau  
2024

Aryane de Paula Cezario

**Inteligência Artificial Para a Predição de Ações no Mercado  
Financeiro: Aplicação no Mercado de Bitcoin**

Trabalho de Conclusão de Curso de Graduação em Engenharia de Controle e Automação do Centro Tecnológico, de Ciências Exatas e Educação da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Engenheiro de Controle e Automação.

Orientador: Prof. Dr. Carlos Roberto Moratelli

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.  
Dados inseridos pelo próprio autor.

Cezario, Aryane de Paula  
Inteligência Artificial Para a Predição de Ações no  
Mercado Financeiro : Aplicação no Mercado de Bitcoin /  
Aryane de Paula Cezario ; orientador, Carlos Roberto  
Moratelli, 2024.  
68 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Campus Blumenau,  
Graduação em Engenharia de Controle e Automação, Blumenau,  
2024.

Inclui referências.

1. Engenharia de Controle e Automação. 2. Inteligência  
Artificial. 3. Rede Neural. 4. LSTM. 5. Bitcoin. I.  
Moratelli, Carlos Roberto. II. Universidade Federal de  
Santa Catarina. Graduação em Engenharia de Controle e  
Automação. III. Título.

Aryane de Paula Cezario

**Inteligência Artificial Para a Predição de Ações no Mercado  
Financeiro: Aplicação no Mercado de Bitcoin**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Engenheiro de Controle e Automação” e aprovado em sua forma final pelo Curso de Graduação em Engenharia de Controle e Automação.

Blumenau, 10 de julho de 2024.

**Banca Examinadora:**

---

Prof. Carlos Roberto Moratelli, Dr.  
Universidade Federal de Santa Catarina

---

Prof. Mauri Ferrandin, Dr.  
Universidade Federal de Santa Catarina

---

Prof. Maiquel de Brito, Dr.  
Universidade Federal de Santa Catarina



## AGRADECIMENTOS

Agradeço à minha família e amigos, por sempre me apoiarem em todos os âmbitos da minha vida, por encorajarem minha criatividade e busca pelo conhecimento, além de proporcionarem todas as ferramentas necessárias para a minha formação acadêmica.

Aos meus amigos e colegas de classe, agradeço pelo apoio e por todos os ensinamentos que foram fundamentais para o meu desenvolvimento dentro e fora da universidade.

Aos meus professores, agradeço por todos os ensinamentos, pela paciência e incentivo. Uma menção especial ao meu orientador Carlos Moratelli, que me auxiliou no desenvolvimento deste trabalho.

Agradeço finalmente ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) que me apoiou no desenvolvimento dessa pesquisa através de uma bolsa de pesquisa, possibilitando que eu pudesse voltar a minha atenção para essa pesquisa em momentos de desafios financeiros.

In your life, you will inevitably misspeak, trust the wrong people, under-react, overreact, hurt the people who didn't deserve it, overthink, not think at all, self sabotage, create a reality where only your experience exists, ruin perfectly good moments for yourself and others, deny any wrongdoing, not take the steps to make it right, feel very guilty, let the guilt eat at you, hit rock bottom, finally address the pain you caused, try to do better next time, rinse, repeat. And I'm not gonna lie, these mistakes will cause you to lose things. I'm trying to tell you that losing things doesn't just mean losing. A lot of the time, when we lose things, we gain things too. (Taylor Alison Swift, 2022)

## RESUMO

As redes neurais LSTM (*Long Short-Term Memory*) são amplamente utilizadas para a predição de valores de ações no mercado financeiro devido à sua capacidade de capturar dependências temporais de longo prazo e lidar com dados sequenciais complexos. Utilizando dados históricos de preços de ações, volumes de negociação e indicadores técnicos, as LSTM's conseguem identificar padrões e tendências que ajudam a prever movimentos futuros dos preços. Sua arquitetura, que inclui unidades de memória e portas de controle, permite filtrar ruídos e focar em informações relevantes, tornando-as robustas para análises financeiras. Apesar da volatilidade e dos desafios inerentes ao mercado financeiro, as LSTM's têm se mostrado eficazes em melhorar a precisão das previsões, auxiliando investidores na tomada de decisões informadas. Assim, esse trabalho apresenta o desenvolvimento de um modelo de LSTM treinado para realizar a predição dos valores das ações da criptomoeda Bitcoin no mercado financeiro, bem como a realização de experimentos para a análise de desempenho de diferentes cenários alterando parâmetros como tipo de função de perda, quantidade de camadas empilhadas, tamanho da sequência de dados de entrada e período de tempo a ser previsto. O projeto resultou em 24 cenários estudados cujo desempenho foi medido através da avaliação da convergência das funções de perda e das métricas de avaliação de regressão  $R^2$  (Coeficiente de Determinação) e MSE (*Mean Squared Error*). Dentre os cenários criados se destacou a utilização da função de perda MAE (*Mean Absolute Error*) (L1) no modelo de 100 camadas e sequência de dados igual à 4 para a predição do valor de fechamento das ações da Bitcoin para as próximas 6 horas.

**Palavras-chave:** Bitcoin; IA; LSTM; Rede neural; Aprendizado de máquina

## ABSTRACT

LSTM (*Long Short-Term Memory*) neural networks are widely used for predicting stock values in the financial market due to their ability to capture long-term temporal dependencies and handle complex sequential data. Using historical data on stock prices, trading volumes and technical indicators, LSTM's can identify patterns and trends that help predict future price movements. It's architecture, which includes memory units and control ports, allows it to filter noise and focus on relevant information, making it robust for financial analysis. Despite the volatility and challenges inherent to the financial market, LSTM's have proven effective in improving the accuracy of forecasts, helping investors make informed decisions. Thus, this work presents the development of an LSTM model trained to predict the values of Bitcoin cryptocurrency shares in the financial market, as well as carrying out experiments to analyze the performance of different scenarios by changing parameters such as the type of loss function, number of stacked layers, size of the input data sequence and period of time to be predicted. The project resulted in 24 studied scenarios whose performance was measured by evaluating the convergence of loss functions and the  $R^2$  (*Coefficient of Determination*) and MSE (*Mean Squared Error*) regression evaluation metrics. Among the scenarios created, the use of the MAE (*Mean Absolute Error*) (L1) loss function in the 100-layer model and data sequence equal to 4 stood out for predicting the closing value of Bitcoin shares for the next 6 hours.

**Keywords:** Bitcoin; IA; LSTM; Neural Network; Machine learning

## LISTA DE FIGURAS

Figura 1 – Participação percentual por instrumento. . . . .	17
Figura 2 – Bitcoin blockchain simplificada. . . . .	19
Figura 3 – Ciclos de mercado da Bitcoin. . . . .	24
Figura 4 – Arquitetura Genérica de redes neurais. . . . .	30
Figura 5 – Classificação de redes neurais. . . . .	31
Figura 6 – Diagrama do fluxo de dados em uma LSTM. . . . .	36
Figura 7 – Diagrama de fluxo do código desenvolvido . . . . .	37
Figura 8 – <i>Candlesticks</i> obtidos na plataforma Binance. . . . .	38
Figura 9 – Casos estudados. . . . .	40
Figura 10 – MAE - Treinamento para 6 horas de previsão . . . . .	42
Figura 11 – MAE - Treinamento para 12 horas de previsão . . . . .	43
Figura 12 – MAE - Treinamento para 24 horas de previsão . . . . .	44
Figura 13 – Convergência da função MAE . . . . .	45
Figura 14 – MSE - Treinamento para 6 horas de previsão . . . . .	46
Figura 15 – MSE - Treinamento para 12 horas de previsão . . . . .	47
Figura 16 – MSE - Treinamento para 24 horas de previsão . . . . .	49
Figura 17 – Convergência da função MSE . . . . .	50
Figura 18 – Huber - Treinamento para 6 horas de previsão . . . . .	51
Figura 19 – Huber - Treinamento para 12 horas de previsão . . . . .	52
Figura 20 – Huber - Treinamento para 24 horas de previsão . . . . .	53
Figura 21 – Convergência da função Huber . . . . .	54
Figura 22 – Smooth L1 - Treinamento para 6 horas de previsão . . . . .	55
Figura 23 – Smooth L1 - Treinamento para 12 horas de previsão . . . . .	56
Figura 24 – Smooth L1 - Treinamento para 24 horas de previsão . . . . .	57
Figura 25 – Convergência da função Smooth L1 . . . . .	58
Figura 26 – Valores previstos pelo modelo para as próximas 6 horas. . . . .	60
Figura 27 – Valores previstos pelo modelo para as próximas 12 horas. . . . .	61
Figura 28 – Valores previstos pelo modelo para as próximas 24 horas. . . . .	62

## LISTA DE QUADROS

Quadro 1 – Indicadores de Mercado Aplicados. . . . .	27
Quadro 2 – Hiperparâmetros Fixos. . . . .	38
Quadro 3 – Cenários criados para análise . . . . .	41

## LISTA DE TABELAS

Tabela 1 – Informações de entrada para o treinamento da rede . . . . .	39
Tabela 2 – Cenários testados . . . . .	59

## LISTA DE ABREVIATURAS E SIGLAS

$R^2$	Coefficiente de Determinação
100L	Número de Camadas Igual à 100
10S	Sequência de Dados de Entrada Igual à 10
4S	Sequência de Dados de Entrada Igual à 4
65L	Número de Camadas Igual à 65
AAII	<i>American Association of Individual Investors</i>
AD	Acumulação/Distribuição
Adam	<i>Adaptive Moment Estimation</i>
ADX	Índice de Direção Média
AES	<i>Adaptive Exponential Smoothing</i>
ANN	<i>Artificial Neural Network</i>
API	Interface de Programação de Programação
APO	Oscilador de Preço Absoluto
ARIMA	<i>Autoregressive Integrated Moving Average</i>
ATR	Média de Amplitude de Variação
BBANDS	Bandas de Bollinger
COMP	<i>Nasdaq Composite</i>
DeFi	Finanças Descentralizadas
DEMA100	Média Móvel Exponencial Dupla 100 Dias
DEMA20	Média Móvel Exponencial Dupla 20 Dias
DEMA200	Média Móvel Exponencial Dupla 200 Dias
DEMA50	Média Móvel Exponencial Dupla 50 Dias
DEX	Exchanges Descentralizadas
DJIA	<i>Dow Jones Industrial Average</i>
EMA100	Média Móvel Exponencial 100 Dias
EMA20	Média Móvel Exponencial 20 Dias
EMA200	Média Móvel Exponencial 200 Dias
EMA50	Média Móvel Exponencial 50 Dias
EPS	Lucros Por Ação ( <i>Earnings Per Share</i> )
FAIR	<i>Fundamental AI Research</i>
GNB	<i>Gaussian Naive Bayes</i>
GPU	Unidade de Processamento Gráfico
IA	Inteligência Artificial
IOTA	<i>Internet of Things Application</i>
LSTM	<i>Long Short-Term Memory</i>
MACD	Média Móvel Convergente e Divergente
MACDHIST	Dados Para Histograma da Média Móvel Convergente e Divergente



MACDSIGNAL	Linha de Sinal da Média Móvel Convergente e Divergente
MAE	<i>Mean Absolute Error</i>
MFI	Índice Fluxo de Dinheiro
MOM	Momento
MSE	<i>Mean Squared Error</i>
NATR	Média de Amplitude de Variação Normalizada
NLP	Processamento de Linguagem Natural
NumPy	<i>Numerical Python</i>
OBV	Volume em Balanço
P/L	Índice Preço/Lucro
PIB	Produto Interno Bruto
PoW	<i>Proof of Work</i>
RAM	<i>Random Access Memory</i>
RNN	<i>Recurrent Neural Network</i>
ROC	Taxa de Variação
RSI	Índice de Força Relativa
S&P500	<i>Standard &amp; Poor's 500</i>
SciPy	<i>Scientific Python</i>
Selic	Sistema Especial de Liquidação e de Custódia
SMA100	Média Móvel Simples 100 Dias
SMA20	Média Móvel Simples 20 Dias
SMA200	Média Móvel Simples 200 Dias
SMA50	Média Móvel Simples 50 Dias
STDDEV	Desvio Padrão
Stoch	Oscilador Estocástico
Stoch RSI	Índice de Força Relativa Estocástico
SVM	<i>Support Vector Machine</i>
TSA	<i>Time Stamp Authority</i>
TSS	<i>Time Stamp Server</i>
VIX	<i>Volatility Index</i>
WILLR	Indicador Williams %R
WMA100	Média Móvel Ponderada 100 Dias
WMA20	Média Móvel Ponderada 20 Dias
WMA200	Média Móvel Ponderada 200 Dias
WMA50	Média Móvel Ponderada 50 Dias

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	OBJETIVO GERAL	15
1.1.1	<b>Objetivos Específicos</b>	<b>15</b>
<b>2</b>	<b>CRIPTOMOEDAS</b>	<b>17</b>
2.1	BLOCKCHAIN	18
2.2	FINANÇAS DESCENTRALIZADAS	22
2.3	CICLOS DE MERCADO	23
2.4	INDICADORES	24
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>29</b>
3.1	REDE NEURAL PROFUNDA	29
3.2	PYTHON E FERRAMENTAS	31
3.3	TRABALHOS RELACIONADOS	33
<b>4</b>	<b>PROPOSTA E IMPLEMENTAÇÃO</b>	<b>35</b>
4.1	IMPLEMENTAÇÃO PARA A PREDIÇÃO DE VALORES	36
<b>5</b>	<b>RESULTADOS</b>	<b>40</b>
5.1	ESTUDOS REALIZADOS	40
5.2	FUNÇÃO DE PERDA MAE	40
5.2.1	<b>MAE: 6 Horas de Previsão</b>	<b>41</b>
5.2.2	<b>MAE: 12 Horas de Previsão</b>	<b>42</b>
5.2.3	<b>MAE: 24 Horas de Previsão</b>	<b>43</b>
5.2.4	<b>MAE: Convergência da Função de Perda</b>	<b>45</b>
5.3	FUNÇÃO DE PERDA <i>MEAN SQUARED ERROR</i> (MSE)	45
5.3.1	<b>MSE: 6 Horas de Previsão</b>	<b>46</b>
5.3.2	<b>MSE: 12 Horas de Previsão</b>	<b>47</b>
5.3.3	<b>MSE: 24 Horas de Previsão</b>	<b>48</b>
5.3.4	<b>MSE: Convergência da Função de Perda</b>	<b>48</b>
5.4	FUNÇÃO DE PERDA HUBER	48
5.4.1	<b>Huber: 6 Horas de Previsão</b>	<b>48</b>
5.4.2	<b>Huber: 12 Horas de Previsão</b>	<b>50</b>
5.4.3	<b>Huber: 24 Horas de Previsão</b>	<b>50</b>
5.4.4	<b>Huber: Convergência da Função de Perda</b>	<b>50</b>
5.5	FUNÇÃO DE PERDA SMOOTH L1	51
5.5.1	<b>Smooth L1: 6 Horas de Previsão</b>	<b>52</b>
5.5.2	<b>Smooth L1: 12 Horas de Previsão</b>	<b>54</b>
5.5.3	<b>Smooth L1: 24 Horas de Previsão</b>	<b>54</b>
5.5.4	<b>Smooth L1: Convergência da Função de Perda</b>	<b>55</b>
5.6	VALORES PREVISTOS PELO MODELO E VALIDAÇÃO	55

6	CONCLUSÃO . . . . .	63
	REFERÊNCIAS . . . . .	65

# 1 INTRODUÇÃO

A predição de ativos no mercado financeiro apresenta inúmeros desafios devido à natureza altamente volátil e não linear. Um dos principais desafios é a complexidade dos fatores que influenciam os preços desses ativos. Esses fatores incluem notícias econômicas, políticas monetárias, eventos geopolíticos, resultados financeiros das empresas, entre outros. A interação entre esses elementos cria um ambiente dinâmico onde os preços dos ativos podem mudar rapidamente, tornando a previsão precisa extremamente difícil, e o aumento do seu comportamento caótico contribui para o aumento da dificuldade do problema de predição (ESFAHANIPOUR; AGHAMIRI, 2010 apud CHOPRA; SHARMA, 2021).

Outro desafio significativo é o comportamento humano e a psicologia dos investidores. Os mercados financeiros são fortemente influenciados pelas decisões dos investidores, que muitas vezes são guiadas por emoções como medo e ganância, além de vieses cognitivos e comportamento de manada. Esses fatores comportamentais podem levar a flutuações de mercado que são difíceis de modelar e prever com precisão, mesmo com ferramentas analíticas avançadas. Além disso, a disseminação de informações falsas ou rumores pode causar movimentos abruptos e inesperados nos preços dos ativos.

A IA (Inteligência Artificial) tem emergido como uma ferramenta poderosa para a predição de ativos no mercado financeiro, transformando a maneira como analistas e investidores abordam a tomada de decisões. Utilizando algoritmos de aprendizado de máquina e redes neurais artificiais, a IA pode analisar grandes volumes de dados em tempo real, identificando padrões e tendências que seriam praticamente impossíveis de detectar manualmente. Essa capacidade de processamento e análise de dados em alta velocidade permite que modelos preditivos baseados em IA façam previsões mais precisas e oportunas sobre os movimentos do mercado.

## 1.1 OBJETIVO GERAL

O objetivo deste trabalho é desenvolver e implementar um modelo IA (Inteligência Artificial) treinado para realizar a predição do valor da Bitcoin no mercado financeiro. Além da elaboração do modelo baseado na arquitetura de LSTM (*Long Short-Term Memory*), esse projeto também visa a avaliação de diferentes cenários da mesma rede, com variações de quantidade de camadas, quantidade de dados na sequência de entrada e função de perda.

### 1.1.1 Objetivos Específicos

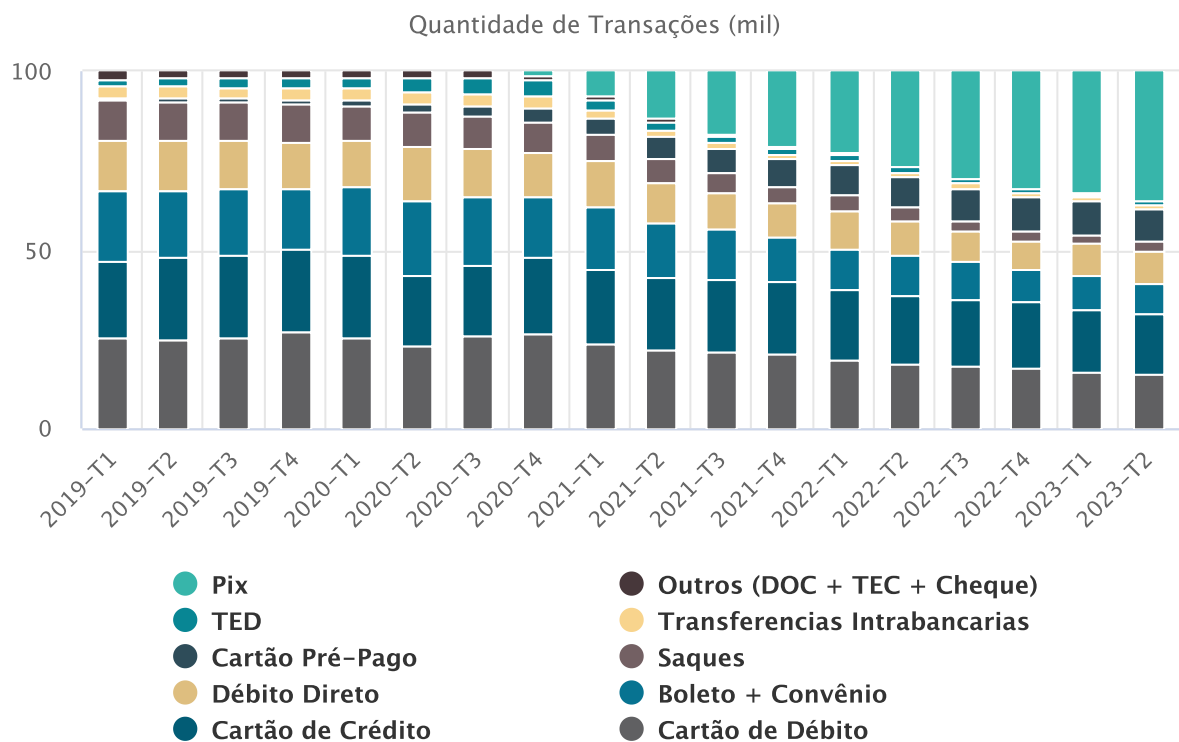
- Aplicar os conceitos de programação e *big data* para a predição de criptoativos no mercado financeiro;
- Definir e implementar o modelo da rede neural;

- Realizar o ajuste dos hiperparâmetros da rede para a obtenção de melhores resultados;
- Comparar diferentes cenários de treinamento.

## 2 CRIPTOMOEDAS

Ao longo dos anos, o avanço tecnológico tem transformado como a sociedade gerencia seus recursos e meios indispensáveis, acompanhando as constantes mudanças no campo tecnológico. Essa evolução reflete diretamente na maneira como as pessoas lidam com finanças: até o início da década de 1950, a maioria das transações eram realizadas utilizando cédulas, principalmente devido à ausência de alternativas. Em 1954, o primeiro cartão de crédito, o *Diner's Club*, foi introduzido no Brasil, marcando o início da popularização desse meio de pagamento nas décadas subsequentes (MORAES, 2008). Em novembro de 2020, o Banco Central brasileiro lançou o Pix, uma ferramenta de transferência simplificada, operando 24 horas por dia, sete dias por semana, e isenta de taxas.

Figura 1 – Participação percentual por instrumento.



Fonte: (BRASIL, 2022).

Conforme os dados apresentados pelo Banco Central do Brasil (BRASIL, 2022), as transações em espécie, representadas por saques, vêm perdendo terreno para as operações digitais, como Pix, cartões de crédito e débito. Conforme evidenciado na Figura 1, no primeiro trimestre de 2019, os saques em cédulas correspondiam a aproximadamente 11% do total de transações no país. Nesse mesmo período, os pagamentos realizados

por meio de cartão de crédito e débito representavam, respectivamente, cerca de 21% e 25%. No segundo trimestre de 2023, porém, as transações em espécie reduziram-se significativamente, alcançando apenas 3% do total, enquanto as operações com cartão de crédito, débito e Pix atingiram, respectivamente, 17%, 15% e 36% (BRASIL, 2022).

A previsão é de que o avanço das novas tecnologias estimule o surgimento de métodos inovadores para a condução de transações, promovendo uma evolução no sistema financeiro, com uma crescente tendência para a digitalização. Um marco inicial nesse processo foi a introdução das criptomoedas: um sistema virtual de cunhagem que opera semelhantemente à moeda convencional, permitindo que os usuários realizem pagamentos virtuais por bens e serviços, sem depender de uma autoridade central (FARELL, 2015).

Os fundamentos das criptomoedas contemporâneas têm suas origens nos primeiros estudos de especialistas em criptografia, como evidenciado em trabalhos como “*Blind Signatures for Untraceable Payments*” (CHAUM, 1983), que introduziu o *eCash*, e “*b-Money*” (DAI, 1998), onde foi proposta a moeda homônima.

Embora os primeiros trabalhos sobre o assunto tenham sido publicados na década de 1980, a primeira criptomoeda bem-sucedida só surgiu em 2008, com a publicação do artigo “*Bitcoin: A Peer-to-Peer Electronic Cash System*” (NAKAMOTO, 2008), atribuído a uma pessoa ou grupo sob o pseudônimo Satoshi Nakamoto. Ao longo dos anos, uma série de outras criptomoedas emergiram no cenário do dinheiro eletrônico, incluindo Litecoin, Solana, Tether, entre outras.

Apesar de suas distinções, as criptomoedas compartilham a essência de serem meios de troca digitais que empregam criptografia ponto-a-ponto, permitindo transações seguras diretamente entre as partes envolvidas, sem a necessidade de intermediação de uma instituição financeira. Essa funcionalidade é viabilizada pela tecnologia digital conhecida como *blockchain* (ZHU, 2021).

## 2.1 BLOCKCHAIN

Uma das principais contribuições de Nakamoto e grande diferencial que auxiliou no sucesso da Bitcoin foi a apresentação da *blockchain*, um método de estabelecer uma sequência cronológica de transações que todos os usuários concordam e sem depender de uma autoridade central. Tabatabaei, Vitenberg e Veeraragavan (2023) identificaram cinco principais definições para *blockchain* na literatura:

[...] Definição 1: *Blockchain* é um sistema baseado na estrutura de dados da Bitcoin, ampliando significativamente sua funcionalidade. Essa concepção é adotada por diversas criptomoedas derivadas da Bitcoin, resultantes de operações de *fork* ou como uma extensão da capacidade limitada de *script* da Bitcoin. Vale ressaltar que essa definição transcende a esfera das criptomoedas, podendo ser aplicada a uma ampla gama de processos lógicos de negócios, por meio da personalização dos módulos e protocolos da *blockchain*.

Definição 2: *Blockchain* é um sistema que mantém uma cadeia de blocos. Essa definição permite uma generalização da Definição 1: permite estruturas de

dados diferentes daquelas utilizadas pela Bitcoin. Ethereum e Hyperledger, por exemplo, condizem com essa definição.

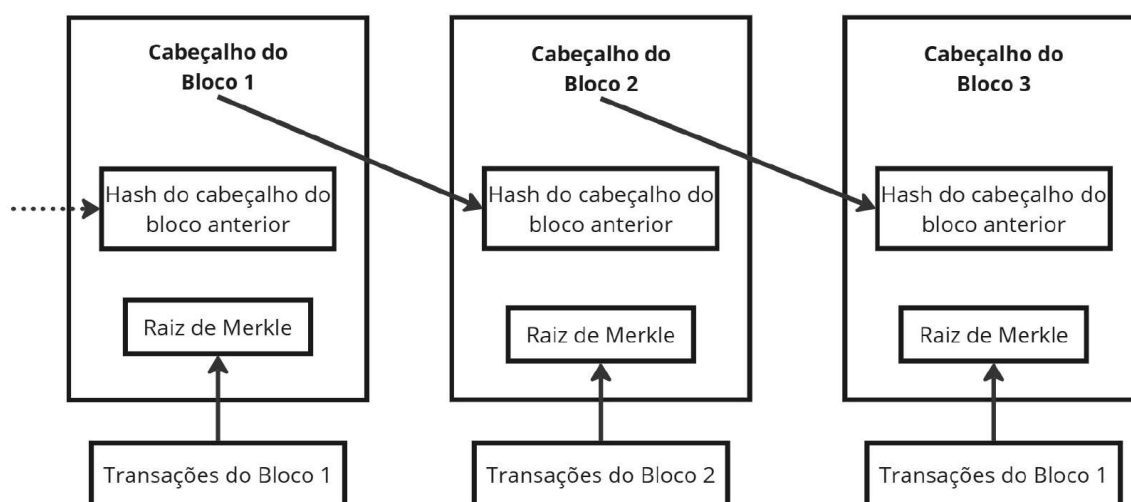
Definição 3: Blockchain é um sistema que mantém um livro-razão de todas as transações. No entanto, esse livro-razão não precisa ser armazenado como uma cadeia de blocos. Um exemplo que adere a essa definição é o sistema IOTA (*Internet of Things Application*).

Definição 4: Blockchain é um sistema em que partes não confiáveis colaboram de maneira distribuída, sem a necessidade de um intermediário confiável. Essa definição destaca a principal propriedade benéfica desse paradigma, sendo inicialmente defendida por Corda (HEARN; BROWN, 2016).

Definição 5: Blockchain é uma plataforma que oferece suporte para contratos inteligentes. Muitos blogs e artigos científicos populares sugerem que o blockchain tem o potencial de substituir contratos em papel e intermediários humanos por contratos inteligentes, embora não abordem completamente a implementação desses contratos. [...] (TABATABAEI; VITENBERG; VEERARAGAVAN, 2023)

Com base nas definições acima, é viável simplificar a descrição da blockchain como uma série de blocos interligados por meio de uma função *hash* criptográfica. Para evitar problemas como gastos duplicados, a blockchain utiliza os registros de data e hora de cada transação como entrada para essa função. Cada bloco na blockchain contém informações cruciais para sua identificação e para assegurar a integridade da rede. Essas informações incluem o *hash* do cabeçalho do bloco anterior para garantir a conexão sequencial na cadeia, um *hash* da raiz da árvore Merkle das transações do bloco para garantir integridade, uma lista de transações desde o último bloco, assinaturas digitais de todos os envolvidos para autenticidade e identificação, e a prova de trabalho. Uma representação simplificada dessa estrutura é apresentada na Figura 2.

Figura 2 – Bitcoin blockchain simplificada.



Fonte: (BITCOIN.ORG, 2023) (adaptado).



A prova de trabalho (PoW (*Proof of Work*)) é um conceito fundamental utilizado para assegurar a segurança e a integridade da rede. Conforme descrito por Nakamoto em seu artigo seminal sobre Bitcoin (NAKAMOTO, 2008), a prova de trabalho implica na busca por um valor cujo *hash*, calculado utilizando a função criptográfica SHA-256, comece com um número específico de bits zero. Nesse contexto, os participantes da rede, também conhecidos como mineiros, competem entre si para encontrar a solução de um problema criptográfico complexo relacionado ao bloco que estão tentando adicionar à blockchain. Esse processo é denominado mineração. A rede é projetada para ajustar dinamicamente a dificuldade do problema com base na taxa global de mineração, garantindo um tempo médio para encontrar uma solução em um intervalo predefinido de tempo, como cerca de 10 minutos no caso do Bitcoin. A solução encontrada pode ser prontamente verificada pelos demais participantes da rede, inserindo o valor descoberto na função *hash* e verificando se o número de bits zero no início corresponde ao requisito especificado pela rede. O mineiro que primeiro resolve o problema é recompensado com uma quantidade específica de criptomoeda.

Um dos problemas enfrentado na concepção de uma nova moeda é garantir que o usuário não possa realizar transações que constituem gasto duplo, ou seja, desembolsar uma quantia maior do que o saldo que possui. É possível evitar esse tipo de situação implementando um sistema onde a cadeia de posse seja verificada, garantindo que o valor de saída da conta de cada usuário não seja maior que o valor de entrada. A solução comumente utilizada para moedas fiduciárias é a definição de uma autoridade central que verifica todas as transações, como instituições financeiras bancárias. Entretanto, devido ao objetivo de assegurar uma moeda descentralizada e transparente, a solução proposta por Nakamoto (2008) para a Bitcoin foi a implementação de um servidor com criptografia baseada em carimbo de tempo (TSS (*Time Stamp Server*)).

Esse padrão de criptografia é utilizado para certificar que o documento foi criado ou modificado em determinada data e horário. Segundo Ansper *et al.* (2001), *time-stamping* é um serviço utilizado pela *Parte Confiável* para provar a relação temporal ao *Verificador*, por exemplo, um juiz. A parte confiável obtém os *timestamps* da TSA (*Time Stamp Authority*) utilizando um protocolo de estampagem. O verificador utiliza um protocolo de verificação para averiguar a legitimidade dos *timestamps* apresentado pela parte confiável. Assim, sempre que um usuário desejar realizar uma nova transação, antes que a mesma seja autorizada o sistema é capaz de verificar o saldo detido na conta através da relação entre o saldo anterior e a transação registrada no último tempo carimbado, evitando o problema de gasto duplo. No caso da Bitcoin, as informações de data e horário de cada transação são registradas no *hash* de cada bloco, onde também é registrado o *timestamp* da transação anterior, criando uma cadeia que onde cada item corrobora a autenticidade do registro anterior (NARAYANAN *et al.*, 2016).

Uma vez que o trabalho é concluído e a prova de trabalho é satisfeita, um bloco

não pode ser modificado sem a necessidade de refazer o trabalho. Devido à natureza de encadeamento sequencial dos blocos, qualquer tentativa de alterar um bloco exigiria a reexecução do trabalho associado à sua criação, bem como de todos os blocos encadeados subsequentemente. Dado o caráter dinâmico, transparente e colaborativo do sistema, a probabilidade de ocorrer um ataque fraudulento na rede diminui exponencialmente à medida que o número de blocos encadeados após o primeiro bloco atacado aumenta. Isso é evidenciado na publicação de Nakamoto (2008) ao ser considerado um cenário onde um agressor tenta gerar uma cadeia de blocos mais veloz que a cadeia honesta. Segundo o autor, a probabilidade da cadeia agressora recuperar a diferença de tamanho em relação à cadeia honesta pode ser interpretada de maneira similar ao problema da Ruína do Jogador (COOLIDGE, 1909) e pode ser calculada pela seguinte expressão:

$$q_z = \begin{cases} 1 & \text{se } p \leq q \\ (q/p)^z & \text{se } p > q \end{cases}$$

Onde:

$p$  = probabilidade de um nó honesto encontrar o próximo bloco

$q$  = probabilidade de um agressor encontrar o próximo bloco

$q_z$  = probabilidade de um agressor alcançar estando  $z$  blocos atrás

Com base nas análises realizadas, o autor deduz que a probabilidade de um ataque bem-sucedido à blockchain diminui exponencialmente com  $z$ , comprovando a robustez da cadeia.

Dessa forma, o funcionamento da rede ocorre da seguinte maneira: as novas transações são divulgadas para todos os nós, e cada nó reúne essas transações em um bloco. Em seguida, cada nó se empenha em encontrar a prova de trabalho correspondente ao seu bloco, e ao concluir essa tarefa, transmite o bloco para todos os outros participantes da rede. Se todas as transações registradas no bloco forem válidas e o valor não tiver sido gasto anteriormente, os demais participantes da rede aceitam o bloco e validam-no, trabalhando na criação do próximo bloco e utilizando o *hash* do bloco aceito como o *hash* anterior.

A essência da blockchain pode ser resumida em três conceitos fundamentais: descentralização, transparência e imutabilidade, conceitos estes abordados em Bashir (2018). A descentralização é assegurada pelo compartilhamento de todos os dados das transações na rede e pelo armazenamento distribuído entre os nós participantes, garantindo a segurança dos dados mesmo se um participante perder informações ou deixar a rede. Além disso, esse modelo assegura que nenhum indivíduo, organização, governo, autoridade ou entidade tenha controle sobre os processos de gravação e armazenamento dos dados. A transparência reflete-se na disponibilidade de todos os dados em um registro público acessível a qualquer interessado. Isso praticamente inviabiliza tentativas de fraude, pois a alteração dos dados

de uma única transação exigiria a modificação de todos os blocos subsequentes em todos os dispositivos conectados à rede globalmente. Por fim, a imutabilidade é assegurada pela aplicação do encadeamento criptografado baseado em *timestamp* e pela utilização da estrutura de árvore de Merkle (LIU, H. *et al.*, 2021). Esses elementos foram concebidos para dificultar ao máximo a alteração de um bloco existente, fortalecendo assim a segurança e a confiança na rede blockchain.

## 2.2 FINANÇAS DESCENTRALIZADAS

As DeFi (Finanças Descentralizadas) representam uma revolução no mundo das finanças. Ao contrário do sistema financeiro tradicional, que depende fortemente de instituições centralizadas como bancos e governos, o DeFi busca criar um ecossistema financeiro completamente descentralizado e baseado em blockchain. Esse tipo de finança refere-se a um sistema financeiro que opera independentemente de intermediários centralizados (BIRRER; AMSTUTZ; WENGER, 2023). Em vez disso, ele utiliza contratos inteligentes e protocolos descentralizados executados em uma blockchain para facilitar transações financeiras, empréstimos, negociações e muito mais. Essencialmente, o DeFi busca eliminar a necessidade de confiança em terceiros, como bancos ou corretoras, ao mesmo tempo em que aumenta a transparência e a acessibilidade financeira.

De acordo com Harvey *et al.* (2021), alguns dos princípios nos quais as DeFi se baseiam são a descentralização, o acesso universal, a transparência e a interoperabilidade. A descentralização consiste na premissa de que os protocolos DeFi são executados em blockchains públicas ou permissionadas, significando não haver uma única entidade controladora. Todas as transações e operações são verificáveis e transparentes para qualquer pessoa com acesso à blockchain. Com o acesso universal, a DeFi visa democratizar o acesso aos serviços financeiros, permitindo que qualquer pessoa com uma conexão à internet participe do ecossistema financeiro global. Isso é especialmente significativo em regiões onde o acesso a serviços financeiros tradicionais é limitado. A transparência se deve ao fato de que todas as transações e contratos inteligentes no DeFi são públicos e imutáveis, o que significa que qualquer pessoa pode auditar e verificar as operações realizadas na blockchain. Isso aumenta a confiança dos usuários no sistema financeiro descentralizado. Além disso, os protocolos DeFi são projetados para serem interoperáveis, o que significa que diferentes aplicativos e plataformas podem se integrar e colaborar entre si. Isso cria um ecossistema financeiro mais robusto e flexível, onde os usuários podem acessar uma variedade de serviços sem sair de uma única plataforma.

Existem inúmeras aplicações para DeFi, sendo algumas delas empréstimos e empréstimos colateralizado, negociação descentralizada, *staking* e *yield farming* e seguros descentralizados. Plataformas DeFi permitem que os usuários emprestem e tomem empréstimos de forma descentralizada, muitas vezes usando criptomoedas como garantia. As DEX (Exchanges Descentralizadas) permitem que os usuários negociem criptomoedas e

tokens sem a necessidade de uma corretora centralizada. Além disso, os usuários podem ganhar recompensas por manterem e apostarem certas criptomoedas em protocolos DeFi, um processo conhecido como *staking*. O *yield farming* envolve a participação em vários protocolos DeFi para maximizar os retornos sobre o investimento. Por fim, os protocolos DeFi oferecem soluções de seguro descentralizadas, onde os usuários podem proteger seus ativos contra perdas e hacks.

Embora as finanças descentralizadas ofereçam muitos benefícios, também enfrentam desafios significativos, incluindo preocupações com segurança, escalabilidade e regulamentação. No entanto, o potencial transformador do DeFi é inegável, e muitos acreditam que ele continuará a crescer e a evoluir, desempenhando um papel cada vez mais importante no futuro das finanças globais.

As finanças descentralizadas representam uma nova era de inovação financeira, onde a confiança é colocada em protocolos transparentes e descentralizados em vez de instituições centralizadas. Embora ainda esteja em seus estágios iniciais, o DeFi promete democratizar o acesso aos serviços financeiros e transformar fundamentalmente a maneira como pensamos sobre o dinheiro e as transações financeiras.

### 2.3 CICLOS DE MERCADO

Os ciclos de mercado são padrões recorrentes de expansão e contração na atividade econômica, refletindo variações na produção, emprego, investimento e outros indicadores-chave. Esses ciclos são uma característica intrínseca das economias de mercado sendo observados ao longo da história econômica.

Assim, o mercado segue um padrão previsível, passando por diferentes fases. Durante a expansão, há um crescimento econômico robusto, com aumento na produção e no emprego, além de um sentimento positivo nos mercados financeiros. No pico, atinge-se o ápice do ciclo, com a demanda começando a diminuir e os custos aumentando. A fase de contração, também conhecida como recessão, é marcada por uma desaceleração econômica, queda na produção e no emprego, afetando a confiança dos consumidores e investidores. Por fim, no vale, atinge-se o ponto mais baixo do ciclo, com uma recuperação gradual da confiança e a possibilidade de uma melhora nas condições econômicas.

Vários fatores exercem influência nos ciclos de mercado. A política monetária, estabelecida pelos bancos centrais, como taxas de juros e oferta de dinheiro, desempenha um papel significativo, afetando o acesso ao crédito e o custo do capital. Paralelamente, as decisões governamentais sobre gastos, impostos e políticas fiscais, conhecidas como política fiscal, também têm impacto na determinação dos ciclos de mercado, influenciando a demanda agregada e a atividade econômica. Além disso, choques externos, como crises financeiras, desastres naturais ou eventos geopolíticos imprevistos, podem desencadear ou agravar os ciclos de mercado, provocando perturbações na economia global. Por fim, as expectativas dos agentes econômicos, incluindo consumidores, empresas e investidores, em

relação ao futuro da economia, também desempenham um papel crucial, afetando suas decisões de gastos, investimentos e poupança.

Embora os ciclos de mercado possam apresentar desafios significativos para indivíduos e empresas, também oferecem oportunidades para aqueles que sabem como identificar e aproveitar as tendências econômicas. Estratégias de diversificação de investimentos, gestão de liquidez e análise cuidadosa dos fundamentos econômicos podem ajudar a mitigar os riscos associados aos ciclos de mercado e a aproveitar as oportunidades de crescimento.

Os ciclos de mercado são uma característica inevitável das economias de mercado, refletindo as flutuações naturais na atividade econômica. Ao compreender as fases dos ciclos de mercado e os fatores que os influenciam, os investidores, empresas e formuladores de políticas podem tomar decisões mais informadas e adaptar suas estratégias para enfrentar os desafios e aproveitar as oportunidades apresentadas pelos ciclos econômicos. De acordo com Alvarez-Ramirez, Rodriguez e Ibarra-Valdez (2018) as dinâmicas do mercado de Bitcoin a longo prazo são diretamente influenciadas por fatores macroeconômicos, de maneira similar ao ocorrido para moedas e ativos tradicionais.

Na Figura 3 é apresentado um gráfico de análise dos ciclos do mercado da Bitcoin, demonstrando as fases de crescimento de mercado (*Bull Run*), retração do mercado (*Bear market*), expansão e acumulação.

Figura 3 – Ciclos de mercado da Bitcoin.



Fonte: Busa (2024).

## 2.4 INDICADORES

Os indicadores de mercado são ferramentas essenciais utilizadas por investidores, analistas e economistas para avaliar e entender o estado atual e as tendências futuras dos

mercados financeiros e da economia em geral. Esses indicadores fornecem conhecimentos valiosos sobre uma ampla gama de aspectos econômicos, desde a atividade do mercado de criptoativos até a saúde da economia de maneira mais universal. Aqui estão alguns dos principais tipos de indicadores de mercado:

- **Indicadores de Mercado de Ações:** Os indicadores de mercado de ações acompanham o desempenho e a volatilidade dos mercados de ações. O índice de referência mais comum é o S&P500 (*Standard & Poor's 500*), que rastreia o desempenho das 500 maiores empresas listadas nas bolsas de valores dos Estados Unidos. Outros indicadores incluem o DJIA (*Dow Jones Industrial Average*) e o COMP (*Nasdaq Composite*). Além disso, métricas como o volume de negociações, o P/L (Índice Preço/Lucro) e os EPS (Lucros Por Ação (*Earnings Per Share*)) são frequentemente monitorados para avaliar a saúde e a direção do mercado de ações.

- **Indicadores Econômicos:** Os indicadores econômicos fornecem informações sobre a saúde da economia de maneira integral. Isso inclui indicadores como o PIB (Produto Interno Bruto), que mede o valor total de todos os bens e serviços produzidos em uma economia, e a taxa de desemprego, que indica a porcentagem da força de trabalho que está desempregada e em busca de ocupação. Outros indicadores econômicos importantes incluem a inflação (medida pelo índice de preços ao consumidor), as vendas no varejo, a produção industrial e os gastos do consumidor.

- **Indicadores de Taxa de Juros:** As taxas de juros desempenham um papel crucial na determinação do comportamento dos consumidores e das empresas. Os indicadores de taxa de juros, são acompanhados de perto pelos participantes do mercado para avaliar as condições de empréstimo e financiamento. Alterações nas taxas de juros podem influenciar o investimento, o consumo e a inflação. Um exemplo desse tipo de indicador é a taxa básica de juros determinada pelo banco central de um país, como o caso do Selic (Sistema Especial de Liquidação e de Custódia) no Brasil.

- **Indicadores de Sentimento do Mercado:** Os indicadores de sentimento do mercado, como o Índice de Sentimento do Investidor, calculado semanalmente pela AII (*American Association of Individual Investors*) e o Índice de Medo e Ganância, medem as opiniões e expectativas dos investidores em relação ao mercado. Esses indicadores podem ajudar a identificar extremos de otimismo ou pessimismo que podem sinalizar possíveis reversões de tendência.

- **Indicadores de Câmbio e Mercado de Commodities:** Os indicadores de câmbio, como o índice do dólar americano, e os indicadores do mercado de commodities, como os preços do petróleo, ouro e outras matérias-primas, fornecem *insights* sobre as condições dos mercados cambiais e de commodities, que por sua vez podem impactar a economia global e os mercados financeiros.

Neste contexto, é crucial esclarecer que há distinção entre os termos “índice” e

“indicador”. Conforme observado por Siche *et al.* (2007) em seu artigo “Índices versus indicadores: precisões conceituais na discussão da sustentabilidade de países”, diferença está em que um índice é o valor agregado final de todo um procedimento de cálculo onde se utilizam, inclusive, indicadores como variáveis que o compõem.

Um índice de mercado é uma representação estatística do desempenho agregado de um conjunto específico de ativos financeiros, como ativos, títulos ou commodities. Geralmente, sua composição é baseada nos preços ou valores de mercado dos ativos incluídos no índice. Esses índices são comumente usados como *benchmarks* para comparar o desempenho de portfólios de investimentos, classes de ativos ou até mesmo a economia na totalidade. Por outro lado, um indicador de mercado é uma medida ou métrica que fornece informações sobre a condição, tendência ou saúde geral de um mercado financeiro. Podendo incluir uma variedade de dados e estatísticas, como volumes de negociação, volatilidade, liquidez, sentimento do investidor, taxas de juros e inflação, esses indicadores auxiliam investidores, analistas e gestores de fundos a compreender as atuais condições de mercado e tomar decisões de investimento embasadas.

No entanto, para simplificar este trabalho, não será feita distinção entre os termos “índice” e “indicadores de mercado”, uma vez que não se encontra no escopo e a aplicação desenvolvida não requer essa diferenciação específica.

Em resumo, os indicadores de mercado desempenham um papel fundamental na análise e tomada de decisões nos mercados financeiros e na economia em geral. Os indicadores de mercado de ativos fornecem informações valiosas sobre o comportamento dos investidores, a volatilidade do mercado e as tendências de preços dos ativos. Alguns dos principais indicadores são índices de mercado, volume de negociação, P/L, dividendo *yield* e volatilidade.

O volume de negociação representa o número de ativos compradas e vendidas em um determinado período. O volume é frequentemente usado como um indicador de liquidez e interesse dos investidores em determinados ativos ou no mercado na totalidade. Altos volumes de negociação podem indicar atividade significativa no mercado, enquanto baixos volumes podem sugerir desinteresse ou falta de confiança. A relação P/L é uma medida que compara o preço atual de uma ação com seus EPS. É calculada dividindo o preço da ação pelo EPS. Um P/L alto pode indicar que uma ação está cara em relação aos seus lucros, enquanto um P/L baixo pode sugerir que uma ação está subvalorizada.

O dividendo *yield* é uma medida que compara o dividendo anual por ação de uma empresa com o preço de mercado da ação. É expresso como uma porcentagem e indica o retorno em dividendos que um investidor pode esperar receber em relação ao preço da ação. Um alto dividendo *yield* pode atrair investidores em busca de renda estável. Por fim, a volatilidade é uma medida da flutuação dos preços dos ativos ao longo do tempo. Um mercado altamente volátil é caracterizado por grandes oscilações nos preços dos ativos, enquanto um mercado de baixa volatilidade é mais estável. A volatilidade é frequentemente

medida pelo VIX (*Volatility Index*), também conhecido como o “índice do medo”.

Esses são apenas alguns dos indicadores de mercado de ativos mais comumente utilizados. Ao analisar esses indicadores em conjunto, os investidores podem obter uma visão mais abrangente do estado atual e das tendências futuras do mercado de ativos, ajudando-os a tomar decisões de investimento informadas. Portanto, visando o treinamento da rede neural baseado para tomada de decisão relacionada ao mercado financeiro, foram utilizados os indicadores apresentados no Quadro 1.

Quadro 1 – Indicadores de Mercado Aplicados.

<b>Categoria</b>	<b>Indicadores</b>
Tendência	SMA20 (Média Móvel Simples 20 Dias)
	SMA50 (Média Móvel Simples 50 Dias)
	SMA100 (Média Móvel Simples 100 Dias)
	SMA200 (Média Móvel Simples 200 Dias)
	EMA20 (Média Móvel Exponencial 20 Dias)
	EMA50 (Média Móvel Exponencial 50 Dias)
	EMA100 (Média Móvel Exponencial 100 Dias)
	EMA200 (Média Móvel Exponencial 200 Dias)
	WMA20 (Média Móvel Ponderada 20 Dias)
	WMA50 (Média Móvel Ponderada 50 Dias)
	WMA100 (Média Móvel Ponderada 100 Dias)
	WMA200 (Média Móvel Ponderada 200 Dias)
	DEMA20 (Média Móvel Exponencial Dupla 20 Dias)
	DEMA50 (Média Móvel Exponencial Dupla 50 Dias)
	DEMA100 (Média Móvel Exponencial Dupla 100 Dias)
DEMA200 (Média Móvel Exponencial Dupla 200 Dias)	
Momento	MACD (Média Móvel Convergente e Divergente)
	MACDSIGNAL (Linha de Sinal da Média Móvel Convergente e Divergente)
	MACDHIST (Dados Para Histograma da Média Móvel Convergente e Divergente)
	ROC (Taxa de Variação)
	Stoch (Oscilador Estocástico)
	RSI (Índice de Força Relativa)
	Stoch RSI (Índice de Força Relativa Estocástico)
	ADX (Índice de Direção Média)
	APO (Oscilador de Preço Absoluto)
	OBV (Volume em Balanço)
	MOM (Momento)
WILLR (Indicador Williams %R)	
MFI (Índice Fluxo de Dinheiro)	
Volatilidade	ATR (Média de Amplitude de Variação)
	NATR (Média de Amplitude de Variação Normalizada)
	STDDEV (Desvio Padrão)
	BBANDS (Bandas de Bollinger)
Volume	AD (Acumulação/Distribuição)



Masiak *et al.* (2018)

### 3 FUNDAMENTAÇÃO TEÓRICA

Esse capítulo é dedicado a apresentação de conceitos relacionados ao projeto desenvolvido neste trabalho, com enfoque em redes neurais profundas, ferramentas utilizadas para sua elaboração, literatura e outros conceitos relacionados.

#### 3.1 REDE NEURAL PROFUNDA

As redes neurais representam uma classe de modelos computacionais que se baseiam no funcionamento do cérebro humano para resolver uma variedade de problemas complexos em diversas áreas, desde reconhecimento de padrões até previsão de séries temporais. Inspiradas na estrutura e no funcionamento dos neurônios biológicos, essas redes são compostas por unidades interconectadas chamadas neurônios artificiais, ou unidades de processamento, que realizam operações matemáticas em entradas para produzir saídas.

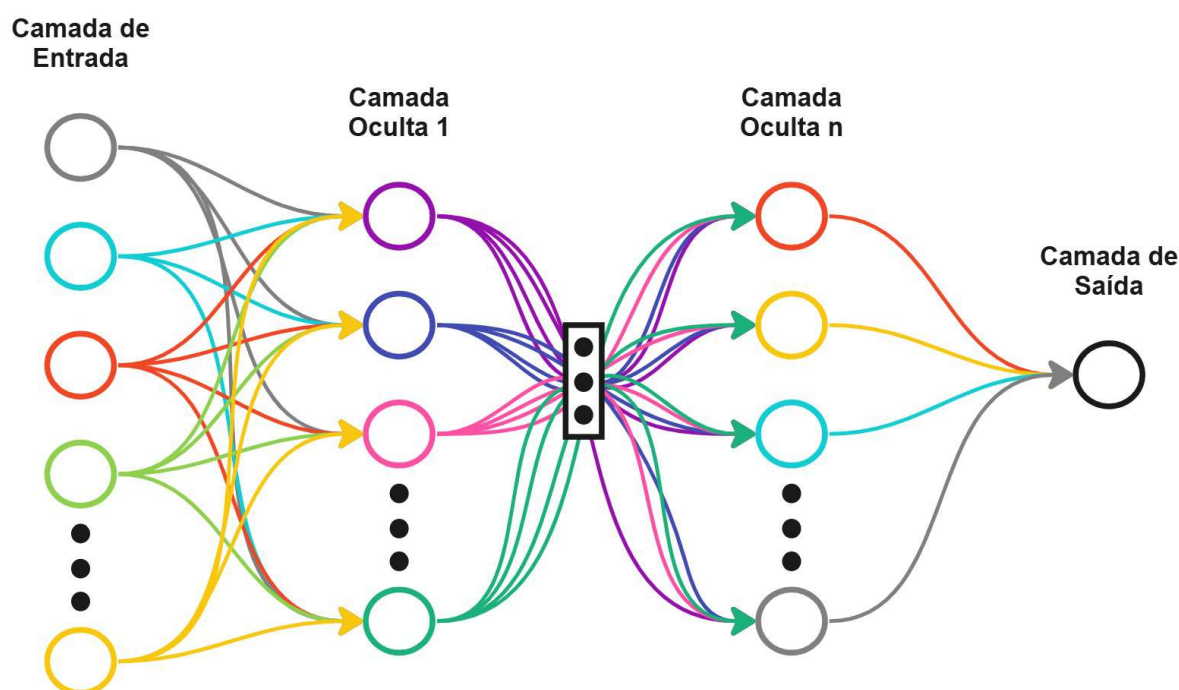
Cada neurônio artificial (nó) recebe um conjunto de entradas, realiza um cálculo com essas entradas, aplicando pesos às conexões, e produz uma saída. Essas saídas são então transmitidas para outros neurônios na rede, formando assim uma estrutura interconectada. Os pesos das conexões entre os neurônios determinam a importância relativa das entradas para a saída de um neurônio, sendo ajustados durante o processo de treinamento da rede.

A principal diferença entre uma rede neural e uma rede neural profunda reside na quantidade de camadas ocultas presentes na estrutura da rede. Uma rede neural tradicional, também conhecida como rede neural rasa, consiste geralmente em três tipos de camadas: a camada de entrada, uma camada oculta (ou poucas camadas ocultas) e a camada de saída. Essa estrutura relativamente simples é adequada para resolver problemas menos complexos e pode ser eficiente em tarefas como classificação linear ou reconhecimento de padrões básicos. No entanto, sua capacidade de modelar relações mais complexas e sutis nos dados é limitada devido ao número reduzido de camadas.

Por outro lado, uma rede neural profunda contém múltiplas camadas ocultas entre a camada de entrada e a camada de saída. Essas camadas adicionais permitem que a rede aprenda representações hierárquicas dos dados, capturando características de alto nível e complexidade progressivamente maiores. Com mais camadas, uma rede neural profunda pode modelar padrões intrincados e realizar tarefas avançadas como reconhecimento de imagem, processamento de linguagem natural e análise preditiva em um nível muito mais sofisticado. A profundidade adicional permite que a rede capture interações não-lineares complexas, tornando-a extremamente poderosa para uma ampla gama de aplicações, mas também exigindo mais dados e maior poder computacional para treinamento eficaz. Na Figura 4 é apresentado um exemplo de rede neural profunda.

As redes neurais têm sido amplamente utilizadas em uma variedade de campos, incluindo reconhecimento de imagem, veículos robóticos, diagnóstico médico, finanças,

Figura 4 – Arquitetura Genérica de redes neurais.



Fonte: (SERVICES, s.d.) (adaptado).

reconhecimento de voz, etc. (RUSSELL; NORVIG, 2009). Entretanto, um dos aspectos mais complicados do aprendizado de máquina é a determinação do melhor modelo para cada problema ou conjunto de dados. Em geral, o problema mais enfrentado é o *overfitting* que aparece durante o ajuste de flexibilidade e indica que o modelo encontra-se mais alinhado com os ruídos presentes nos dados de treinamento do que com o sinal previamente modelado (HALLS-MOORE, 2017).

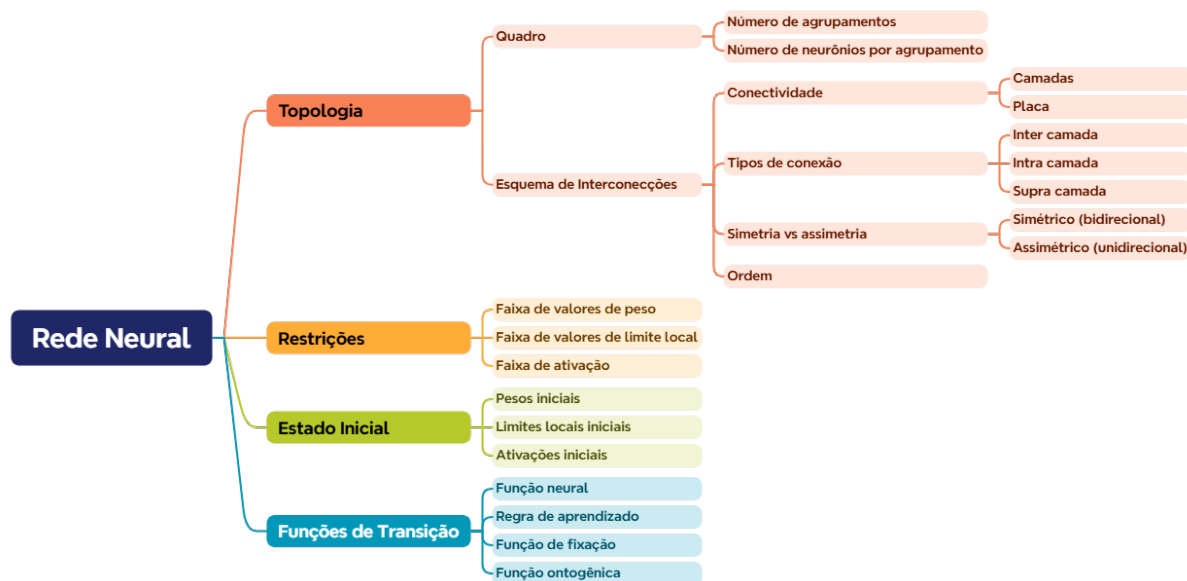
Uma das características mais poderosas das redes neurais é sua capacidade de aprendizado a partir dos dados. Durante o treinamento, a rede é exposta a um conjunto de exemplos de entrada e saída esperada, ajustando os pesos das conexões para minimizar a diferença entre as saídas produzidas pela rede e as saídas esperadas. O algoritmo de aprendizado mais comumente utilizado é o de retro propagação (*backpropagation*), que calcula o gradiente da função de perda em relação aos pesos das conexões e utiliza esse gradiente para atualizar os pesos para minimizar a perda.

Existem várias arquiteturas de redes neurais, cada uma com suas próprias características e aplicações específicas. As redes neurais *feedforward* consistem em camadas de neurônios onde as informações fluem em uma única direção, da entrada para a saída, sem ciclos. As redes neurais recorrentes têm conexões retroativas que permitem que informações sejam propagadas em ciclos, sendo adequadas para modelar sequências temporais. Já as redes neurais convolucionais são eficazes em tarefas de visão computacional, onde a

estrutura espacial dos dados é importante.

Devido à variedade de possibilidades ao definir e modelar uma rede neural, existem certos desafios ao realizar a classificação de cada arquitetura. Fiesler (1994) sugere uma padronização para a determinação do tipo de rede neural baseado em algumas de suas características como topologia, restrições, estado inicial e funções de transição Figura 5.

Figura 5 – Classificação de redes neurais.



Fonte: (FIESLER, 1994) (adaptado)

## 3.2 PYTHON E FERRAMENTAS

Python é uma linguagem de programação de alto nível, orientada a objetos, criada por Guido van Rossum. Segundo Srinath (2017) o Python se tornou a linguagem de programação com o crescimento mais rápido e define a linguagem como “uma linguagem de código aberto rápida, simples, poderosa e portátil que suporta outras tecnologias”. De acordo com uma pesquisa realizada em 2023 e respondida por mais de 85000 pessoas no Stack Overflow (OVERFLOW, 2023), plataforma que hospeda a maior comunidade online de desenvolvedores do mundo, o Python é a terceira linguagem de programação mais utilizada entre todos os entrevistados e a mais utilizada entre desenvolvedores não profissionais ou aqueles que estão aprendendo a programar.

Amplamente utilizada em uma variedade de domínios, desde desenvolvimento web até ciência de dados e aprendizado de máquina, Python ganhou popularidade devido à sua vasta coleção de bibliotecas e frameworks poderosos. Alguns dos pacotes mais utilizadas são o NumPy (*Numerical Python*), o SciPy (*Scientific Python*), o Pandas, o Matplotlib e o Scikit-learn (FAREEZ; THANGARAJAH; SAABITH, 2020).

Quando se trata de aprendizado de máquina, o PyTorch é um dos poucos frameworks disponíveis para o Python. Essa biblioteca, que foi desenvolvida pelo FAIR (*Fundamental AI Research*) e lançada em 2016, é baseada no Torch e possui uma vasta quantidade de ferramentas e bibliotecas que suporta desenvolvimento para áreas como visão computacional e NLP (Processamento de Linguagem Natural). O PyTorch é amplamente utilizado para desenvolver aplicativos de aprendizado profundo, como redes neurais profundas e modelos de aprendizado de máquina. Ele oferece uma interface flexível e intuitiva que facilita a experimentação e a prototipagem rápida de modelos complexos. Além disso, o PyTorch possui uma comunidade ativa e em crescimento, o que significa que os usuários têm acesso a uma ampla gama de recursos, tutoriais e suporte online.

Uma das características distintivas do PyTorch é sua ênfase na programação dinâmica. Ao contrário de algumas outras estruturas de aprendizado profundo que adotam uma abordagem estática de gráficos computacionais, o PyTorch permite a definição e execução de grafos de computação de forma dinâmica durante o tempo de execução. Isso oferece maior flexibilidade e facilita a depuração e o desenvolvimento iterativo de modelos de aprendizado profundo.

Outro aspecto importante do PyTorch é seu suporte à GPU (Unidade de Processamento Gráfico), que permite acelerar significativamente o treinamento de modelos em grandes conjuntos de dados. Essa capacidade de aproveitar o poder de processamento paralelo da GPU é fundamental para o treinamento eficiente de redes neurais profundas e modelos de aprendizado de máquina em escala.

Por fim, para a criação de uma base de dados com informações coletadas diretamente do mercado financeiro, a Binance, uma das maiores e mais populares plataformas de câmbio de criptomoedas do mundo, disponibiliza uma API (Interface de Programação de Programação) através da qual os desenvolvedores podem acessar e interagir com dados de mercado em tempo real, fazer negociações automatizadas, gerenciar contas de usuário e muito mais.

A Binance API oferece uma ampla gama de recursos e funcionalidades para desenvolvedores interessados em construir aplicativos e serviços que interagem com a plataforma Binance. Isso inclui *endpoints* para obter informações sobre pares de negociação, consultar históricos de preços, fazer pedidos de compra e venda, acessar dados de saldo de conta e muito mais.

Ao combinar Python, PyTorch e Binance API, os desenvolvedores podem criar aplicativos sofisticados e poderosos para análise de dados de mercado, execução de estratégias de negociação automatizadas e até mesmo desenvolvimento de modelos de aprendizado de máquina para prever tendências de mercado e otimizar estratégias de negociação.

A flexibilidade e a simplicidade do Python, juntamente com a poderosa estrutura de aprendizado profundo do PyTorch e a ampla gama de recursos oferecidos pela Binance API, tornam essa combinação uma escolha popular para desenvolvedores e *traders* interessados

em explorar o mundo das criptomoedas e da negociação algorítmica.

No entanto, é importante observar que o uso da Binance API para negociação automatizada e acesso a dados de mercado em tempo real requer cuidado e atenção às melhores práticas de segurança e conformidade regulatória. Os desenvolvedores devem garantir que seus aplicativos e sistemas estejam devidamente protegidos contra vulnerabilidades de segurança e cumpram todas as regulamentações relevantes.

### 3.3 TRABALHOS RELACIONADOS

Conforme os resultados publicados por de Faria *et al.* (2009), ao comparar a predição de ativos através de ANN (*Artificial Neural Network*) e por meio do AES (*Adaptive Exponential Smoothing*) (TRIGG; LEACH, 1967) no mercado brasileiro, ambos os métodos demonstram resultados similares em relação à predição dos retornos do índice, entretanto a ANN supera o método de AES (*Adaptive Exponential Smoothing*) quando se trata da previsão dos movimentos de mercado, com taxas de acerto relativo semelhante às encontradas em mercados desenvolvidos.

Um dos principais benefícios da IA na predição de ativos é sua capacidade de aprender e se adaptar continuamente. Modelos de aprendizado profundo, como as redes neurais convolucionais e recorrentes, podem ser treinados com vastos conjuntos de dados históricos, incluindo preços de ativos, volumes de negociação, notícias econômicas e dados de redes sociais. Ao incorporar essa vasta gama de informações, esses modelos podem ajustar suas previsões com base em novos dados e padrões emergentes, melhorando continuamente sua precisão. Além disso, técnicas como o aprendizado por reforço permitem que a IA desenvolva estratégias de negociação adaptativas que podem responder rapidamente a mudanças nas condições do mercado.

Sendo assim, apesar da quantidade de publicações relacionadas ao mercado de criptomoedas ser limitada, existem diversos trabalhos que investigam a aplicação do aprendizado de máquina para a previsão do valor de ativos no mercado financeiro. No estudo de Oliveira, Nobre e Zárate (2013), por exemplo, é apresentado a aplicação de um modelo baseado em perceptron multicamadas *feedforward* para a predição da ação PETR4 da empresa brasileira Petrobras, comercializada na BM&FBOVESPA. Apesar de não se aprofundar nas especificações do modelo de ANN utilizada, os autores demonstram a viabilidade da aplicação desse método para a previsão do comportamento e tendências do mercado de criptoativos.

No contexto da arquitetura de redes neurais, a RNN (*Recurrent Neural Network*) se sobressai em aplicações que envolvem dados sequenciais ou séries temporais, graças ao seu desempenho e adequação específicos. Essa característica a torna essencial para a previsão de dados no mercado financeiro, onde a análise temporal é um requisito fundamental. Entretanto, as RNN's tendem a apresentar dois principais problemas: gradientes explosivos e gradientes que desaparecem, fatores que dificultam a aprendizagem eficaz quando há

a necessidade de reter informações relevantes por longos períodos. A arquitetura LSTM (*Long Short-Term Memory*), desenvolvida por Hochreiter e Schmidhuber (1997), oferece uma solução a esse problema por meio de uma estrutura de rede mais sofisticada que permite a retenção de informações por períodos prolongados, tornando-as especialmente adequadas para tarefas onde o contexto histórico é crucial.

Nos resultados apresentados por Jaquart, Dann e Weinhardt (2021), Agarwal e Muppalaneni (2022) e Mokhtari, Yen e Jin Liu (2021) é possível verificar o desempenho da LSTM em diferentes estudos de caso e sua comparação com outros modelos como RNN, ARIMA (*Autoregressive Integrated Moving Average*), SVM (*Support Vector Machine*) e GNB (*Gaussian Naive Bayes*).

## 4 PROPOSTA E IMPLEMENTAÇÃO

A arquitetura da LSTM é constituída por uma série de “células de memória”, cada uma equipada com três portas: a porta de entrada, a porta de esquecimento e a porta de saída. A porta de entrada determina quais valores das novas entradas serão usados para atualizar o estado da célula, enquanto a porta de esquecimento decide quais informações do estado anterior devem ser descartadas. Finalmente, a porta de saída controla quais partes do estado da célula serão utilizadas para gerar a saída atual. Essas portas são implementadas usando funções sigmóides, que produzem valores entre 0 e 1, permitindo um controle preciso sobre o fluxo de informações dentro da célula. O fluxo de dados em uma LSTM é demonstrado na Figura 6.

As portas da LSTM operam por meio de funções sigmóides e operações ponto a ponto que garantem um fluxo de informações controlado e eficiente. A porta de entrada, via uma função sigmoide, filtra quais valores das entradas atuais devem influenciar o estado da célula. A porta de esquecimento também utiliza uma função sigmoide para decidir quais informações armazenadas anteriormente podem ser descartadas, permitindo que a rede “esqueça” informações que não são mais relevantes. A porta de saída, por sua vez, aplica uma função sigmoide ao estado da célula, determinando quais partes do estado atual contribuem para a saída da célula. Essa estrutura permite que as LSTM’s mantenham um equilíbrio entre a retenção de informações cruciais e a eficiência computacional.

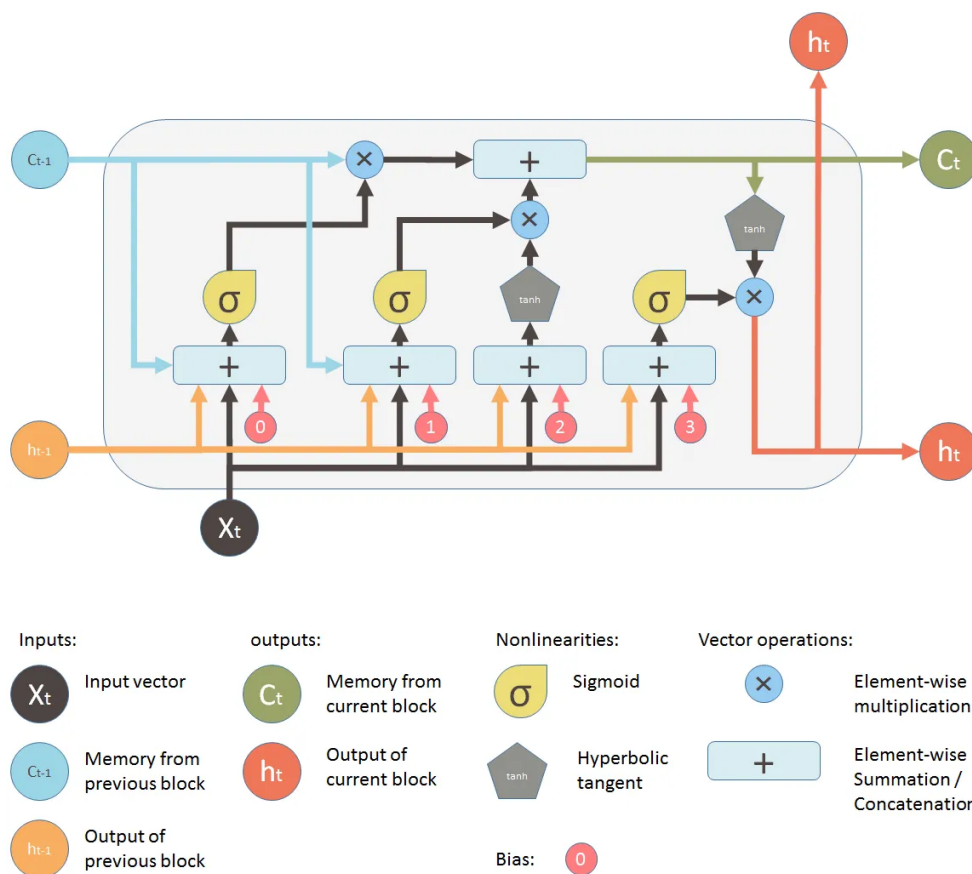
LSTM’s são amplamente utilizadas em uma variedade de aplicações que envolvem dados sequenciais e temporais. No reconhecimento de fala, por exemplo, as LSTM’s capturam dependências de longo prazo necessárias para interpretar o contexto de uma conversação de forma mais precisa. Na tradução automática, elas ajudam a manter a coerência e o significado ao traduzir frases inteiras, mesmo quando as palavras podem estar distantes umas das outras. Além disso, em tarefas de análise de séries temporais, como previsão de preços de ativos ou detecção de anomalias, as LSTM’s são capazes de prever tendências futuras com base em padrões históricos.

As LSTM’s oferecem várias vantagens significativas, principalmente na sua capacidade de lidar com dependências de longo prazo sem sofrer com os problemas comuns das RNN’s tradicionais, como o desaparecimento do gradiente. No entanto, essa capacidade vem com um custo computacional. As LSTM’s são complexas e exigem recursos consideráveis para treinamento, especialmente quando aplicadas a grandes conjuntos de dados. Ajustar uma LSTM para um problema específico pode ser uma tarefa desafiadora, exigindo experimentação cuidadosa e afinação de hiperparâmetros. Apesar dessas limitações, as LSTM’s continuam sendo uma ferramenta essencial em aprendizado de máquina para tarefas que envolvem sequências de dados, graças à sua robustez e capacidade de capturar e utilizar dependências temporais complexas.

Portanto, esse trabalho propõe a implementação de uma rede neural do tipo LSTM



Figura 6 – Diagrama do fluxo de dados em uma LSTM.



Fonte: (YAN, 2016).

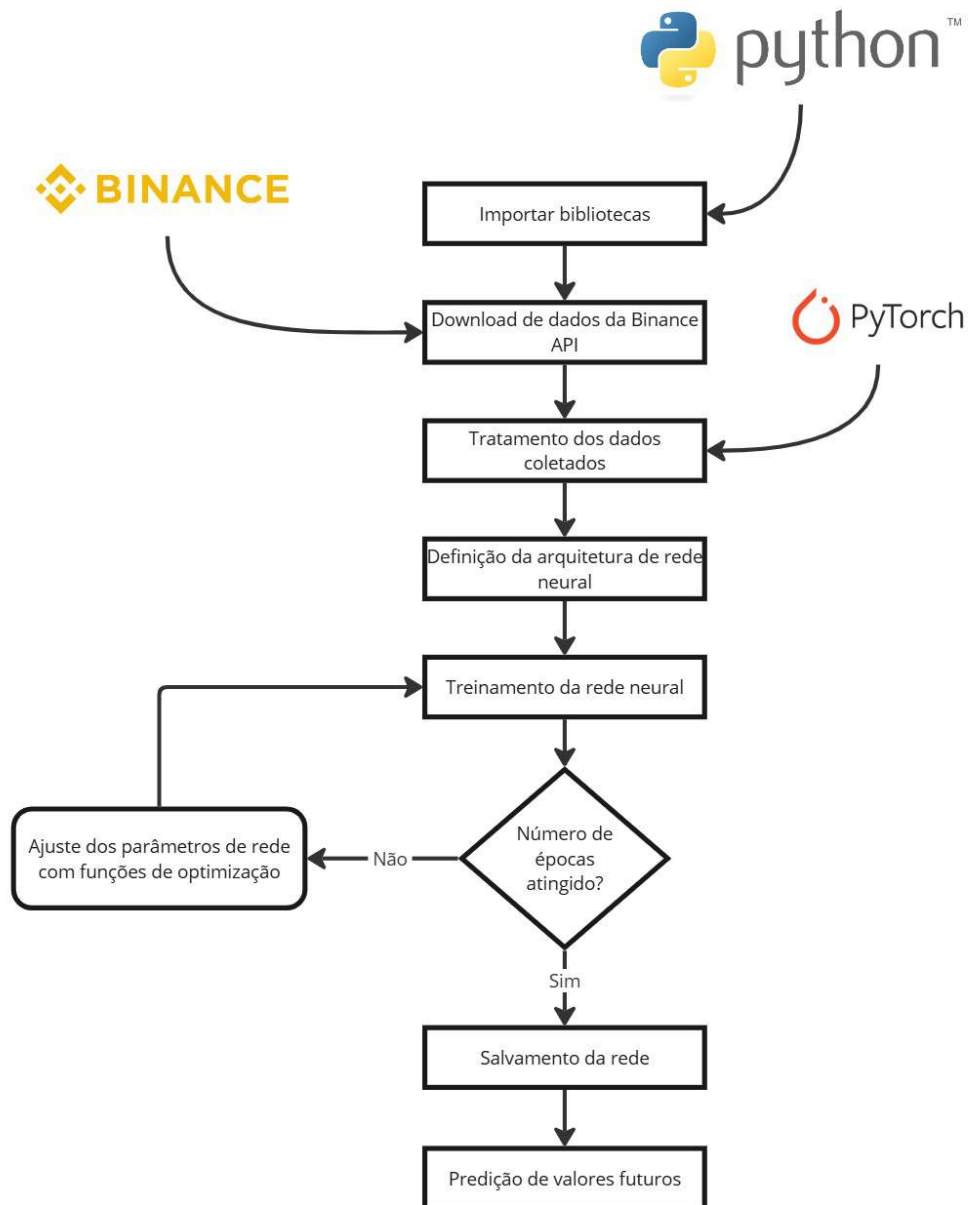
para a realização da predição dos valores de fechamento da Bitcoin no mercado financeiro de ativos baseado no histórico de valores coletados.

#### 4.1 IMPLEMENTAÇÃO PARA A PREDIÇÃO DE VALORES

A definição e o treinamento da rede neural desenvolvida foram realizados utilizando a linguagem de programação Python. O código seguiu a seguinte lógica: inclusão das bibliotecas necessárias, utilização da API da corretora Binance para o download dos dados históricos, tratamento dos dados coletados, definição da arquitetura da LSTM, treinamento e salvamento da rede, e, finalmente, predição dos valores futuros, como apresentado na Figura 7.

Devido às características do aprendizado de máquina, os resultados obtidos são fortemente influenciados pela qualidade dos dados utilizados no treinamento, o que torna a etapa de pré-processamento de dados essencial. Nas publicações disponíveis na literatura, os conjuntos de dados incluem frequentemente dados históricos, como valor de abertura, valor máximo, valor mínimo e volume negociado do ativo analisado, também utilizados

Figura 7 – Diagrama de fluxo do código desenvolvido



para a representação gráfica tradicional de *candlesticks* demonstrada na Figura 8. No entanto, para alcançar resultados mais precisos, além dos dados históricos presentes nos *candlesticks* este projeto também utilizou como informações de entrada da rede os indicadores apresentados no Quadro 1, que ajudam os negociadores na tomada de decisões no mercado de ativos. Além disso, o conjunto de dados foi verificado para eliminar informações ruidosas ou ausentes, e passou por processos de normalização e rearranjo em sequências para serem inseridas no modelo durante o treinamento. A última etapa do pré-processamento de dados é a criação de três diferentes conjuntos: treinamento, validação e avaliação.

Na Tabela 1 é apresentado um exemplo dos dados inseridos na entrada do modelo

Figura 8 – *Candlesticks* obtidos na plataforma Binance.

desenvolvido contendo os valores de *candlesticks* obtidos através da API da Binance e os indicadores calculados para as primeiras 3 horas de informações. Esse projeto utilizou um conjunto de dados contendo informações de aproximadamente 58.200 horas.

No caso das redes LSTM, a qualidade dos resultados é diretamente afetada pelos hiperparâmetros e funções de perda empregados. Para minimizar as funções de erro e maximizar a eficiência do modelo, são utilizados algoritmos de otimização; no entanto, essas funções podem acarretar um alto custo computacional. Devido à sua eficiência em problemas que empregam uma alta quantidade de parâmetros e dados, o algoritmo de otimização utilizado foi o Adam (*Adaptive Moment Estimation*). Certos parâmetros do modelo, conforme apresentados no Quadro 2, foram fixados para viabilizar a comparação dos resultados obtidos.

Quadro 2 – Hiperparâmetros Fixos.

Parâmetro	Valor
Épocas	6000
Taxa de aprendizado	0.001
Ativo	BTCUSDT
Camadas Escondidas	10
Início do conjunto de dados de treinamento	21/08/2017
Início do conjunto de dados de validação	01/04/2024
Intervalo de tempo	1h

Por fim, foram determinados diferentes estudos de caso que diferenciam entre si em relação às características de tempo de antecipação da predição, quantidade de camadas empilhadas, tamanho da sequência de dados inserida no modelo e função de perda aplicada.

Tabela 1 – Informações de entrada para o treinamento da rede

<i>Parâmetros</i>	<i>Valores</i>		
<i>Datetime</i>	2017 – 09 – 02 18 : 00 : 00	2017 – 09 – 02 19 : 00 : 00	2017 – 09 – 02 20 : 00 : 00
<i>Open</i>	4554,98	4486,06	4509,77
<i>High</i>	4554,98	4509,96	4509,96
<i>Low</i>	4480,81	4428,66	4286,87
<i>Close</i>	4486,06	4509,77	4351,19
<i>Volume</i>	31,20194	23,438896	57,065876
SMA20	4652,674499999999	4636,417499999999	4609,699499999999
SMA50	4704,9962000000005	4702,180600000001	4695,201800000001
SMA100	4642,949899999996	4642,347599999996	4640,820299999996
SMA200	4481,52585	4482,26835	4482,47585
EMA20	4609,98867967415	4600,444043514707	4576,705563179973
EMA50	4656,55010123362	4650,794018832302	4639,044841623192
EMA100	4614,865671409152	4612,784569005011	4607,604478529664
EMA200	4507,4360767449025	4507,459299861868	4505,904381455282
WMA20	4582,22438095247	4568,614428571518	4541,449904761995
WMA50	4681,253772549002	4673,597843137238	4659,833505882335
WMA100	4671,147324752494	4668,510099009919	4662,744601980216
WMA200	4570,773709950254	4571,054746268662	4569,750484079606
DEMA20	4523,549469210583	4513,601515617698	4476,656079541729
DEMA50	4662,536198359186	4651,015013371285	4627,968744547972
DEMA100	4727,220708672694	4720,874861589968	4708,476854854926
DEMA200	4720,914447295394	4718,836499562485	4713,638878855841
MACD	-70,26499149304163	-69,19695627835154	-80,22186365968173
MACDSIGNAL	-49,98589487805234	-53,82810715811218	-59,10685845842609
MACDHIST	-20,279096614989292	-15,368849120239354	-21,11500520125564
ROC	-3,084992849211232	-2,0471240350823816	-5,931606119042899
Stoch <i>slowk</i>	46.054239523048466	63,086562398409335	50,42002988371541
Stoch <i>slowd</i>	37.81613878817063	49,70644188931825	53,1869439350578
RSI	39.14364636736051	41,18174539435932	33,17832908261295
Stoch RSI <i>fastk</i>	69.37935007642497	94,51647548203795	0,0
Stoch RSI <i>fastd</i>	50.80852910073978	82,31402092808577	54,63194185282095
ADX	44.01379863971664	44,42557918020572	45,52662824067462
APO	-119.6932051282065	-124,46256410256592	-131,45211538461717
OBV	180.08470500000004	203,52360100000004	146,45772500000004
MOM	-142.79999999999927	-94,25	-274,37000000000008
WILLR	-71.41334841136926	-63,52523566609449	-86,95678624297858
MFI	49.07035841442875	52,48493005381364	41,42217197386073
ATR	114,82578939220905	112,43109014990841	120,33529799634354
NATR	2,5596133219843034	2,4930559684841667	2,7655721307583336
STDDEV	30,41892634585945	32,778384707572435	63,69741049712548
BBANDS <i>Uper Band</i>	4948,2446016324575	4925,808081247966	4900,816396210152
BBANDS <i>Middle Band</i>	4652,674499999999	4636,417499999999	4609,699499999999
BBANDS <i>Lower Band</i>	4357,104398367541	4347,026918752032	4318,582603789846
AD	247,89792670676238	271,2272682113136	247,06718937227987

Para cada cenário apresentado na Figura 9 foram realizados o treinamento de uma rede neural.

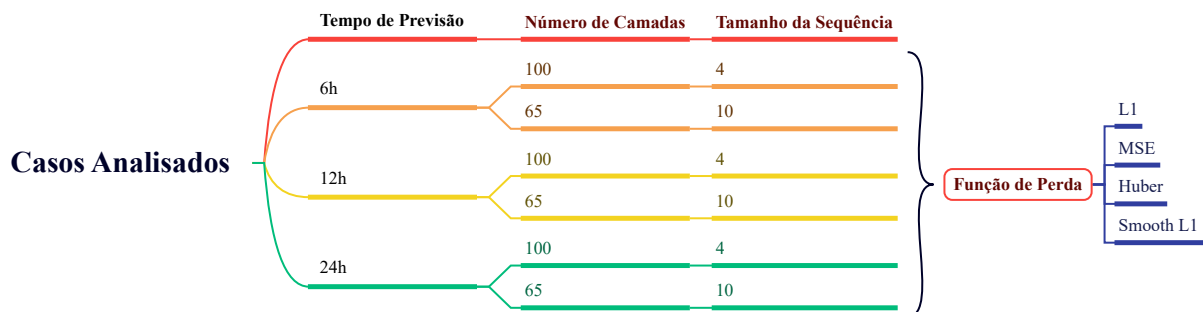
## 5 RESULTADOS

A partir da definição dos casos a serem analisados, foram realizados treinamentos com as diferentes características e coletados os dados de valores previstos durante as fases de treinamento, teste e validação, convergência da função de perda e as métricas de avaliação de regressão:  $R^2$  (Coeficiente de Determinação) e MSE (*Mean Squared Error*). Esses dados foram utilizados para realizar a avaliação de cada modelo e os resultados são apresentados ao longo desse capítulo.

### 5.1 ESTUDOS REALIZADOS

Tomando em consideração a arquitetura de LSTM desenvolvida, foram definidos 24 diferentes cenários para serem testados e analisados variando parâmetros como tempo de previsão, número de camadas, tamanho da sequência de dados inserida no modelo e função de perda utilizada, como demonstrado na Figura 9. Portanto, foi realizado o treinamento de uma rede neural para cada cenário apresentado no Quadro 3 e os resultados obtidos serão apresentados ao longo desse Capítulo 5, separando por tipo de função de perda e tempo de previsão para cada função.

Figura 9 – Casos estudados.



### 5.2 FUNÇÃO DE PERDA MAE

O primeiro caso analisado foi o treinamento da LSTM utilizando a função de perda L1, também conhecida como MAE (*Mean Absolute Error*). Essa função calcula a diferença absoluta entre os valores previstos e real conforme a Equação (1).

$$J_{MAE} = (1/n) \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

Onde:

- $J_{MAE}$  = resultado da função de perda MAE

Quadro 3 – Cenários criados para análise

Cenário	Função de Perda	Tempo de Previsão	Número de Camadas	Tamanho da Sequência
1	MAE	6h	100	4
2			65	10
3		12h	100	4
4			65	10
5		24h	100	4
6			65	10
7	MSE	6h	100	4
8			65	10
9		12h	100	4
10			65	10
11		24h	100	4
12			65	10
13	Huber	6h	100	4
14			65	10
15		12h	100	4
16			65	10
17		24h	100	4
18			65	10
19	Smooth MAE	6h	100	4
20			65	10
21		12h	100	4
22			65	10
23		24h	100	4
24			65	10

- $y_i$  = valor real de fechamento do ativo
- $\hat{y}_i$  = valor previsto pelo modelo
- $n$  = número de dados no *dataset*

### 5.2.1 MAE: 6 Horas de Previsão

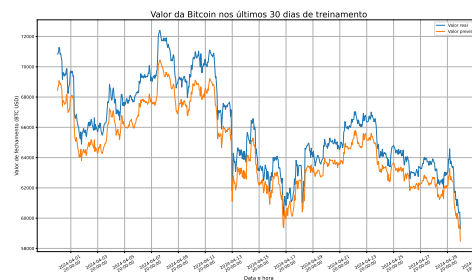
Utilizando a função de perda MAE e realizando a previsão dos valores de fechamento do ativo Bitcoin para as próximas 6 horas, o modelo foi treinado com 100L (Número de Camadas Igual à 100) e 4S (Sequência de Dados de Entrada Igual à 4) (Figura 10a) e 65L (Número de Camadas Igual à 65) e 10S (Sequência de Dados de Entrada Igual à 10) (Figura 10c).

Para melhor visualização da proximidade entre os dados reais e os dados previstos durante o treinamento para cada caso de previsão de 6 horas com a função MAE, as Figuras 10b e 10d apresentam uma comparação nos últimos 30 dias de dados.

Figura 10 – MAE - Treinamento para 6 horas de previsão



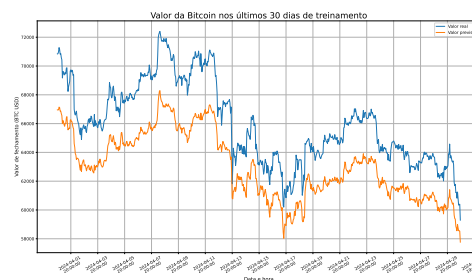
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



(d) 65L e 10S - Gráfico parcial

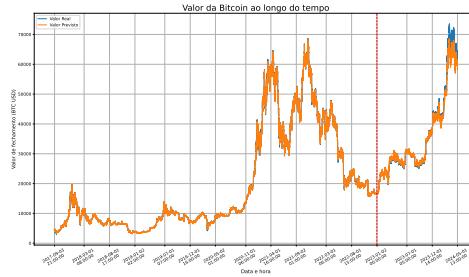
### 5.2.2 MAE: 12 Horas de Previsão

Em seguida, foram realizados treinamentos para prever os valores de fechamento dos criptoativos de Bitcoin para as próximas 12 horas, o modelo também foi treinado com 100L e 4S (Figura 11a) e 65L e 10S (Figura 11c).

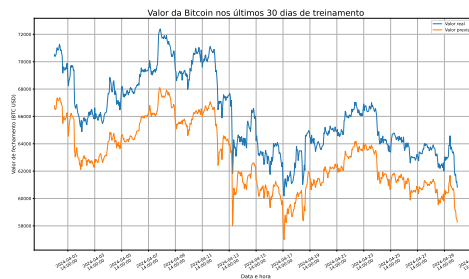
Assim como realizado para o primeiro período de previsão, também foram criados gráficos demonstrando cada modelo treinado para a previsão de 12 horas com a função

MAE com os dados dos últimos 30 dias de treinamento, apresentado nas Figuras 11b e 11d.

Figura 11 – MAE - Treinamento para 12 horas de previsão



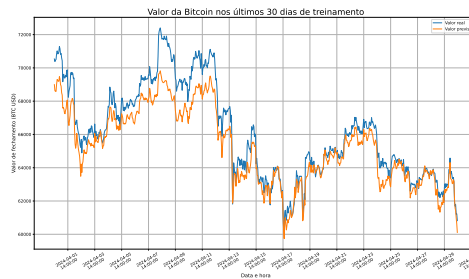
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



(d) 65L e 10S - Gráfico parcial

### 5.2.3 MAE: 24 Horas de Previsão

Por fim, para cada função de perda também foram realizados treinamentos para a previsão dos valores de fechamento dos criptoativos de Bitcoin para as próximas 24



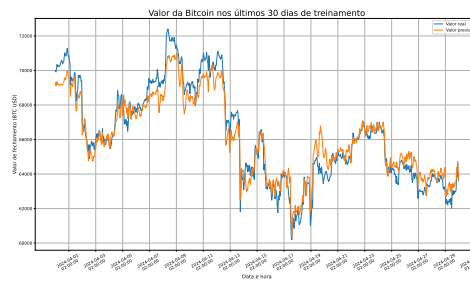
horas, também foram considerados os modelos com 100L e 4S (Figura 12a) e 65L e 10S (Figura 12c).

As Figuras 12b e 12d apresentam a comparação entre os dados reais e os dados previstos pelo modelo durante a fase de treinamento para os últimos 30 dias de informações presente no *dataset*.

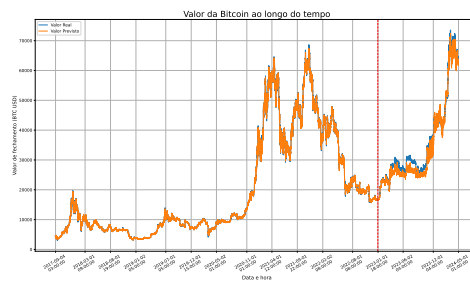
Figura 12 – MAE - Treinamento para 24 horas de previsão



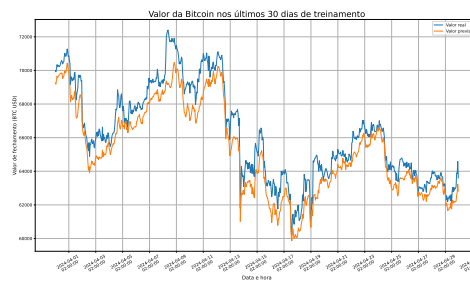
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



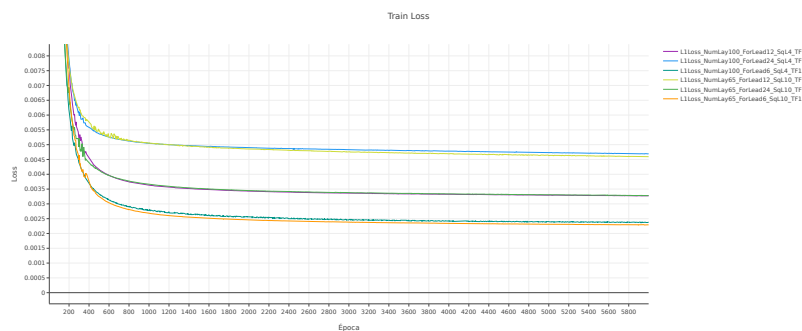
(d) 65L e 10S - Gráfico parcial

### 5.2.4 MAE: Convergência da Função de Perda

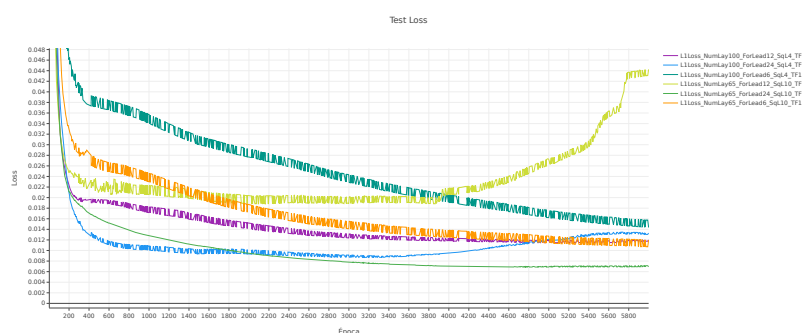
Um dos métodos de avaliação do comportamento da rede neural durante o treinamento é a análise dos gráficos de função de perda. Ao observar a convergência dessa função é possível detectar se o modelo está passando por um bom aprendizado, sendo subajustado (*underfitting*) ou sendo superajustado (*overfitting*). Situações de ocorrência de subajuste, superajuste ou grandes oscilações são indicativos que ajuste dos hiperparâmetros, como taxa de aprendizado e complexidade do modelo, é necessário.

No caso da utilização da função de perda MAE, as Figuras 13a e 13b apresentam a convergência durante as fases de treinamento e teste, respectivamente. Como é possível observar, de maneira geral, a função apresenta queda no valor de perda e certa estabilização para todos os casos criados. Entretanto, no gráfico da função durante a fase de teste é possível observar algumas discrepâncias, com destaque para previsão de 12 horas com 65L e 10S.

Figura 13 – Convergência da função MAE



(a) Fase de treinamento



(b) Fase de teste.

## 5.3 FUNÇÃO DE PERDA *MEAN SQUARED ERROR* (MSE)

O próximo caso analisado foi o treinamento da LSTM utilizando a função de perda MSE (*Mean Squared Error*). Essa função calcula a média quadrática da diferença entre os valores previstos e real conforme a Equação (2).

$$J_{MSE} = (1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Onde:

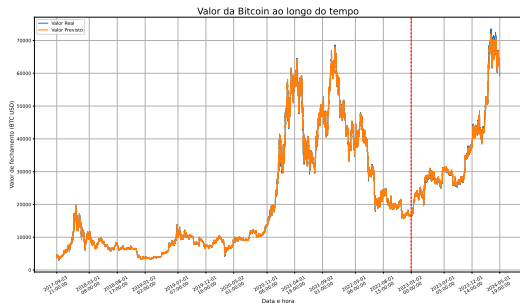
- $J_{MSE}$  = resultado da função de perda MSE
- $y_i$  = valor real de fechamento do criptoativo
- $\hat{y}_i$  = valor previsto pelo modelo
- $n$  = número de dados no *dataset*

### 5.3.1 MSE: 6 Horas de Previsão

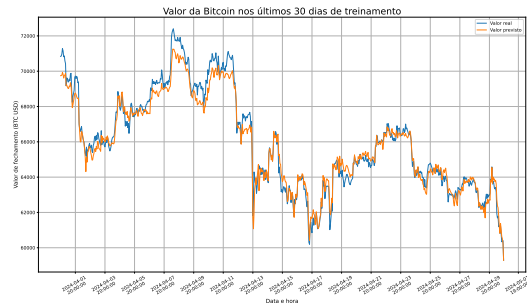
Utilizando a função de perda MSE e realizando a previsão dos valores de fechamento dos criptoativos de Bitcoin para as próximas 6 horas, o modelo foi treinado com 100L e 4S (Figura 14a) e 65L e 10S (Figura 14c).

Os gráficos apresentados nas Figuras 14b e 14d apresentam cada modelo treinado com a função MSE para a previsão de 12 horas com os dados dos últimos 30 dias de treinamento.

Figura 14 – MSE - Treinamento para 6 horas de previsão



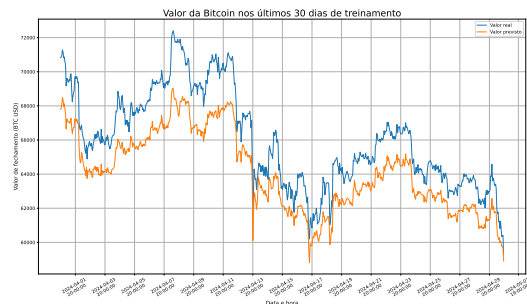
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



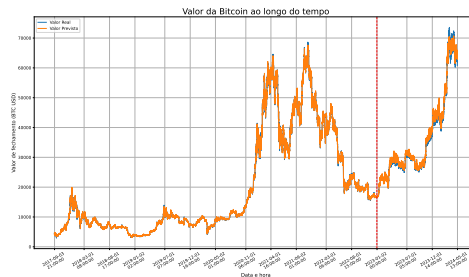
(d) 65L e 10S - Gráfico parcial

### 5.3.2 MSE: 12 Horas de Previsão

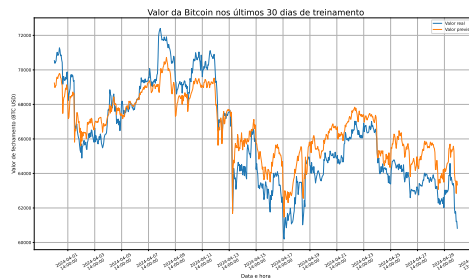
Em seguida, também foram treinadas duas diferentes redes neurais para prever os valores de fechamento dos criptoativos de Bitcoin para as próximas 12 horas, o modelo também foi treinado com 100L e 4S (Figura 15a) e 65L e 10S (Figura 15c).

Os gráficos com a representação dos dados parciais de treinamento, para os últimos 30 dias, com a função MSE são apresentados nas Figuras 15b e 15d.

Figura 15 – MSE - Treinamento para 12 horas de previsão



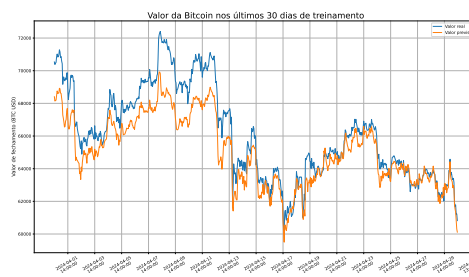
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



(d) 65L e 10S - Gráfico parcial

### 5.3.3 MSE: 24 Horas de Previsão

Assim como realizado para a função de perda MAE, no caso da função MSE também foram treinados modelos para a previsão dos valores de fechamento dos criptoativos de Bitcoin para as próximas 24 horas considerando 100L e 4S (Figura 16a) e 65L e 10S (Figura 16c).

As Figuras 16b e 16d apresentam a comparação entre os dados reais e os dados previstos pelo modelo durante a fase de treinamento para os últimos 30 dias de informações presente no *dataset*.

### 5.3.4 MSE: Convergência da Função de Perda

As Figuras 17a e 17b mostram a convergência dos modelos durante as fases de treinamento e teste, respectivamente, utilizando a função de perda MSE. De modo geral, observa-se uma redução e estabilização da perda em todos os casos. No entanto, no gráfico da fase de teste, notam-se algumas discrepâncias, especialmente ao prever os valores para as próximas 24 horas.

## 5.4 FUNÇÃO DE PERDA HUBER

O próximo caso analisado foi o treinamento da LSTM utilizando a função de perda Huber. Essa função é uma combinação entre as funções MSE e MAE e é geralmente utilizada para lidar com *datasets* que possuem dados discrepantes. Ela é calculada conforme a Equação (3).

$$J_{Huber} = \frac{1}{n} \sum_{i=1}^n \begin{cases} \frac{1}{2} (y_i - \hat{y}_i)^2, & |y_i - \hat{y}_i| \leq \delta \\ \delta (|y_i - \hat{y}_i| - \frac{1}{2}\delta), & |y_i - \hat{y}_i| > \delta \end{cases} \quad (3)$$

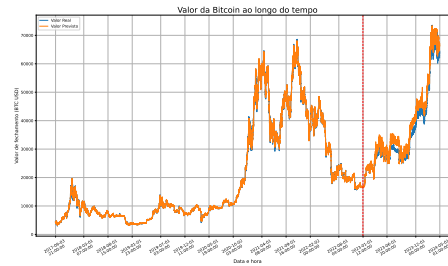
Onde:

- $J_{Huber}$  = resultado da função de perda Huber
- $y_i$  = valor real de fechamento do ativo
- $\hat{y}_i$  = valor previsto pelo modelo
- $n$  = número de dados no *dataset*
- $\delta$  = Hiperparâmetro que determina o ponto de transição da função entre quadrática e linear (Nesse trabalho  $\delta = 1$ )

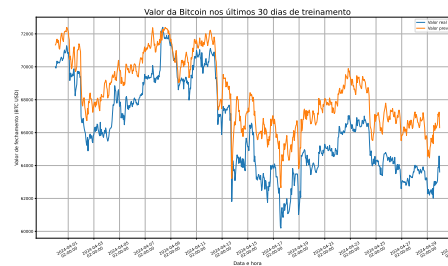
### 5.4.1 Huber: 6 Horas de Previsão

Utilizando a função de perda Huber, o modelo foi treinado para prever os valores de fechamento dos criptoativos de Bitcoin para as próximas 6 horas. O treinamento envolveu

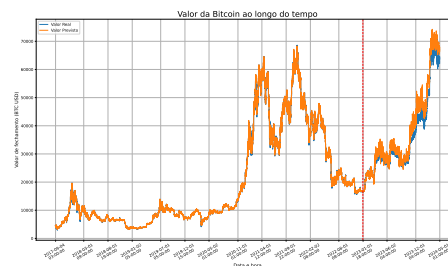
Figura 16 – MSE - Treinamento para 24 horas de previsão



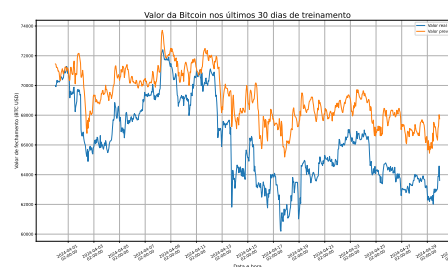
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S

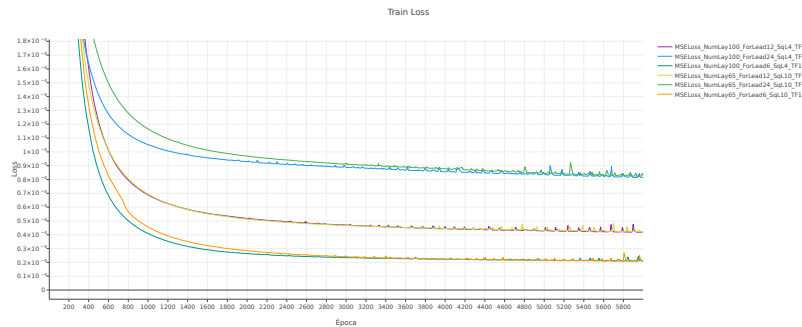


(d) 65L e 10S - Gráfico parcial

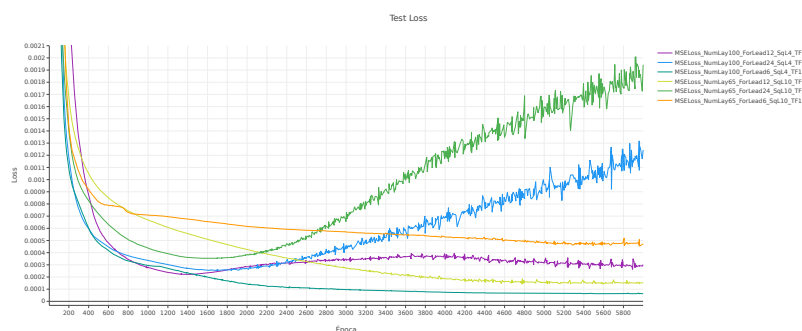
os dois casos treinados para as funções de perda MAE e MSE: 100L e 4S (Figura 18a) e 65L e 10S (Figura 18c).

As Figuras 18b e 18d apresentam uma comparação entre os dados reais e previstos para os últimos 30 dias de treinamento utilizando a função Huber e com a previsão para as próximas 6 horas.

Figura 17 – Convergência da função MSE



(a) Fase de treinamento



(b) Fase de teste.

#### 5.4.2 Huber: 12 Horas de Previsão

Também foram treinadas duas diferentes redes neurais para prever os valores de fechamento dos criptoativos de Bitcoin para as próximas 12 horas utilizando a função de perda Huber, o modelo também foi treinado com 100L e 4S (Figura 19a) e 65L e 10S (Figura 19c).

Os gráficos com a representação dos dados parciais de treinamento, para os últimos 30 dias, com a função Huber são apresentados nas Figuras 19b e 19c.

#### 5.4.3 Huber: 24 Horas de Previsão

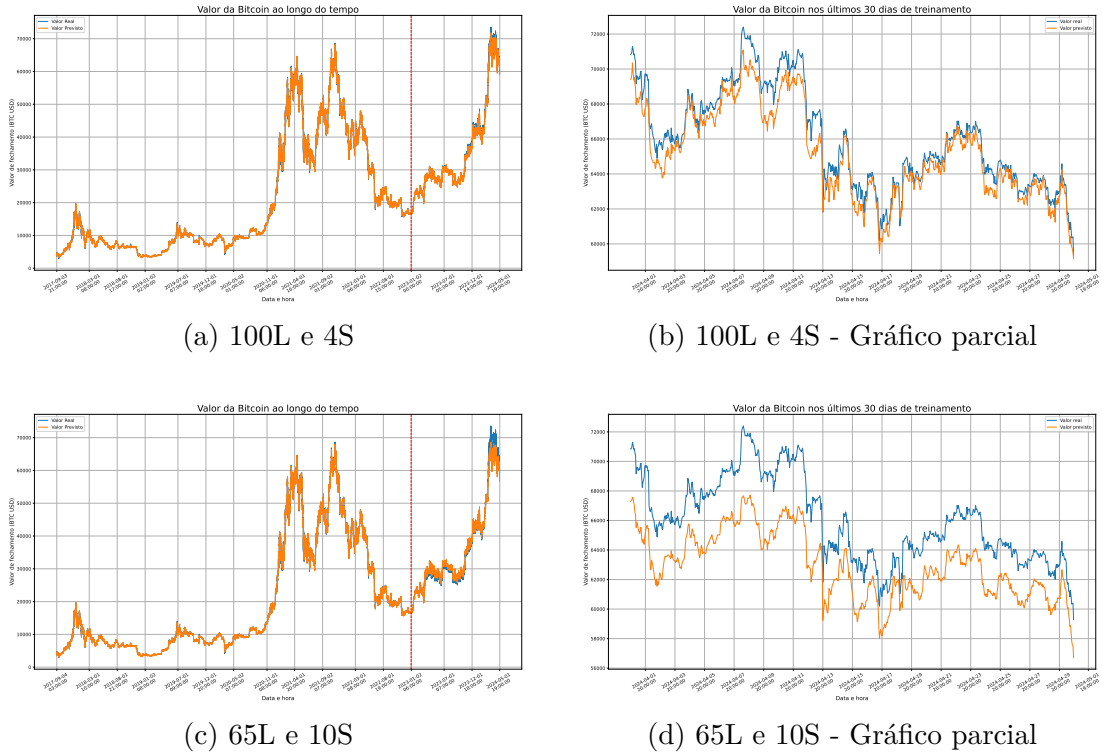
O último cenário testado com a função de perda Huber foi o treinamento do modelo para a previsão das próximas 24 horas utilizando os mesmos parâmetros mencionados anteriormente: 100L e 4S (Figura 20a) e 65L e 10S (Figura 20c).

Os gráficos do treinamento parcial desse cenário evidenciando os últimos 30 dias de dados presente no *dataset* são apresentados nas Figuras 20b e 20d.

#### 5.4.4 Huber: Convergência da Função de Perda

A avaliação da convergência dos cenários utilizando a função de perda Huber demonstra que durante a fase de treinamento (Figura 21a) o modelo apresenta pequenas

Figura 18 – Huber - Treinamento para 6 horas de previsão



oscilações em altas quantidades de épocas, podendo sinalizar a necessidade de ajuste da taxa de aprendizado. Já durante a fase de teste, a Figura 21b indica que esse cenário apresenta resultados semelhantes àqueles evidenciados na utilização da função de perda MSE: aumento da taxa de perda com muita oscilação para prever os valores para as próximas 24 horas.

## 5.5 FUNÇÃO DE PERDA SMOOTH L1

Por fim, a última função de perda utilizada na criação de cenários para o treinamento da LSTM foi a Smooth L1. Essa função também é uma combinação entre as funções MSE e MAE e é geralmente utilizada para evitar casos de gradiente explosivo. Ela é calculada conforme a Equação (4).

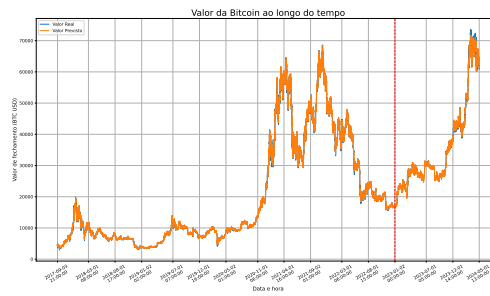
$$J_{SmoothL1} = \frac{1}{n} \sum_{i=1}^n \begin{cases} \frac{1}{2} (y_i - \hat{y}_i)^2 \frac{1}{\beta}, & |y_i - \hat{y}_i| \leq \beta \\ |y_i - \hat{y}_i| - \frac{1}{2}\beta & |y_i - \hat{y}_i| > \beta \end{cases} \quad (4)$$

Onde:

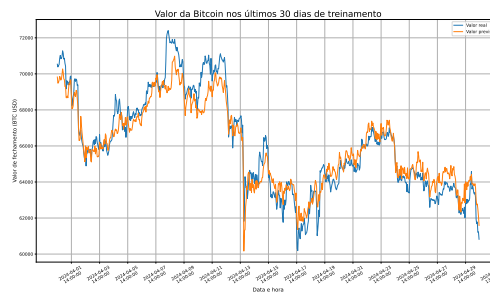
- $J_{SmoothL1}$  = resultado da função de perda Smooth L1
- $y_i$  = valor real de fechamento do ativo
- $\hat{y}_i$  = valor previsto pelo modelo
- $n$  = número de dados no *dataset*



Figura 19 – Huber - Treinamento para 12 horas de previsão



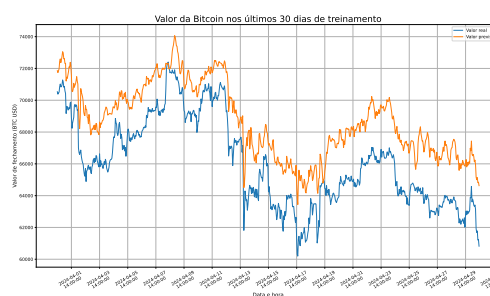
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



(d) 65L e 10S - Gráfico parcial

- $\beta$  = Hiperparâmetro de ajuste da função (Nesse trabalho  $\beta = 1$ )

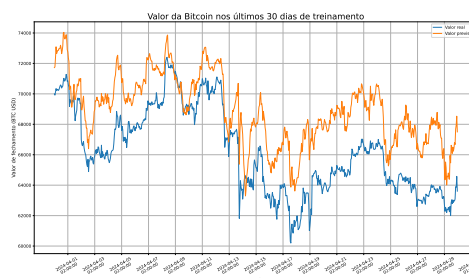
### 5.5.1 Smooth L1: 6 Horas de Previsão

Utilizando a função de perda Smooth L1 e prevendo os valores de fechamento dos criptoativos de Bitcoin para as próximas 6 horas, o modelo foi treinado com 100L e 4S

Figura 20 – Huber - Treinamento para 24 horas de previsão



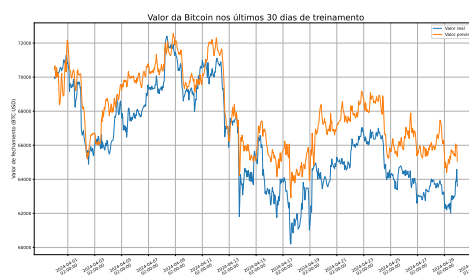
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S

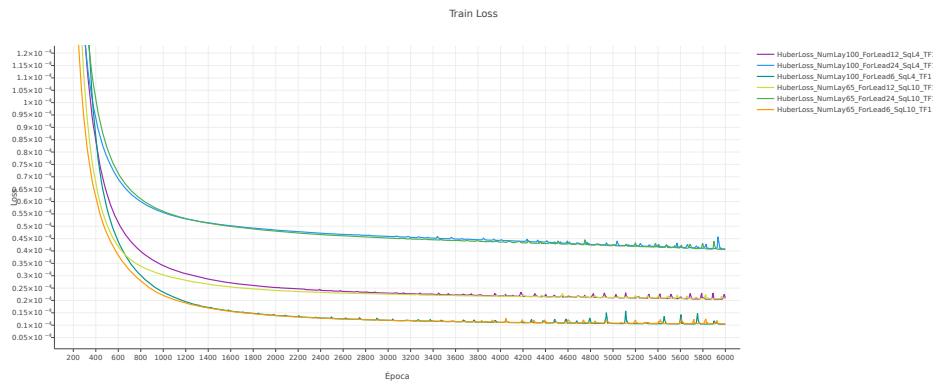


(d) 65L e 10S - Gráfico parcial

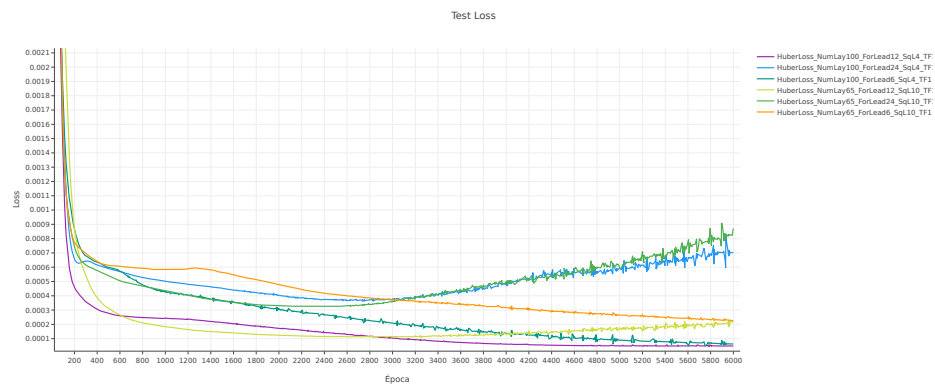
(Figura 22a) e 65L e 10S (Figura 22c).

Os gráficos destacando a representação dos dados parciais de treinamento, para os últimos 30 dias, com a função Smooth L1 são apresentados nas Figuras 22b e 22d.

Figura 21 – Convergência da função Huber



(a) Fase de treinamento



(b) Fase de teste.

### 5.5.2 Smooth L1: 12 Horas de Previsão

Também foram treinados modelos da rede neural criada para prever os valores de fechamento dos criptoativos de Bitcoin para as próximas 12 horas utilizando a função Smooth L1 utilizando os parâmetros: 100L e 4S (Figura 23a) e 65L e 10S (Figura 23c).

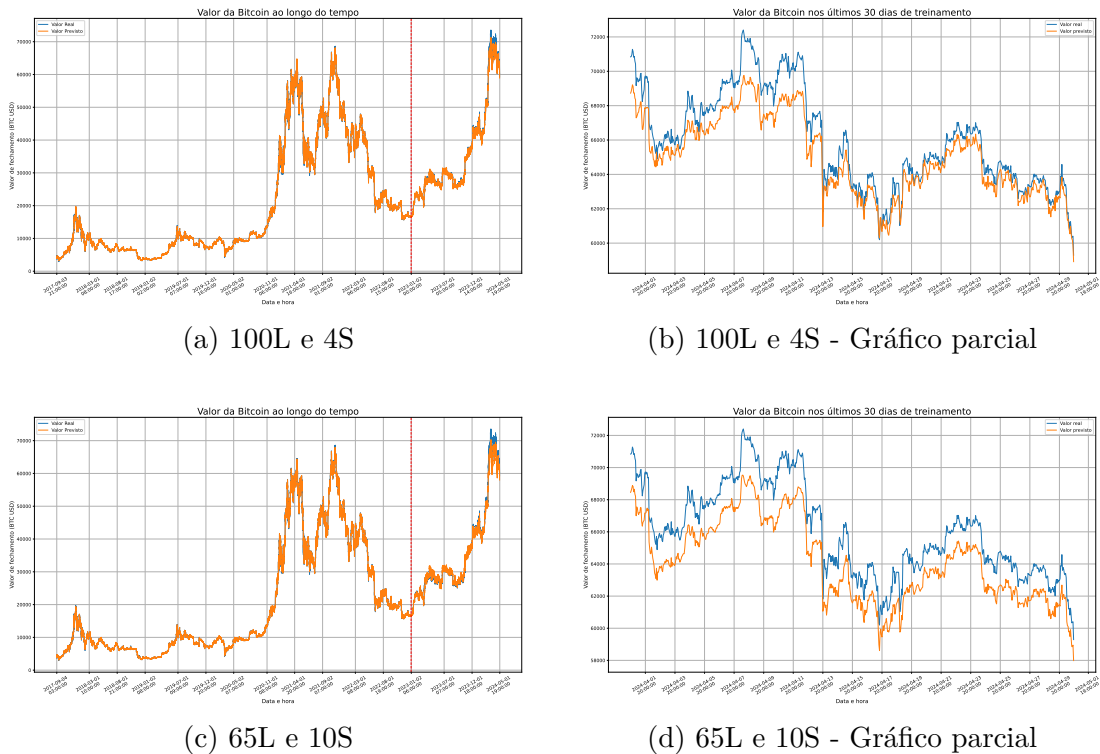
Assim como realizado para o período de 6 horas de previsão, foram criados gráficos demonstrando cada modelo treinado para a previsão de 12 horas com a função Smooth L1 evidenciando os dados dos últimos 30 dias de treinamento, apresentado nas Figuras 23b e 23d.

### 5.5.3 Smooth L1: 24 Horas de Previsão

O último caso treinado foi para a tentativa de previsão das próximas 24 horas de valores de fechamento da moeda analisada com a utilização da função Smooth L1. A Figura 24a apresenta os valores reais e previstos durante o treinamento utilizando 100L e 4S. Para 65L e 10S os resultados obtidos durante o treinamento são apresentados na Figura 24c.

Os gráficos demonstrando cada modelo treinado para a previsão de 24 horas com a

Figura 22 – Smooth L1 - Treinamento para 6 horas de previsão



função Smooth L1 evidenciando os dados dos últimos 30 dias de treinamento é apresentado nas Figuras 24b e 24d.

#### 5.5.4 Smooth L1: Convergência da Função de Perda

Assim como exibido para a função Huber, o gráfico da função Smooth L1 também demonstra que durante a fase de treinamento (Figura 25a) o modelo apresenta pequenas oscilações em altas quantidades de épocas, podendo sinalizar a necessidade de ajuste da taxa de aprendizado. Entretanto, durante a fase de teste, a Figura 25b indica que esse cenário exibe diversos pontos de oscilações e convergência apenas para a previsão das 6 horas seguintes. Essas informações indicam que os hiperparâmetros necessitam de ajuste e que essa função requer maior quantidade de épocas para ser capaz de se estabilizar e realizar um treinamento de melhor qualidade.

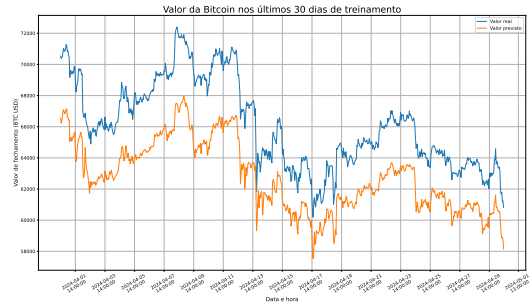
#### 5.6 VALORES PREVISTOS PELO MODELO E VALIDAÇÃO

Além da avaliação da convergência das funções de perda, outras maneiras utilizadas para a determinação da qualidade do modelo treinado são os valores obtidos através das métricas de regressão  $R^2$ , determinado pelas Equações (5) e (7), e MSE, definido pela Equação (6).

Figura 23 – Smooth L1 - Treinamento para 12 horas de previsão



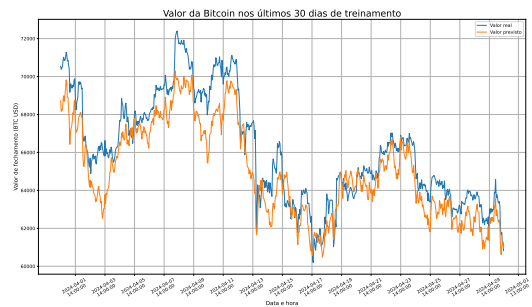
(a) 100L e 4S



(b) 100L e 4S - Gráfico parcial



(c) 65L e 10S



(d) 65L e 10S - Gráfico parcial

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (5)$$

$$MSE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

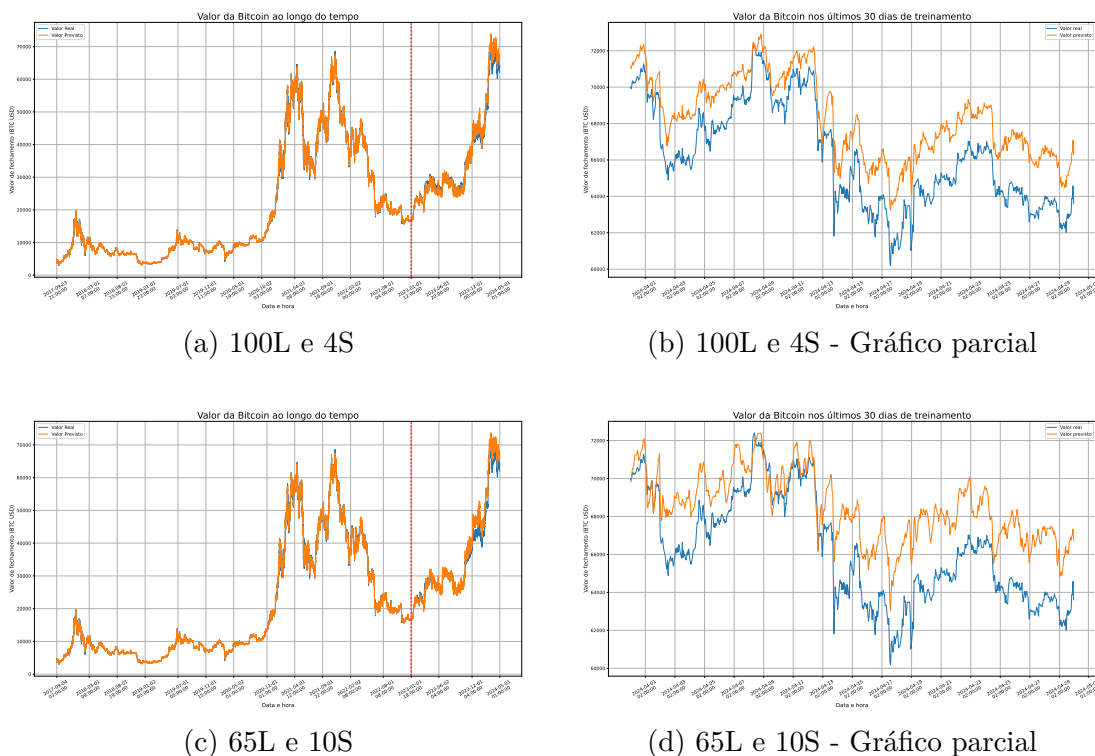
Onde:

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^n y_i \quad (7)$$

- $y_i$  = valor real de fechamento do ativo
- $\hat{y}_i$  = valor previsto pelo modelo
- $n$  = número de dados no *dataset*

Na Tabela 2 são apresentados os valores dessas métricas para cada um dos 24 cenários para os quais foram criados, treinados e testados modelos. Como é possível notar, todos os modelos registraram valores satisfatórios e semelhantes de  $R^2$ , se aproximando de 1 em todos os casos treinados. Entretanto, a partir dos resultados obtidos para a métrica MSE é possível notar que de maneira geral, ao utilizar uma quantidade de camadas menor, priorizando a quantidade de dados em sequência a serem colocados na entrada, o modelo

Figura 24 – Smooth L1 - Treinamento para 24 horas de previsão



desenvolvido possui dificuldades prevendo valores para um futuro distante, nesse caso 24 horas.

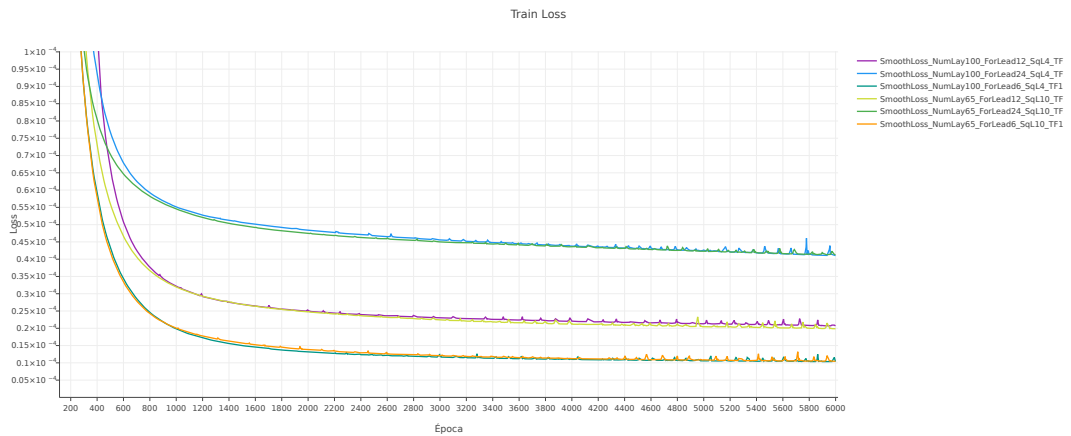
De maneira semelhante, esses dados também evidenciam uma melhor qualidade para a previsão a curto prazo (6 horas), com destaque para a utilização das funções MSE e Smooth L1 com os parâmetros de 100L e 4S.

Com o intuito de realizar a validação dos modelos treinados, foram realizadas as previsões dos valores de fechamento da Bitcoin para cada período treinado (6, 12 e 24 horas) utilizando informações não inseridas durante as fases de treinamento e teste e separadas especificamente para a fase de validação. Os resultados obtidos por cada função e conjunto de parâmetros foram inseridos em um gráfico juntamente com os valores reais.

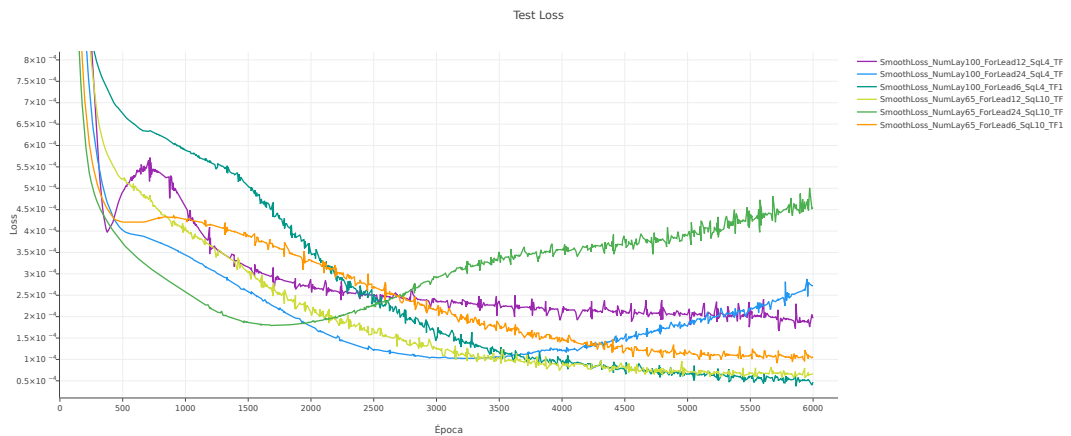
As Figuras 26a e 26b apresentam os valores previstos para o período de 6 horas utilizando modelos com 100L e 4S e 65L e 10S, respectivamente. Em ambos casos os resultados que mais se aproximaram dos valores reais foram aqueles obtidos com a utilização da função MAE. Além disso, quando comparado os parâmetros utilizados, notou-se melhor desempenho com 100L e 4S.

No caso de previsão para o período de 12 horas, os valores previstos encontram-se nas Figuras 27a e 27b. Como é possível notar, em ambos casos o pior desempenho foi do modelo utilizando a função de perda Huber, entretanto com a utilização de 100L e 4S os melhores resultados variam entre as funções MAE e Smooth L1, mas utilizando 65L e 10S a função MSE apresentou o melhor desempenho.

Figura 25 – Convergência da função Smooth L1



(a) Fase de treinamento



(b) Fase de teste.

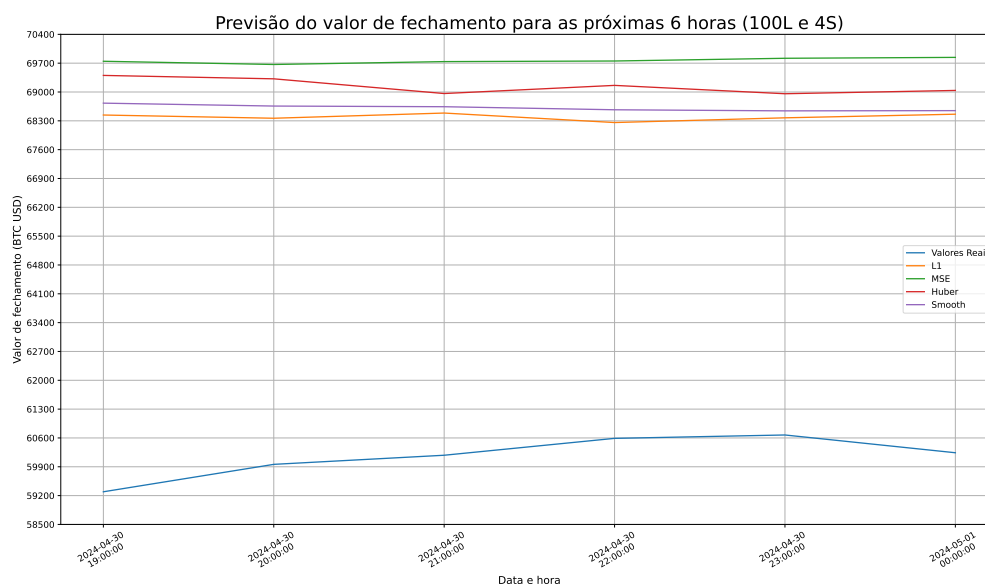
Por fim, os resultados obtidos para a previsão do período de 24 horas (Figuras 28a e 28b) são aqueles que apresentam maior diferença entre os valores reais e obtidos através da LSTM. Em ambos cenários analisados a função de perda que apresentou o melhor desempenho foi a MAE e utilizando 100L e 4S é possível notar maior disparidade nos resultados de cada função.

Tabela 2 – Cenários testados

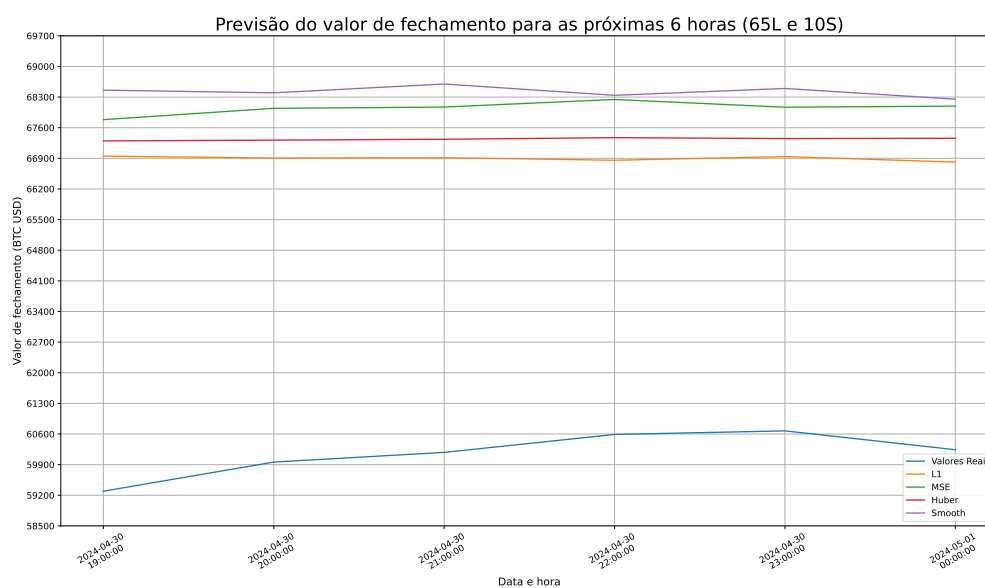
Cenário	Função de perda	Tempo de Previsão	Parâmetros da rede	$r^2$	MSE
1	MAE	6h	100L e 4S	$99,9177112 \times 10^{-2}$	$5,0445629 \times 10^{-5}$
2			65L e 10S	$99,8905999 \times 10^{-2}$	$6,7065048 \times 10^{-5}$
3		12h	100L e 4S	$99,8938837 \times 10^{-2}$	$6,5028369 \times 10^{-5}$
4			65L e 10S	$99,9730272 \times 10^{-2}$	$1,6528849 \times 10^{-5}$
5		24h	100L e 4S	$99,9823539 \times 10^{-2}$	$1,0804278 \times 10^{-5}$
6			65L e 10S	$99,7426720 \times 10^{-2}$	$15,7554998 \times 10^{-5}$
7	MSE	6h	100L e 4S	$99,9879201 \times 10^{-2}$	$0,7405363 \times 10^{-5}$
8			65L e 10S	$99,8511152 \times 10^{-2}$	$9,1270158 \times 10^{-5}$
9		12h	100L e 4S	$99,9198552 \times 10^{-2}$	$4,9112954 \times 10^{-5}$
10			65L e 10S	$99,9623867 \times 10^{-2}$	$2,3049275 \times 10^{-5}$
11		24h	100L e 4S	$99,6081996 \times 10^{-2}$	$23,9891160 \times 10^{-5}$
12			65L e 10S	$99,4422198 \times 10^{-2}$	$34,1513805 \times 10^{-5}$
13	Huber	6h	100L e 4S	$99,9650083 \times 10^{-2}$	$2,1450984 \times 10^{-5}$
14			65L e 10S	$99,8568458 \times 10^{-2}$	$8,7757173 \times 10^{-5}$
15		12h	100L e 4S	$99,9773129 \times 10^{-2}$	$1,3902726 \times 10^{-5}$
16			65L e 10S	$99,8678807 \times 10^{-2}$	$8,0962221 \times 10^{-5}$
17		24h	100L e 4S	$99,5502369 \times 10^{-2}$	$27,5380560 \times 10^{-5}$
18			65L e 10S	$99,4022479 \times 10^{-2}$	$36,5987449 \times 10^{-5}$
19	Smooth L1	6h	100L e 4S	$99,9743729 \times 10^{-2}$	$1,5710222 \times 10^{-5}$
20			65L e 10S	$99,9409697 \times 10^{-2}$	$3,6187044 \times 10^{-5}$
21		12h	100L e 4S	$99,8830109 \times 10^{-2}$	$7,1691233 \times 10^{-5}$
22			65L e 10S	$99,9650820 \times 10^{-2}$	$2,1397616 \times 10^{-5}$
23		24h	100L e 4S	$99,8387377 \times 10^{-2}$	$9,8737510 \times 10^{-5}$
24			65L e 10S	$99,7006198 \times 10^{-2}$	$18,3302429 \times 10^{-5}$



Figura 26 – Valores previstos pelo modelo para as próximas 6 horas.

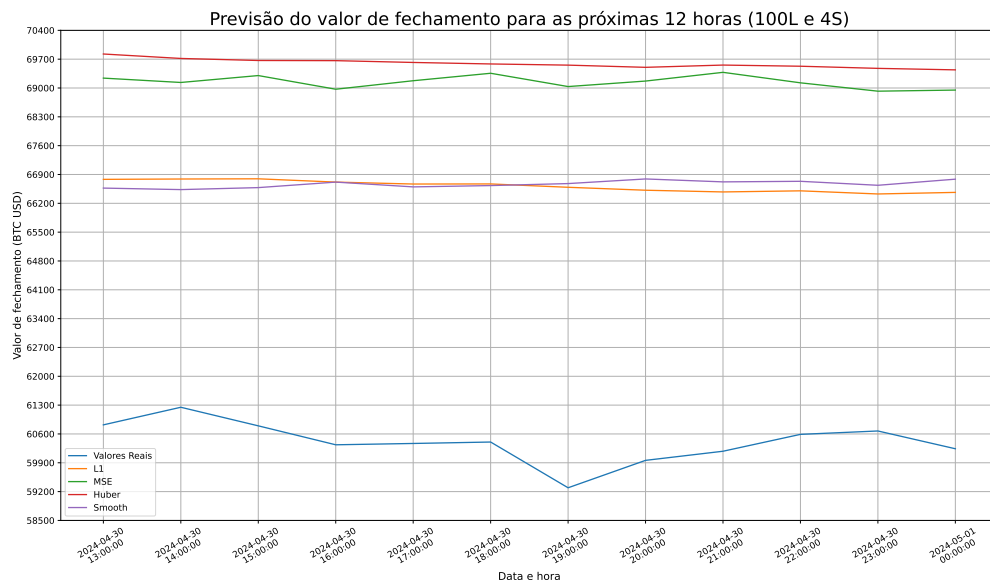


(a) 100L e 4S.

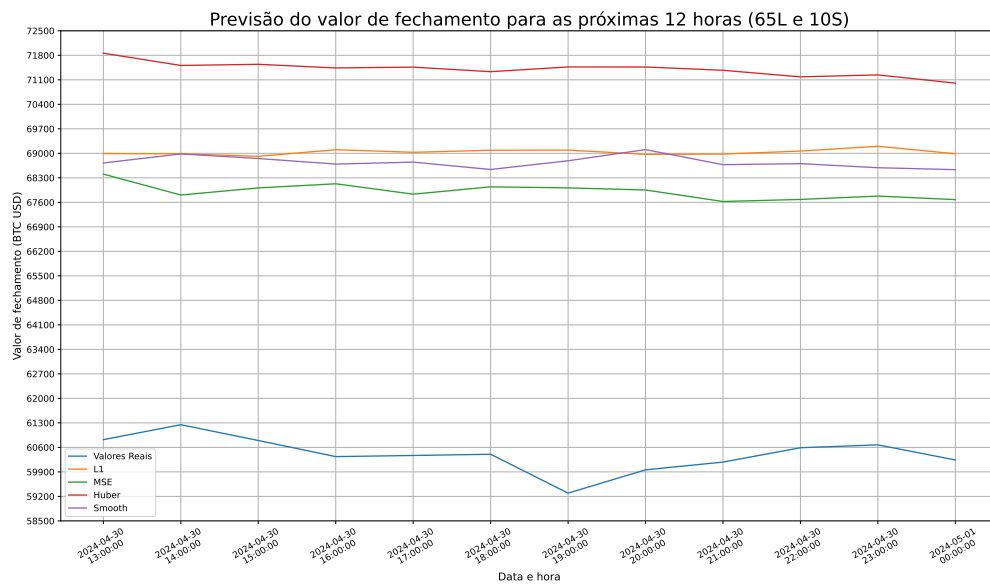


(b) 65L e 10S.

Figura 27 – Valores previstos pelo modelo para as próximas 12 horas.

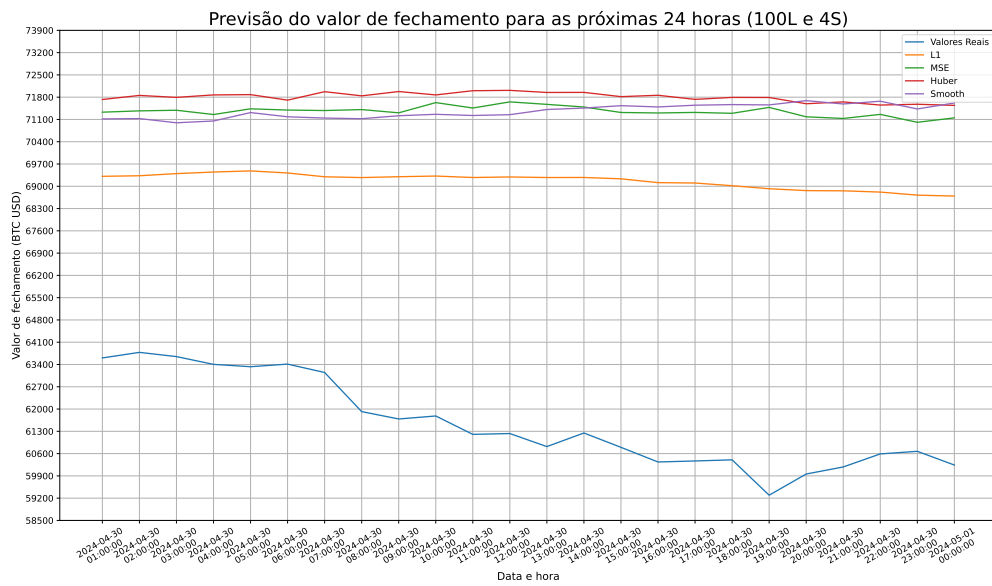


(a) 100L e 4S.

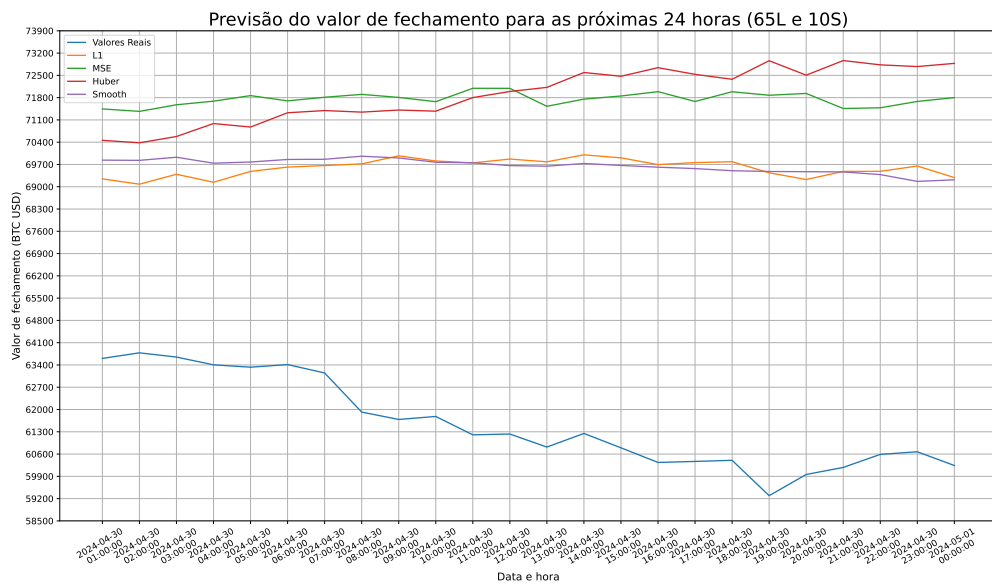


(b) 65L e 10S.

Figura 28 – Valores previstos pelo modelo para as próximas 24 horas.



(a) 100L e 4S.



(b) 65L e 10S.

## 6 CONCLUSÃO

Com base nas informações apresentadas, é possível afirmar que os resultados obtidos foram coerentes com o esperado. A arquitetura de rede neural do tipo LSTM, criada e treinada, foi capaz de prever aproximadamente os valores de fechamento dos criptoativos de Bitcoin. Além disso, os cenários testados permitiram determinar as melhores condições para a função de perda, período de predição, quantidade de camadas e sequência de dados inseridas no modelo.

Em relação ao período de predição, os resultados obtidos condizem com a literatura de redes neurais no sentido de que quanto mais distante é o prazo futuro que se deseja prever, maiores as taxas de erro e mais distantes os resultados obtidos são dos valores reais. Assim, a predição mais aproximada dos valores foi a realizada na tentativa de saber os dados de fechamento da moeda nas próximas 6 horas.

Considerando a quantidade de camadas e quantidade de sequência de dados inseridos na entrada do modelo, o melhor desempenho geral foi obtido utilizando 100 camadas e 4 de sequência. Esse comportamento é observado, pois a quantidade de vezes em que cada subconjunto de dados é processada é diretamente relacionada à quantidade de camadas. Além disso, foi possível notar que apesar do tamanho da sequência de dados influenciar nos resultados, o peso dessa influência é significativamente menor que a interferência causada pela quantidade de camadas, permitindo concluir que para casos que possuem limitação de *hardware*, como esse projeto, é mais favorável a priorização da quantidade de camadas.

Tendo em vista a função de perda aplicada durante o treinamento para o ajuste do modelo, os melhores resultados foram obtidos através da função MAE, com evidência para o modelo utilizando 65 camadas e sequência de dados igual à 10. Entretanto, ao analisar os gráficos de convergência das funções de perda é possível notar que durante a etapa de teste todas as funções utilizadas apresentaram grandes oscilações ou crescimento do valor de perda com o aumento da quantidade de épocas treinadas. Dessa forma, é possível concluir que todos os modelos treinados possuem certo ajuste fino de parâmetros a ser feito para que seu desempenho seja melhorado, criando a possibilidade de que outras funções, além da MAE, apresentem valores de predição mais precisos.

Por fim, é importante destacar a dificuldade do modelo em produzir resultados exatamente iguais aos valores reais, o que impede que decisões no mercado de ativos sejam tomadas exclusivamente com base nos resultados da rede neural. Entretanto, o modelo desenvolvido é capaz de prever certas oscilações nos valores, possibilitando uma visão ampla que pode auxiliar a previsão de quando os valores no mercado vão crescer ou decrescer.

Essas discrepâncias entre os valores reais e os valores previstos podem ocorrer por diversos fatores, tais como ajuste dos hiperparâmetros da rede neural ou até mesmo pela inexistência de arquiteturas capazes de prever situações tão incertas quanto o mercado

de ativos. Em projetos desenvolvidos futuramente recomenda-se a utilização de *hardware* como memórias RAM (*Random Access Memory*) e principalmente GPU com melhores capacidades de processamento. Dessa forma é possível realizar o ajuste fino dos parâmetros utilizados na LSTM e também aumentar as quantidades de camadas e épocas utilizadas durante o treinamento, fatores que melhoram de maneira significativa o desempenho do modelo.

## REFERÊNCIAS

- AGARWAL, Shiva; MUPPALANENI, Naresh Babu. Stock Market Price and Cryptocurrency Price Prediction. *In: 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. [S.l.: s.n.], 2022. P. 1–6.
- ALVAREZ-RAMIREZ, J.; RODRIGUEZ, E.; IBARRA-VALDEZ, C. Long-range correlations and asymmetry in the Bitcoin market. **Physica A: Statistical Mechanics and its Applications**, v. 492, p. 948–955, 2018. ISSN 0378-4371.
- ANSPER, Arne; BULDAS, Ahto; SAAREPERA, Märt; WILLEMSON, Jan. Improving the Availability of Time-Stamping Services. *In: VARADHARAJAN, Vijay; MU, Yi (Ed.)*. **Information Security and Privacy**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001. P. 360–375.
- BASHIR, Imran. **Mastering Blockchain**. 2. ed. Birmingham, England: Packt Publishing, 2018.
- BIRRER, Thomas K.; AMSTUTZ, Dennis; WENGER, Patrick. **Decentralized Finance: From Core Concepts to DeFi Protocols for Financial Transactions**. [S.l.]: Springer Fachmedien Wiesbaden, 2023. ISBN 9783658398729.
- BITCOIN.ORG. **Block Chain**. Acessado em 30 de março de 2024. 2023. Disponível em: [https://developer.bitcoin.org/devguide/block\\_chain.html#id1](https://developer.bitcoin.org/devguide/block_chain.html#id1).
- BRASIL, Banco Central do. **Estatísticas de Meios de Pagamentos**. Acessado em 30 de março de 2024. 2022. Disponível em: <https://www.bcb.gov.br/estatisticas/spbadendos?ano=2022>.
- BUSA, Tee. **BTC Market Cycle | Repetitive and Predictable Market Cycle**. [S.l.], 2024. Disponível em: <https://www.tradingview.com/chart/BLX/IUYIOsNk-BTC-Market-Cycle-Repetitive-and-Predictable-Market-Cycle/>. Acesso em: 1 mai. 2024.
- CHAUM, David. Blind Signatures for Untraceable Payments. *In: CHAUM, David; RIVEST, Ronald L.; SHERMAN, Alan T. (Ed.)*. **Advances in Cryptology: Proceedings of Crypto 82**. Boston, MA: Springer US, 1983. P. 199–203.

CHOPRA, Ritika; SHARMA, Gagan Deep. Application of Artificial Intelligence in Stock Market Forecasting: A Critique, Review, and Research Agenda. **Journal of Risk and Financial Management**, v. 14, n. 11, 2021. ISSN 1911-8074.

COOLIDGE, J. L. The Gambler's Ruin. **The Annals of Mathematics**, JSTOR, v. 10, n. 4, p. 181, jul. 1909. ISSN 0003-486X.

DAI, Wei. **B-money**. [S.l.: s.n.], 1998.

DE FARIA, E.L.; ALBUQUERQUE, Marcelo P.; GONZALEZ, J.L.; CAVALCANTE, J.T.P.; ALBUQUERQUE, Marcio P. Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. **Expert Systems with Applications**, v. 36, n. 10, p. 12506–12509, 2009. ISSN 0957-4174.

DE OLIVEIRA, Fagner A.; NOBRE, Cristiane N.; ZÁRATE, Luis E. Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index – Case study of PETR4, Petrobras, Brazil. **Expert Systems with Applications**, v. 40, n. 18, p. 7596–7606, 2013. ISSN 0957-4174.

ESFAHANIPOUR, Akbar; AGHAMIRI, Werya. Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis. **Expert Systems with Applications**, v. 37, n. 7, p. 4742–4748, 2010. ISSN 0957-4174.

FAREEZ, MMM; THANGARAJAH, Vinothraj; SAABITH, Sayeth. Popular Python Libraries and Their Applications Domains, nov. 2020.

FARELL, Ryan. An analysis of the cryptocurrency industry. **Wharton Research Scholars**, v. 130, p. 1–23, 2015.

FIESLER, E. Neural network classification and formalization. **Computer Standards & Interfaces**, v. 16, n. 3, p. 231–239, 1994. ISSN 0920-5489.

HALLS-MOORE, Michael L. Advanced algorithmic trading. **QuantStart, United Kingdom**, 2017.

HARVEY, Campbell R.; DE VRIES, David; RAMACHANDRAN, Ashwin; SANTORO, Joey. **Defi and the future of Finance**. [S.l.]: Ascent Audio, 2021.

- HEARN, Mike; BROWN, Richard Gendal. Corda: A distributed ledger. **Corda Technical White Paper**, v. 2016, 2016.
- HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997.
- JAQUART, Patrick; DANN, David; WEINHARDT, Christof. Short-term bitcoin market prediction via machine learning. **The Journal of Finance and Data Science**, v. 7, p. 45–66, 2021. ISSN 2405-9188.
- LIU, Haojun; LUO, Xinbo; LIU, Hongrui; XIA, Xubo. Merkle Tree: A Fundamental Component of Blockchains. *In*: 2021 International Conference on Electronic Information Engineering and Computer Science (EIECS). [S.l.: s.n.], 2021. P. 556–561.
- MASIAK, Christian; BLOCK, Joern H; MASIAK, Tobias; NEUENKIRCH, Matthias; PIELEN, Katja N. **The market cycles of ICOs, bitcoin, and ether**. [S.l.]: SSRN, 2018.
- MOKHTARI, Sohrab; YEN, Kang K.; LIU, Jin. Effectiveness of Artificial Intelligence in Stock Market Prediction based on Machine Learning. arXiv, 2021.
- MORAES, Dalila de. **Modelagem de fraude em cartão de crédito**. [S.l.: s.n.], 2008.
- NAKAMOTO, Satoshi. Bitcoin: A peer-to-peer electronic cash system. **Decentralized business review**, 2008.
- NARAYANAN, Arvind; BONNEAU, Joseph; FELTEN, Edward; MILLER, Andrew; GOLDFEDER, Steven. **Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction**. [S.l.]: Princeton University Press, 2016. ISBN 978-0691171692.
- OVERFLOW, Stack. **2023 Developer Survey**. Acessado em 30 de março de 2024. 2023. Disponível em: <https://survey.stackoverflow.co/2023/#technology>.
- RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence**. 3. ed. Upper Saddle River, NJ: Pearson, dez. 2009.



SERVICES, Amazon Web. **O que é uma rede neural? – Explicação sobre rede neural artificial – AWS** — [aws.amazon.com](https://aws.amazon.com). [S.l.: s.n.].

<https://aws.amazon.com/pt/what-is/neural-network/>. [Accessed 27-05-2024].

SICHE, Raúl; AGOSTINHO, Feni; ORTEGA, Enrique; ROMEIRO, Ademar. Índices versus indicadores: precisões conceituais na discussão da sustentabilidade de países. **Ambiente & Sociedade**, FapUNIFESP (SciELO), v. 10, n. 2, p. 137–148, dez. 2007. ISSN 1414-753X.

SRINATH, KR. Python—the fastest growing programming language. **International Research Journal of Engineering and Technology**, v. 4, n. 12, p. 354–357, 2017.

TABATABAEI, Mohammad Hossein; VITENBERG, Roman; VEERARAGAVAN, Narasimha Raghavan. Understanding blockchain: Definitions, architecture, design, and system comparison. **Computer Science Review**, v. 50, p. 100575, 2023. ISSN 1574-0137.

TRIGG, D. W.; LEACH, A. G. Exponential Smoothing with an Adaptive Response Rate. **Journal of the Operational Research Society**, Informa UK Limited, v. 18, n. 1, p. 53–59, mar. 1967. ISSN 1476-9360.

YAN, Shi. **Understanding LSTM and its diagrams**. [S.l.], 2016. Disponível em: <https://blog.mlreview.com/understanding-lstm-and-its-diagrams-37e2f46f1714>. Acesso em: 16 mar. 2024.

ZHU, Yuhan. Research on Digital Finance Based on Blockchain Technology. *In*: 2021 International Conference on Computer, Blockchain and Financial Development (CBFD). [S.l.: s.n.], 2021. P. 410–414.