

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE ENGENHARIA DE COMPUTAÇÃO

RODOLFO HELFENSTEIN

**Um modelo de Rede Neural de classificação
para predição da tendência do preço Bitcoin**

Monografia apresentada como requisito parcial
para a obtenção do grau de Bacharel em
Engenharia da Computação

Orientador: Prof^a. Dr^a. Renata Galante

Porto Alegre
2022

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões Mendes

Vice-Reitora: Prof^a. Patricia Helena Lucas Pranke

Pró-Reitor de Graduação: Prof^a. Cíntia Inês Boll

Diretora do Instituto de Informática: Prof^a. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Engenharia de Computação: Prof. Walter Fetter Lages

Bibliotecário-chefe do Instituto de Informática: Alexsander Borges Ribeiro

Bibliotecária-chefe da Escola de Engenharia: Rosane Beatriz Allegretti Borges

AGRADECIMENTOS

Agradeço, primeiramente, para a minha família. Especialmente para minha mãe, Nara Lúcia de Brito Helfenstein, meu pai, Wilson Frederico Helfenstein, e minha irmã, Bruna Helfenstein os quais sempre me apoiaram nas diversas dificuldades no caminho da graduação e sempre me motivaram a continuar e nunca desistir dos meus sonhos independente de quais forem. Seria impossível a conclusão desse trabalho sem o esforço deles para eu conseguir atingir meus objetivos.

Agradeço muito aos meu colegas de faculdades que se tornaram amigos e percorreram essa árdua caminhada da graduação junto comigo. Sempre me ajudaram e estavam juntos comigo nos momentos mais difíceis da faculdade. Tornaram muito mais fácil e divertido essa difícil tarefa de se formar

Agradeço, também, aos professores que tive durante minha jornada na graduação da UFRGS que me ensinaram muito sobre diversos tópicos diferentes da computação e da vida. Mesmo aqueles que às vezes faziam a jornada mais difícil, me ensinaram lições de resiliência e trabalho duro.

Agradeço especialmente à professora Renata de Matos Galante, orientadora deste trabalho, que demonstrou todos os dias empatia e preocupação não só comigo mas com todos os seus alunos e fez deste trabalho uma experiência agradável e menos assustadora.

Muito obrigado.

RESUMO

Os estudos sobre séries temporais financeiras vem aumentando cada vez mais ao longo dos anos principalmente com o uso de modelos de aprendizado de máquina para tentar prever preços e movimentações para extrair ao máximo as oportunidades que surgem nesse mercado. Há pouco tempo, uma nova classe de ativos foi criada e vem crescendo exponencialmente nos últimos anos: criptomoedas. Esses ativos são moedas virtuais que utilizam criptografia para manter a segurança das transações e tem como principal representante o Bitcoin. Por serem relativamente jovens, comparadas ao mercado tradicional como por exemplo *commodities*, ações, forex, existe uma volatilidade muito grande no preço devido à adoção por pessoas e instituições de todo o mundo desses ativos, mas que podem ser analisados da mesma forma. Neste trabalho, é proposto um modelo de rede neural de classificação utilizando células LSTM para predição da tendência no preço do Bitcoin utilizando como base de dados os valores do *candlestick* semanal junto com o volume de transação. São testadas configurações diferentes para o modelo, variando-se separadamente *timesteps*, *epochs* e a utilização de dados adicionais da macroeconomia. Após encontrar a configuração que obtém os melhores resultados de retorno de investimento e acurácia, o modelo é comparado a estratégias de negociação simples.

Palavras-chave: Rede neural. predição. LSTM. bitcoin. macroeconomia. séries temporais.

Neural network model applied in Bitcoin price trend prediction

ABSTRACT

The studies on financial time series have been increasing over the years, mainly with the use of machine learning models to try to predict prices and movements to extract the most of the opportunities that arise in this market. Not long ago, a new asset class was created and has been growing exponentially in recent years: cryptocurrencies. These assets are virtual currencies that use cryptography to maintain the security of transactions and their main representative is Bitcoin. Because they are relatively young, compared to the traditional market such as commodities, stocks, forex, there is a very high volatility in the price due to the adoption by people and institutions from all over the world of these assets, but they can be analyzed in the same way. In this work, a classification neural network model is proposed using LSTM cells to predict trends in Bitcoin price using weekly candlestick values along with transaction volume as a database. Different configurations for the model are tested, varying separately timesteps, epochs and using additional macroeconomic data. After finding the configuration that obtains the best results of return on investment and accuracy, the model is compared to other simple trading strategies.

Keywords: neural network, bitcoin, forecasting, time series.

LISTA DE FIGURAS

Figura 2.1 Representação de um <i>Candlestick</i> de alta e de baixa, esquerda e direita respectivamente.....	14
Figura 2.2 Série temporal do preço do Bitcoin em escala logarítmica	18
Figura 2.3 Representação de uma célula LSTM	20
Figura 4.1 Resumo do modelo LSTM	26
Figura 5.1 Metodologia do experimento	29
Figura 5.2 Exemplo de evolução do capital investido com ROI negativo e positivo	30
Figura 5.3 Resultados da Configuração Padrão	32
Figura 5.4 Resultados utilizando a métrica ROI	35
Figura 5.5 Resultados utilizando a métrica acurácia.....	36
Figura A.1 Índice histórico da NASDAQ	41
Figura A.2 Índice histórico da NYA	41
Figura A.3 Índice histórico do S&P500.....	42
Figura B.1 Preço histórico do ouro	43
Figura B.2 Preço histórico da prata.....	43

LISTA DE TABELAS

Tabela 2.1	Tabela de correlação	16
Tabela 3.1	Comparativo entre os trabalhos que utilizam redes neurais para predição de ativos financeiros	24
Tabela 5.1	Tabela resultados variando <i>timesteps</i>	33
Tabela 5.2	Tabela resultados variando <i>epochs</i>	33
Tabela 5.3	Tabela resultados variando dados de entradas	34

LISTA DE ABREVIATURAS E SIGLAS

LSTM	<i>Long Short-Term Memory</i>
RNN	<i>Recurent Neural Network</i>
ANN	<i>Artificial Neural Network</i>
GRU	<i>Gated Recurrent Unit</i>
ROI	<i>Return of Investment</i>
IFR	Índice de Força Relativa

SUMÁRIO

1 INTRODUÇÃO	10
2 CONCEITOS E TECNOLOGIAS UTILIZADAS	13
2.1 Mercado Financeiro	13
2.1.1 <i>Candlestick</i>	13
2.1.2 Negociação de Ativos	14
2.1.3 Indicadores técnicos	15
2.1.4 Indicadores Macroeconômicos	16
2.1.5 Séries Temporais	17
2.2 Redes neurais artificiais.....	18
2.3 Considerações Finais	20
3 TRABALHOS RELACIONADOS	21
3.1 Criptomoeadas	21
3.2 Mercado Tradicional.....	22
3.3 Análise Comparativa	23
4 MODELO LSTM DE CLASSIFICAÇÃO PARA PREDIÇÃO DA TENDÊN- CIA DO PREÇO DO BITCOIN.....	25
4.1 Arquitetura do Modelo Proposto.....	25
4.2 Recursos Computacionais	26
5 AVALIAÇÃO EXPERIMENTAL	27
5.1 Configuração dos Experimentos.....	27
5.1.1 Tratamento dos dados	27
5.1.2 Métricas.....	28
5.1.3 Metodologia dos Experimentos	29
5.1.4 Conjunto de dados.....	31
5.2 Descrição dos Experimentos	31
5.2.1 Configuração Padrão	31
5.2.2 <i>Timesteps</i>	32
5.2.3 <i>Epochs</i>	33
5.2.4 Adicionando dados da Macroeconomia.....	33
5.2.5 Comparação com outras estratégias.....	34
5.3 Considerações Finais	37
6 CONCLUSÃO	38
REFERÊNCIAS.....	39
APÊNDICE A — DADOS DA BOLSA AMERICANA	41
APÊNDICE B — <i>COMMODITIES</i>	43

1 INTRODUÇÃO

Bitcoin é a maior criptomoeda em capitalização de mercado (coinmarketcap.com, 2022), um ativo completamente digital que permite transações *peer-to-peer* as quais são encriptadas para garantir a segurança da rede. Utiliza *blockchain*, uma estrutura de dados que é formada por uma cadeia de registros denominados blocos (YAGA et al., 2019), conectados através da criptografia para armazenar todas as movimentações de saldo entre as carteiras, muito semelhante ao livro-razão de uma contabilidade de uma empresa e, uma vez que alguma transação é registrada na *blockchain*, ela não pode ser desfeita. Todas as transações são de domínio público e são processadas de forma descentralizada por mineradores de todo o mundo através da Internet, ou seja, não existe um órgão regulador centralizado (NAKAMOTO, 2009). O Bitcoin já chegou a ter mais de 1 trilhão de dólares de valor de mercado e hoje se encontra na 8ª posição no ranking dessa característica estando acima de diversas empresas multinacionais como a Visa e o Facebook (Companies Marketcap, 2022). O mercado de criptomoedas está sempre aberto, diferentemente do mercado tradicional, por exemplo, os ativos da IBOVESPA só podem ser negociados das 10:00 às 17:00 horas em dias comerciais.

A representação do preço do Bitcoin, assim como de outros ativos financeiros, é feita através de séries temporais. Séries temporais (KIM; CHOI; KANG, 2019) são dados coletados em intervalos de tempo regulares com dependência de ordem. Séries temporais vêm se tornando cada vez mais populares quando o assunto se trata de como prevê-las com base no histórico do objeto estudado, o que não é nada simples visto que a predição, nesse caso, está fora do intervalo dos dados conhecidos. As áreas onde essas séries são aplicadas são diversas, como, por exemplo, demografia, epidemiologia, meteorologia entre outros. Porém, o que mais se destaca são os estudos de predição de ativos financeiros com o intuito de comprá-los ou vendê-los em momentos oportunos para conseguir um retorno maior do que o capital investido, ou seja, lucrar.

A tarefa de predição não é simples de ser feita, visto que muitos agentes econômicos entram no mercado com o mesmo objetivo fazendo com que o volume de transações diário seja muito grande e, conseqüentemente, aumentando a alta volatilidade. Esse fato se agrava quando se trata de criptomoedas. Além disso, existem muitas variáveis que têm influência nesse mercado, como sazonalidade e macroeconomia, tornando-o um problema não-linear. Por ser um mercado relativamente muito novo e os estudos preditivos ainda estarem nos seus estágios iniciais, a literatura sobre a predição de criptomoedas acabam

por serem limitadas e com resultados preliminares, quase que não seguindo as normas padrões acadêmicas e falta de transparência na apresentação do modelo para poder replicá-lo utilizando os mesmos hiperparâmetros, *benchmark* e modelos de treino. Além desse fator, ainda se tem pouca orientação sobre a natureza da formação do preço do Bitcoin.

Diversos tipos de abordagens já foram utilizados para tentar prever séries temporais financeiras como (MOGHAR; HAMICHE, 2020) onde é proposto um modelo de LSTM para predição de ativos da bolsa americana e (SAAD; MOHAISEN, 2018) onde foram propostos modelos de regressão múltipla. Comumente são utilizados modelos de aprendizado de máquina no intuito de encontrar padrões na flutuação dos preços desses ativos como regressão linear, floresta aleatória, gradiente descendente, máquina de vetores de suporte, redes neurais, entre outros. Especificamente, as redes neurais são as que têm melhor desempenho na predição, visto que elas têm um desempenho bom para problemas não-lineares em comparação a outros modelos. Além disso, as redes neurais podem receber vários tipos de entrada para o cálculo da saída possibilitando que diversos dados diferentes sejam usados para o treinamento do modelo, além dos valores do próprio preço do Bitcoin.

O objetivo deste trabalho é propor um modelo de rede neural de classificação utilizando células LSTM que obtenha resultados satisfatórios em relação à acurácia e retorno do investimento para fazer a predição da tendência do preço do Bitcoin. O modelo é construído com 8 camadas, sendo quatro delas de LSTM, três de *dropout* e uma camada de *dense* e para seu treinamento e teste são utilizados como base de dados os valores do *candlestick* semanal junto com o volume de transação. São testadas configurações diferentes para o modelo, variando-se separadamente *timesteps*, *epochs* e a utilização de dados adicionais da macroeconomia para verificar qual obtém os melhores resultados de acurácia e retorno de investimento. Após encontrar a melhor configuração, são feitas comparações com outras estratégias de negociação. Obteve-se uma acurácia de 55,13% e um retorno de investimento de 238,25% utilizando o modelo proposto. Esse trabalho contribui para o campo da inteligência artificial já que obteve resultados positivos utilizando redes neurais LSTM como classificador da tendência do Bitcoin, podendo ser feita uma implementação de um algoritmo de compra e venda que esteja vinculado a uma corretora utilizando esse modelo como classificador.

O restante do texto será organizado da seguinte forma: O Capítulo 2 apresenta os conceitos do mercado financeiro, séries temporais, métricas e redes neurais. O Capítulo 3 apresenta os trabalhos que se relacionam com o assunto abordado neste trabalho, sendo

dividido em 2 Seções. A primeira contendo trabalhos que utilizam como base de dados criptomoedas e a outra que utiliza dados do mercado tradicional. O Capítulo 4 apresenta modelo proposto para predição da tendência do preço do Bitcoin. O Capítulo 5 apresenta uma metodologia de experimentos para avaliar o modelo juntamente com os resultados obtidos. Finalmente, o Capítulo 6 apresenta as conclusões dos resultados obtidos.

2 CONCEITOS E TECNOLOGIAS UTILIZADAS

Neste Capítulo, são apresentados os conceitos que foram utilizados como base para esse trabalho. Conceitos sobre mercado financeiro, redes neurais e LSTM que foram usados para construir o modelo e posteriormente treiná-lo e fazer os experimentos necessários para averiguar sua eficiência são explicados neste Capítulo.

2.1 Mercado Financeiro

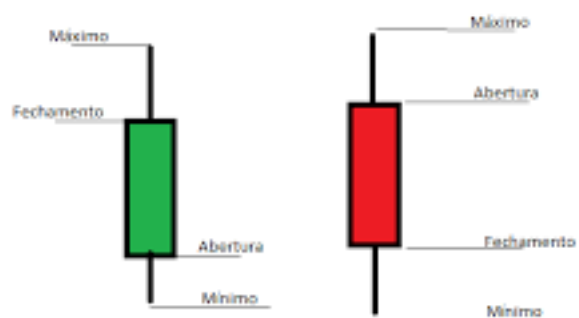
Grande parte dos conceitos utilizados neste trabalho são do mercado financeiro, e comumente são desconhecidos pela maioria das pessoas. O objetivo desta seção é esclarecer esses conceitos para melhor entendimento do trabalho. Os conceitos apresentados nessa seção já estão bem consolidados na área e são utilizadas referências da literatura que já estão bem definidas (ELDER, 1993; ELDER, 2002). Existem diversas outras referências mas os conceitos utilizados em todas são os mesmos.

2.1.1 *Candlestick*

Um *candlestick* nada mais é do que uma representação gráfica do comportamento do preço do ativo em um determinado período de tempo e recebe esse nome por ser visualmente semelhante a uma vela. Cada um deles possui quatro características que o definem, as quais são abertura, fechamento, máxima e mínima do preço naquele intervalo de tempo, como mostrado na Figura 2.1. A coloração do corpo define se houve uma variação positiva, preço de abertura menor que o preço de fechamento, ou negativa, preço de abertura maior do que o de fechamento, recebendo, respectivamente, as cores verde e vermelho. Os pavios, ou sombras, que ficam acima e abaixo do corpo da vela são os preços máximos e mínimos que o ativo atingiu naquele intervalo.

O histórico da cotação do Bitcoin utilizado neste trabalho cobrirá o período de 2014 até os dias atuais (2022) e foi obtido com os dados fornecidos pela Yahoo Finance (yahoo, 2022)

Figura 2.1 – Representação de um *Candlestick* de alta e de baixa, esquerda e direita respectivamente



Fonte: (bussoladoinvestidor, 2022)

2.1.2 Negociação de Ativos

No mercado financeiro, os ativos financeiros podem ser negociados de duas formas distintas:

- **Compra e Venda normal**- esse é o método mais utilizado pelos agentes econômicos. O ativo é comprado em um determinado intervalo de tempo e é vendido num momento posterior à compra. Essa ação pode gerar lucro, se o preço de compra do ativo for menor que o preço que o ativo foi vendido, ou pode gerar prejuízo, caso o preço de compra seja maior do que o preço de venda.
- **Compra e Venda a descoberto**- esse método é semelhante ao anterior, porém mais incomum de ser utilizado. Nessa forma de negociação, o agente econômico vende o ativo sem efetivamente tê-lo e recompra o ativo após um período de tempo. Nesse caso, o agente obtém lucro caso o preço de venda seja maior do que o preço de compra e obtém prejuízo caso o preço de venda seja menor do que o preço de compra.

Uma estratégia de compra e venda de ativos pode ser caracterizada por compras e/ou vendas tanto normais quanto descobertas, variando dependendo do agente econômico.

Existem diversos tipos de abordagens para comprar e vender ativos no mercado. Algumas estratégias podem ser simplesmente comprar o ativo e segurá-lo por um longo prazo, mas também existem estratégias mais complexas que podem envolver diversos

tipos de análise técnica. Uma das mais famosas no mundo de criptomoedas é o *Buy and Hold* que como o próprio nome diz é comprar e segurar o ativo. Essa estratégia, apesar de ser simples, funciona muito bem para o Bitcoin, visto que desde a sua criação o Bitcoin cresce exponencialmente de preço. Outra estratégia simples, como é auto intitulada, é a estratégia trivial a qual consiste em comprar ou vender o ativo de acordo com a variação do último período, comprando caso a variação tenha sido positiva e vendendo caso contrário.

Também é possível fazer compras e vendas a partir de indicadores técnicos. Uma estratégia comum é comprar um ativo quando seu preço ultrapassa uma média móvel tanto para cima como para baixo podendo indicar uma virada de tendência do preço desse ativo. Outros indicadores técnicos também podem ser utilizados de maneira semelhante como o índice de força relativa . Algumas dessas estratégias serão comparadas aos resultados do modelo proposto para analisar a eficiência tanto em relação à acurácia quanto em relação ao retorno do investimento.

2.1.3 Indicadores técnicos

Outro conceito que é utilizado neste trabalho são os indicadores técnicos (HUANG; HUANG; NI, 2018) que são modelos matemáticos que utilizam os valores passados do ativo para tentar mostrar alguma tendência no preço. A seguir são listados alguns desses indicadores:

- Médias móveis- o indicador mais usado e mais simples é a média móvel tanto na sua forma simples quanto exponencial. O cálculo desse indicador é feito fazendo a média dos n valores mais recentes do gráfico, o que ajuda a filtrar ruídos e falsas mudanças de tendência.
- Índice de força relativa (IFR)- esse índice também é muito utilizado no mercado financeiro para analisar tendências de forma técnica. Para calculá-lo é utilizada a fórmula a seguir:

$$IFR = 100 - \frac{100}{\frac{A}{B}}$$

Onde A é média das cotações que o preço aumentou dos últimos n períodos e B a média das cotações que o preço diminuiu no mesmo intervalo. O resultado da equação varia de 0 a 100 e indica se o mercado está sobre-comprado (com excesso de compras, indicando uma possível desvalorização do preço nos próximos períodos)

caso o valor esteja mais próximo de 100 ou sobre-vendido (com excesso de vendas, indicando uma possível valorização do preço nos próximos períodos) caso esteja mais próximo de 0.

2.1.4 Indicadores Macroeconômicos

Pelo fato de ser um mercado jovem, as criptomoedas acabam por sofrerem impactos da macroeconomia (JAQUART; DANN; WEINHARDT, 2021), principalmente das grande potências econômicas, como os Estados Unidos. Adicionar variáveis desse tipo pode acrescentar informações relevantes para a predição desse mercado já que ele possui uma alta correlação com o mercado financeiro global como é mostrada na Tabela 2.1 onde os dados utilizados foram correlacionados utilizando o intervalo de tempo semanal.

Tabela 2.1 – Tabela de correlação

	NYA	S&P500	NASDAQ
BTC	0.905932	0.910538	0.918350

Um dos indicadores mais conhecidos é o S&P500, um índice composto por 500 maiores ações com base no tamanho de mercado das bolsas americanas NYSE e NASDAQ. Como o dólar é a moeda mais forte e estável do mercado, esse índice tem uma grande relevância para indicar a situação da economia global, já que o índice engloba as maiores empresas da maior potência econômica do mundo. Outro indicador dos Estados Unidos que pode ser relevante é a taxa de juros americana, já que ela tem uma grande importância na definição de momentos de retração ou expansão econômica devido a diminuição ou aumento da taxa de acesso à crédito.

O ouro também tem grande importância na economia, já que ele é considerado a reserva de valor mais sólida devido a sua escassez natural, facilidade de negociação e pelo seu valor crescente a longo prazo. Em momentos de incerteza econômica, muitos agentes econômicos tendem a migrar seu capital para o ouro como uma forma de se proteger de instabilidades econômicas e esse fato pode ajudar a analisar tendências e comportamentos da macroeconomia. Além do ouro, outras *commodities* como gás natural, petróleo, prata, gasolina, aço entre outros têm grande relevância na macroeconomia visto que estes são a matéria prima ou são essenciais para de diversos setores industriais.

2.1.5 Séries Temporais

Séries temporais (KIM; CHOI; KANG, 2019) são dados coletados em intervalos de tempo regulares com dependência de ordem. Os valores da série temporal só podem variar com o tempo ou com variáveis que sejam dependentes do tempo. Uma série temporal pode ser decomposta nos seguintes itens:

- Tendência- capta elementos de longo prazo relacionados com à série de tempo;
- Ciclo- longas ondas, mais ou menos regulares, em torno de uma linha de tendência;
- Sazonalidade- capta os padrões regulares da série de tempo; e
- Aleatório- capta todos os efeitos que não foram incorporados pela série de tempo via os três componentes anteriormente citados, ou seja, é o resíduo.

Vale ressaltar que não são todas as séries temporais que apresentam todos esses itens. A notação geral para séries temporais é a seguinte:

$$\{X_t\}, t \in T$$

onde cada observação X_t ocorre em um determinado tempo t que pertence ao conjunto T de tempos que pode ser discreto ou contínuo.

Séries Temporais vem se tornando cada vez mais populares quando o assunto se trata de como prevê-las com base no histórico do objeto estudado. As áreas onde essas séries são aplicadas são diversas, como, por exemplo, demografia, epidemiologia, meteorologia entre outros. Porém, o que mais se destaca são os estudos de previsão de ativos financeiros com o intuito de comprá-los ou vendê-los em momentos oportunos para conseguir um retorno maior do que o capital investido, ou seja, lucrar.

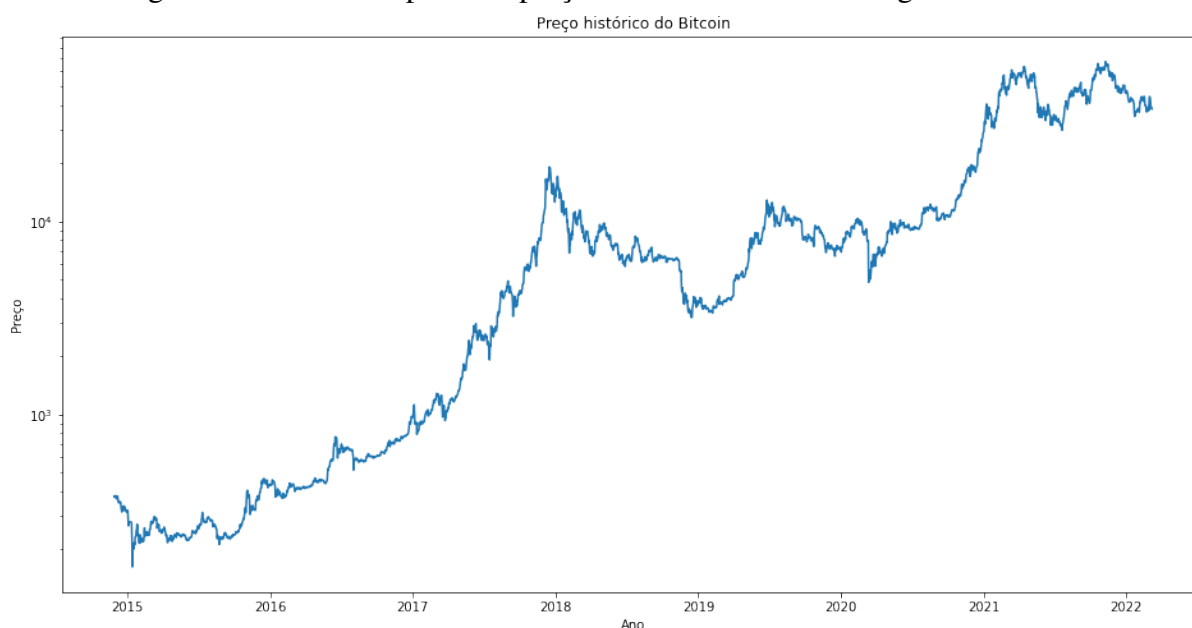
Séries temporais financeiras modelam os valores de um determinado ativo ou indicador econômico, como, por exemplo, o PIB de um país, índice de desemprego, taxa de juros, preço do ouro e também o preço de criptomoedas. Além de apresentar tendência e sazonalidade, esses tipos de series temporais apresentam outras características especiais:

- Pontos influentes- valores que destoam do padrão da série que podem ser causados por momentos de instabilidade econômica onde existe uma grande volatilidade, mas que volta a estabilidade posteriormente;
- Heteroscedasticidade condicional- a volatilidade não é constante e é dependente ou é condicionada à volatilidade do período anterior ;

- Não-linearidade - não é possível modelar a série temporal em uma função linear devido a seu comportamento estocástico e sua complexidade.

A Figura 2.2 mostra a série temporal do preço do Bitcoin desde o final do ano de 2014 até os dias atuais. O gráfico está na escala logarítmica, pois o preço do Bitcoin valorizou muito ao longo desses anos, sendo necessário o uso dessa escala para melhor visualização do comportamento do preço ao longo desses anos.

Figura 2.2 – Série temporal do preço do Bitcoin em escala logarítmica



Fonte: Próprio Autor

2.2 Redes neurais artificiais

Redes neurais artificiais são um tipo específico de técnica de aprendizado de máquina que é inspirado na estrutura neural de organismos inteligentes. Uma rede neural artificial é constituída de diversos neurônios, os quais são unidades de processamento que implementam uma função aritmética elementar. Os neurônios são organizados em camadas onde cada um deles recebe uma informação de entrada, processa essa informação e envia o resultado para o próximo neurônio que recebe como entrada até que chegue na última camada da rede que fornece o resultado final. Enquanto o sistema está funcionando, ele melhora sua performance ajustando os pesos entre as conexões de cada nodo.

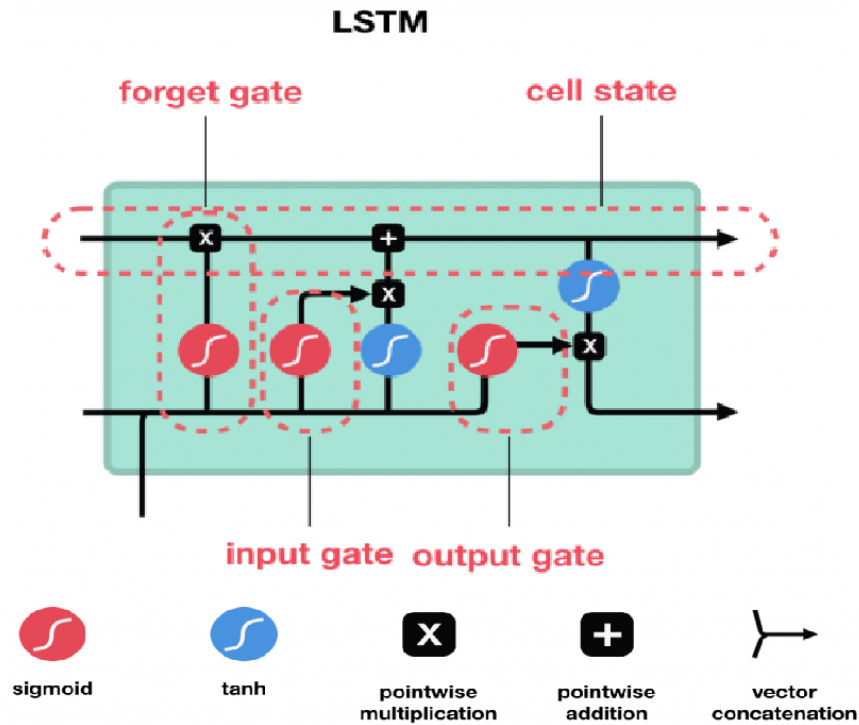
Os principais tipos de rede neural são: rede neural direta, perceptron multicamadas, redes neurais artificiais de função de base radial, redes neurais recorrentes e redes

neurais modulares. A rede neural recorrente (RNN) é a mais utilizada nos trabalhos relacionados e tem o melhor desempenho devido a sua capacidade de ser um aproximador universal se possuir um número considerável de nodos ocultos e também funções de ativação não linear, o que satisfaz o problema da alta volatilidade do Bitcoin. Por esse motivo, foi escolhida a rede neural recorrente para predição nesse trabalho. RNNs possuem a característica de serem grafos cíclicos, já que possibilitam que exista um transporte de informações relevantes das etapas iniciais para as etapas finais do processo da rede neural. Existem dois tipos de RNN: *Long short-term memory* (LSTM) e *Gated Recurrent Unit* (GRU). As LSTMs são utilizadas em diversos problemas como reconhecimento de caligrafia, geração de música, reconhecimento de fala e análise de dados financeiros.

A Figura 2.3 representa uma célula LSTM. Da mesma forma que outras redes neurais, ela processa os dados de forma sequencial onde a saída de um nodo é utilizada como entrada para o próximo até chegar no último nodo que fornece a saída final da rede neural. A grande diferença se encontra nas operações que estão dentro de cada célula, ou seja, a maneira como ela processa os dados. A *cell state* é encarregada de transportar as informações para as próximas células da sequência, servindo como uma memória da rede. As informações que esse estado carrega podem ser adicionadas ou removidas pelos *gates* que são treinados para decidirem se os dados são relevantes ou não. Essas possuem funções de ativação sigmóide as quais fazem uma normalização dos dados para que todos os valores fiquem no intervalo de 0 a 1, onde quanto mais próximo de 1 mais relevante o dado é e menos relevante caso o valor se aproxime de 0.

Existem três tipos de *gates*: *forget gate*, *input gate* e *output gate*. O primeiro decide se as informações devem ser mantidas ou descartadas recebendo como entradas as informações do *hidden state* anterior e da entrada atual, que são as entradas de todos os *gates*. A *cell gate* é calculada multiplicando o valor da *cell gate* anterior multiplicado pela saída da *forget gate*. O segundo faz o mesmo processo que o anterior com a adição de que as mesmas entradas passam por uma função tangente a qual normaliza os valores entre -1 e 1 e as duas saídas dessas funções são multiplicadas entre si e o resultado é adicionado com o valor da *cell gate* gerando o valor que será passado para a próxima célula. O último é encarregado de calcular o *hidden state*, que possui informações de células anteriores, que serão transmitidas para a próxima célula. O *hidden state* é calculado através da multiplicação do resultado da função sigmóide aplicada na informação da *hidden state* junto com a entrada atual com o resultado da aplicação da função tangente na informação da *cell gate* que foi modificada.

Figura 2.3 – Representação de uma célula LSTM



Fonte: (deeplearningbook, 2022)

2.3 Considerações Finais

Neste capítulo foram apresentados os conceitos de redes neurais, principalmente de células LSTM, que foram utilizados no desenvolvimento deste trabalho para a construção e execução do modelo preditivo. Os conceitos sobre o funcionamento do mercado financeiro são fundamentais para a configuração dos parâmetros da rede neural.

3 TRABALHOS RELACIONADOS

Neste Capítulo, são apresentados os principais trabalhos relacionados. Os trabalhos estão divididos em duas seções, onde a primeira utiliza criptomoedas como base de dados e a segunda utiliza o mercado tradicional como base de dados.

3.1 Criptomoedas

No artigo de Zhengyang *et al.* (ZHENGYANG *et al.*, 2019) é feita uma comparação de dois modelos de aprendizado de máquina: *Fully-connected Artificial Neural Network* (ANN) e *Long-Short-Term-Memory* (LSTM). Esses modelos foram aplicados no gráfico do Bitcoin utilizando como alvo o preço mais alto do dia. Notou-se que o primeiro teve um erro menor tanto utilizando a métrica de Raiz Quadrada do Erro Médio (RMSE) quanto Erro Médio Absoluto (MAE). Posteriormente, foram utilizadas outras criptomoedas junto com o Bitcoin para testar a hipótese de que a predição seria melhorada, pois as criptomoedas podem compartilhar informações que podem melhorar a predição de outras tendo um resultado positivo.

Outra abordagem muito comum quando se trata de predição de séries temporais é o uso de ensemble de redes neurais. Em Sin *et al.* (SIN; WANG, 2017) é proposto um modelo híbrido de *Genetic Algorithm* baseado em *Selective Neural Network Ensemble* (GASEN) com o intuito de prever qual a direção do preço do Bitcoin no dia posterior. O ensemble foi criado com 5 *Multi-Layered Perceptron* (MLP) com as mesmas especificações variando somente o número de nodos nas duas camadas ocultas de cada modelo. Todos os modelos têm 190 nodos de entrada o que demandou um poder computacional altíssimo que foi amenizado com o uso do algoritmo de Levenberg–Marquardt o que aumentou tanto a otimização quanto a acurácia em comparação a outros algoritmos heurísticos.

Um problema recorrente no uso de algoritmos de aprendizado de máquina é a memória de curto prazo que faz com que as entradas iniciais que alimentam o algoritmo acabam por serem menos consideradas do que as últimas em sequências de dados maiores, efeito conhecido como *vanishing gradient*. Para resolvê-lo, existem dois tipos de modelos: LSTM e GRU, os quais possuem células que mantêm apenas informações relevantes para a predição e as que não forem relevantes são descartadas. Jaquart *et al.* (JAQUART; DANN; WEINHARDT, 2021) utilizam esses dois modelos para predição de

movimentações curtas, de 1 a 60 minutos, da série do Bitcoin do ano de 2019. Além dos modelos já citados, também são utilizados outros algoritmos como rede neural direta, árvore de decisão e floresta aleatória.

Saad *et al.* (SAAD; MOHAISEN, 2018) implementaram modelos baseados em regressão múltipla para prever os preços do Bitcoin usando pontos de dados anteriores a 2017 (regressão linear, floresta aleatória e gradiente descendente). Dados entre 2016 e 2017 foram utilizados para treinar os modelos. Os resultados incluem erro quadrático médio (RMSE), valor de erro absoluto médio (MAE) e acurácia.

Livieris *et al.* (LIVIERIS et al., 2021), usam um modelo híbrido com Convolutional Neural Networks (CNN) com LSTM utilizando 3 anos de dados (janeiro de 2017 a outubro de 2020) de preços em uma estrutura de rede neural profunda de múltiplas entradas. As defasagens utilizadas no experimento são de 7 e 14 dias, e os modelos múltiplos alteram o número de suas camadas convolucionais (16 e 32, respectivamente), bem como o número de neurônios em sua camada densa (64 e 128, respectivamente). O modelo MICDL (LSTM + CNN) consiste em 3 camadas convolucionais com 16 filtros, cada um recebendo uma entrada de dados de série temporal únicos (pooling médio de tamanho 2) e uma camada LSTM de 50 unidades.

3.2 Mercado Tradicional

Como o mercado de criptomoedas é relativamente jovem, existem relativamente poucos artigos acadêmicos que englobam o uso de redes neurais para a predição do preço desses ativos. Porém, o mercado tradicional já existe a muitos anos e por esse motivo os estudos de redes neurais aplicadas em ações de empresas, principalmente nas ações norte-americanas, são encontrados em centenas. Como tanto o preço de criptomoedas quanto de ações são representadas por séries temporais, é muito conveniente utilizar como base trabalhos que utilizam ações como dados para treinar modelos de redes neurais para predição do preço desses ativos. Porém, vale ressaltar que as características de ações e criptomoedas como ativos são bem distintas, sendo ações um mercado muito mais consolidado, logo, menos volátil e mais estável do que as criptomoedas que são ativos novos e por tal fato apresentam uma grande flutuação no seu preço.

Hsu *et al.* (HSU et al., 2016) demonstrou que métodos de aprendizado de máquina podem ter resultados mais precisos que métodos econômicos quando se trata de predição do mercado financeiro. Outro fato que ele comprovou neste trabalho foi que o efeito da

predição do mercado financeiro foi afetado pela maturidade do mercado, pela variável de entrada, pelo período analisado e pelo método de predição.

No artigo de Moghar *et al.* (MOGHAR; HAMICHE, 2020) são utilizadas as cotações da abertura diária das ações das empresas Google e Nike, GOOGL e NKE respectivamente, como dados para o treinamento do modelo LSTM para predição do valor futuro dessas ações. O objetivo principal foi analisar qual precisão do modelo conseguia alcançar e quanto às épocas podem influenciar no modelo. Os dados utilizados cobrem um período de 15 anos para a NKE e de 9 anos para a GOOGL. Os dados de teste e treino foram particionados em 20% e 80% e as quantidades de épocas de treino foram 12,25,50 e 100 para cada modelo. Obtiveram como conclusões que tanto o número de épocas quanto o período dos dados têm grande influência na precisão do modelo.

Em Pang *et al.* (PANG et al., 2020) também é utilizado uma RNN do tipo LSTM, mas com duas abordagens diferentes. A primeira utiliza uma camada incorporada e a segunda utiliza um codificador automático para vetorizar os dados para fazer uma predição via LSTM. Os dados utilizados foram as ações da Shanghai A-shares tanto como seu próprio índice tendo um resultado melhor na predição no índice da bolsa chinesa.

3.3 Análise Comparativa

Neste capítulo, foram apresentados trabalhos relacionados e foi explorado o que esses trabalhos discutem e como eles se relacionam com o trabalho desenvolvido. A Tabela 3.1 mostra um quadro comparativo com as principais características de cada estudo. Existe uma lacuna a ser explorada no uso de LSTM para predição do preço do Bitcoin com o uso de *features* que não estão diretamente ligadas ao preço do Bitcoin, mas que podem ajudar a aumentar a acurácia da predição, como, por exemplo, os indicadores macroeconômicos. Outra abordagem não utilizada é a redução da predição para um problema de classificação, já que a alta volatilidade do Bitcoin torna difícil a predição exata do seu preço. Esses são os diferenciais que serão testados neste trabalho com o intuito de alcançar uma maior acurácia e ganho de capital.

Tabela 3.1 – Comparativo entre os trabalhos que utilizam redes neurais para predição de ativos financeiros

Trabalho	Tipo de Rede Neural	Série Temporal	Métricas
(ZHENG YANG et al., 2019)	Fully-connected ANN e LSTM	Bitcoin (2015 - 2018)	RMSE e MAE
(SIN; WANG, 2017)	MLP e GASEN	Bitcoin (2015 - 2017)	Acurácia e ROI
(JAQUART; DANN; WEINHARDT, 2021)	LSTM, GRU, rede neural direta, árvore de decisão e árvore aleatória	Bitcoin (2019)	Acurácia
(SAAD; MOHAISEN, 2018)	Regressão Linear, árvore aleatória e gradiente descendente	Bitcoin e Ethereum (2016 - 2018)	RMSE, MAE e acurácia
(LIVIERIS et al., 2021)	CNN e LSTM	Bitcoin (2017 - 2020)	MAE, RMSE e Acurácia
(MOGHAR; HAMICHE, 2020)	LSTM	GOOGL (2004 - 2019) e NKE (2010 - 2019)	função de perda e análise gráfica
(PANG et al., 2020)	LSTM	Shanghai A-share (2006 - 2016)	MSE e acurácia
Modelo LSTM de classificação	LSTM	Bitcoin (2014 - 2022)	ROI e acurácia

4 MODELO LSTM DE CLASSIFICAÇÃO PARA PREDIÇÃO DA TENDÊNCIA DO PREÇO DO BITCOIN

Neste Capítulo, é apresentado como foi construído o modelo de predição utilizando células LSTM para prever o movimento do preço Bitcoin no próximo intervalo de tempo em forma de classificação, tendo como resposta se o preço tende a subir ou cair. Inicialmente, é descrito como foi feita a preparação dos dados para serem utilizados na instanciação do modelo. Os recursos computacionais utilizados para a preparação dos dados e implementação do modelo também são descritos.

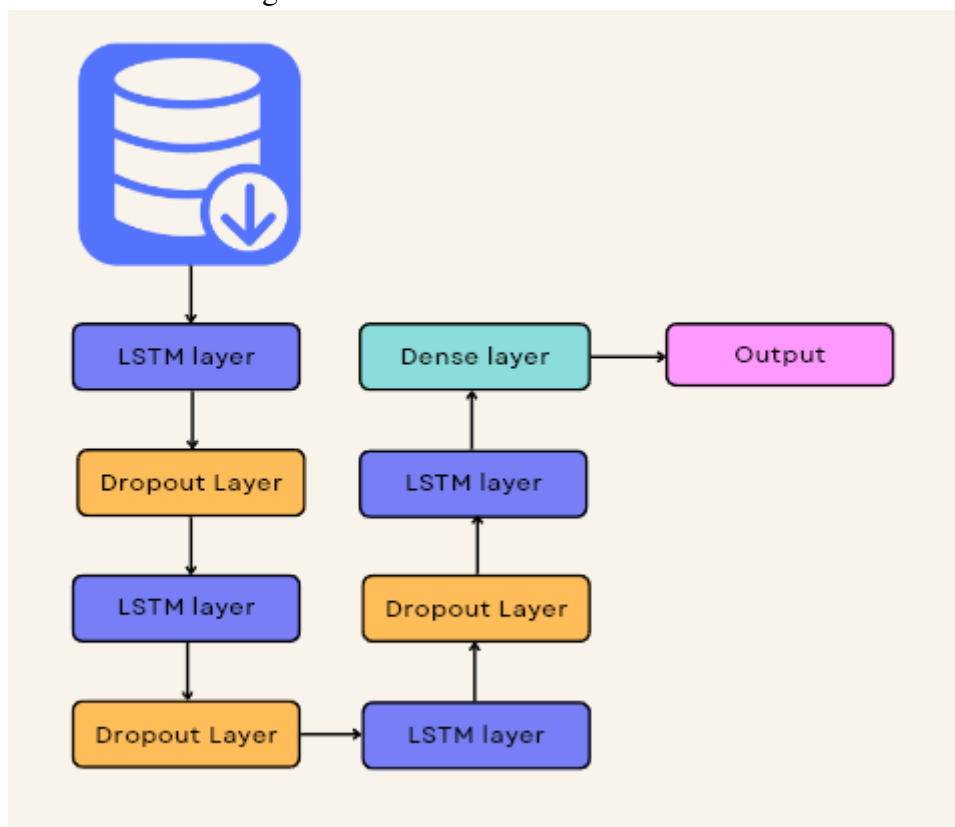
4.1 Arquitetura do Modelo Proposto

O modelo proposto está representado na Figura 4.1 e tem como objetivo fazer a predição da tendência do preço do Bitcoin. O modelo utiliza redes neurais com células LSTM e faz a classificação se o preço tende a subir ou descer. As células do modelo de LSTM seguiram os moldes do modelo apresentado no trabalho de Hochreiter (HOCHREITER; SCHMIDHUBER, 1997).

Devido à complexidade do problema, são utilizadas quatro camadas de células LSTM visto que foi o valor que obteve os melhores resultados nos testes realizados, onde o modelo foi testado utilizando de duas a seis camadas. O modelo foi inicializado com 100 unidades na dimensão da saída, esse número foi escolhido para que o modelo lembre de padrões complexos que compõem a série temporal. Cada camada recebe como entrada um tensor com três dimensões: *batch*, *timesteps*, *features*. Entre cada camada de LSTM existe uma camada de *dropout* que randomicamente ignora uma parcela, nesse caso o valor definido foi 0,2 ou 20%, das saídas de cada camada. A camada de *dropout* ajuda na redução do sobre-ajuste do modelo. Finalmente, foi adicionada uma camada densa que faz uma multiplicação matricial para que a saída do modelo seja da forma desejada.

O modelo foi treinado com uma função de perda da forma *binary crossentropy*, já que a classificação do problema é binária além de usar um o algoritmo de otimização Adam que é uma extensão da descida de gradiente estocástica. O modelo é treinado com diferentes épocas de treinamento (12, 25, 50, 100 e 200) para verificar qual seria o valor que obtém os melhores resultados. O modelo é treinado para prever a movimentação do Bitcoin para o próximo intervalo de tempo, onde o *target* é feito um *one hot encode*

Figura 4.1 – Resumo do modelo LSTM



Fonte: Próprio Autor

na forma de lista binária onde $[0,1]$ indica uma queda no preço no próximo intervalo de tempo e $[1,0]$ indica uma subida no preço no próximo intervalo de tempo.

4.2 Recursos Computacionais

Para a construção do modelo, foi utilizada a linguagem de programação *python* escrita na forma de *jupyter notebooks*. A biblioteca *Keras* (Keras, 2022) é utilizada para fazer as camadas da arquitetura do modelo, ou seja, as camadas LSTM, *dropout* e *dense* e também para o treinamento e teste do modelo.

Os dados foram processados tanto utilizando a biblioteca *pandas* (Pandas, 2022), que é muito utilizada para trabalhar com dados, como a biblioteca *numpy* (Numpy, 2022) que possui diversos recursos para manipulação de dados. Para a representação gráfica tanto dos dados utilizados como dos resultados obtido foi utilizada a biblioteca *matplotlib* (Matplotlib, 2022). Também foi utilizada a biblioteca *yfinance* (Pypi, 2022) para coletar os dados a serem utilizados no modelo.

5 AVALIAÇÃO EXPERIMENTAL

Este Capítulo descreve os experimentos desenvolvidos para avaliar a eficácia do modelo LSTM de classificação na predição da tendência do preço do Bitcoin variando hiperparâmetros do modelo para chegar em uma configuração que obtém os melhores resultados. A Seção 5.1 apresenta as métricas que são utilizadas para analisar os resultados obtidos pelo modelo proposto. A Seção 5.2 apresenta a metodologia usada para a execução dos experimentos. Finalmente, a Seção 5.3 descreve os experimentos que são feitos e mostra os resultados obtidos.

5.1 Configuração dos Experimentos

5.1.1 Tratamento dos dados

Os dados utilizados são séries temporais que representam os preços dos ativos em um intervalo de tempo determinado. Como é comum no mercado financeiro, essas séries temporais são representadas na forma de *candlesticks*, as quais possuem quatro valores: abertura, fechamento, máxima e mínima do preço. Esses valores são utilizados como dados de entrada do modelo, tanto com o preço do Bitcoin como no preço dos indicadores macroeconômicos. Os dados utilizados no trabalho compreendem as datas de 15 de setembro de 2014 até dia 15 de setembro de 2022, somando 8 anos de dados dos preços dos ativos e indicadores analisados. A periodicidade escolhida foi a semanal, totalizando 418 registros.

Cada amostra de preços corresponde à flutuação do preço no intervalo de uma semana, ou seja, os preços máximos e mínimos e também o preço de abertura e fechamento da semana em específico. Outra métrica que foi utilizada foi o volume total de transações naquele período que corresponde ao montante financeiro de um ativo que foi comprado e vendido. Foi feita uma normalização dos dados para que os valores de cada atributo separadamente fique no intervalo de 0 a 1, onde o valor máximo do atributo corresponde a 1 e o valor mínimo corresponde a 0. A normalização foi feita para evitar que valores muito elevados não pesem na hora do treinamento do modelo.

Após a normalização dos dados, é necessário formatá-los de tal maneira que eles possam ser aceitos como entradas do modelo. Cada amostra tem um conjunto de valores dos últimos x intervalos de tempo, *timesteps*, que varia de acordo com cada experimento

junto com o valor alvo que é se o preço do Bitcoin aumentou ou diminuiu no próximo intervalo de tempo. Além disso, os dados são separados em dados de treino, para que o modelo aprenda os padrões da série temporal e dados de teste para posteriormente ao treino o modelo seja testado com valores diferentes do treino para evitar sobre-ajuste e verificar os resultados da predição do modelo. A proporção dos dados de treino e teste foram 80% e 20%, respectivamente. Esses valores foram escolhidos devido a maioria dos trabalhos relacionados utilizarem esse valores e obterem bons resultados.

5.1.2 Métricas

São utilizadas duas métricas para a avaliação de cada experimento: acurácia e retorno do investimento. O cálculo da acurácia é obtido pela soma da quantidade de predições que foram corretas dividido pela soma total da quantidade de predições como mostra na fórmula a seguir:

$$Acurácia (\%) = \frac{Total\ de\ acertos}{Total\ de\ predições} * 100$$

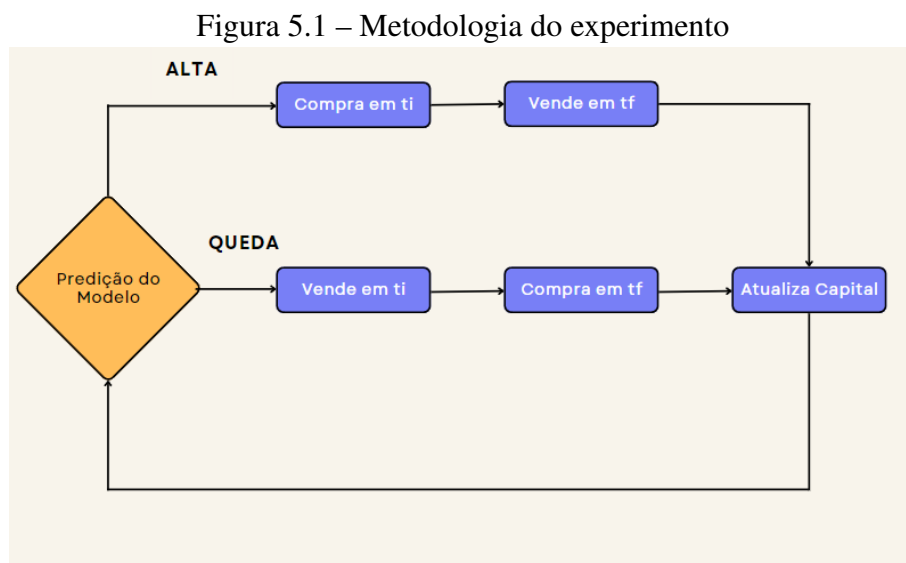
O retorno do investimento (ROI - *return of investment*) é a métrica que calcula a diferença percentual do capital investido inicialmente com o percentual do capital final após terem sido feitas compras e vendas de acordo com alguma estratégia. Essa métrica é muito relevante visto que no mercado financeiro o que mais importa é quanto lucro ou prejuízo foram obtidos nas negociações. É preferível que o modelo tenha um ROI alto e uma acurácia baixa do que um ROI baixo e uma acurácia alta. Vale ressaltar que no mercado financeiro cada negociação possui taxas para cada negociação que variam dependendo da corretora que está sendo feita. Em nenhum trabalho relacionado às taxas são consideradas e por isso não são utilizadas no cálculo do ROI neste trabalho. O ROI é calculado da seguinte maneira:

$$ROI (\%) = \frac{Capital\ Final - Capital\ Inicial}{Capital\ Inicial} * 100$$

5.1.3 Metodologia dos Experimentos

Cada configuração do modelo foi treinada e avaliada 20 vezes, sendo calculada a média dos resultados com o intuito de diminuir possíveis variações em cada experimento e verificar a estabilidade dos resultados.

Para cada predição do modelo foram seguidos os passos do fluxograma da Figura 5.1:



Fonte: Próprio Autor

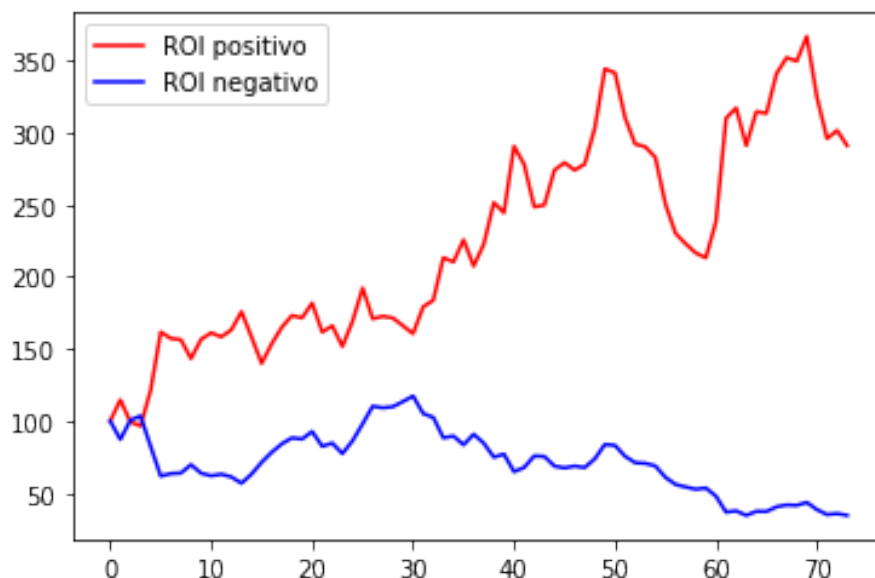
Caso o modelo preveja uma alta no preço do Bitcoin no próximo período de tempo, é realizada uma operação de compra no preço de abertura do Bitcoin e uma venda no preço de fechamento do Bitcoin. A operação retorna lucro se o preço de fechamento for maior que o preço de abertura daquele período e retorna um prejuízo, caso contrário.

Inversamente, se o modelo prevê uma baixa no próximo período de tempo no preço do Bitcoin, é realizada uma operação de venda na abertura do preço do Bitcoin e uma compra no preço de fechamento. Nesse caso, se o preço de fechamento for maior que o preço de abertura, a operação tem um prejuízo e obtém um lucro, caso contrário.

Ambas possibilidades, após terem sido feitas suas compras e vendas, atualizam o capital somando o lucro obtido na operação ou subtraindo o prejuízo do capital investido. O processo se repete para toda a base de dados de treino até que a última operação tenha sido feita e o capital final é o resultado de todas essas operações. O número total de operações é igual ao número de dados de teste menos o valor de *timesteps* utilizado no modelo. Por exemplo, se o número de *timesteps* for igual a 10, serão feitas 74 operações já que o número de dados de testes é igual a 84. O gráfico da Figura 5.2 mostra um

exemplo da variação do capital utilizando esse método onde o ROI foi positivo e outro foi negativo, utilizando um capital inicial de 100 dólares :

Figura 5.2 – Exemplo de evolução do capital investido com ROI negativo e positivo



Fonte: Próprio Autor

Na Figura 5.2, as operações da predição do ROI positivo obtiveram, em sua maioria, lucro obtendo um capital final de 290.93 dólares e um ROI igual a 190.93%. Em contrapartida, a maioria das operações do ROI negativo obtiveram prejuízo, tendo como resultado final um capital de 34.58 dólares e um ROI de -65.42%

Tanto o ROI como a acurácia do modelo proposto foram obtidos através da média dos 20 treinamentos que o modelo passou. A média das métricas também foi utilizada para outras estratégias que têm resultados estocásticos, como a estratégia aleatória que será apresentada posteriormente.

Os resultados serão considerados satisfatórios caso a estratégia de compra e venda que seguirá as predições do modelo obtiver lucro após terem sido feitas todas as negociações. Quanto maior for o ganho de capital em relação ao capital inicial investido, melhor o resultado. Além disso, é preferível que o índice de acertos das previsões seja maior que 50%, já que um modelo aleatório obteria esse mesmo resultado sem demandar muito poder computacional.

Para fazer tanto o treinamento do modelo quanto a fase de testes foi utilizado um computador pessoal com 8GB de RAM com processador Intel Core i5-7400 (Intel, 2022).

5.1.4 Conjunto de dados

Os dados utilizados no trabalho compreendem as datas de 15 de setembro de 2014 até dia 15 de setembro de 2022, somando 8 anos de dados dos preços dos ativos e indicadores analisados. A periodicidade escolhida foi o semanal, totalizando 418 registros.

Os dados de treino do modelo estão no intervalo de 15 de setembro de 2014 até 1 de fevereiro de 2021, 80% do volume total de dados, enquanto que os dados de teste estão no intervalo de 8 de fevereiro de 2021 até 15 de setembro de 2022, 20% do volume total de dados. O número de registros utilizados no treino do modelo é igual a 334 enquanto que o número de registros utilizados no teste do modelo é igual a 84. Serão utilizados os quatro valores dos *candlestick* e também o volume de transações no período semanal referente ao Bitcoin. Os dados adicionais do experimento da Seção 5.2.2 também possuem as mesmas *features* e estão no mesmo intervalo de tempo dos dados do Bitcoin

5.2 Descrição dos Experimentos

Nesta seção, são detalhados os experimentos realizados no trabalho. São eles a execução e avaliação do modelo por meio das métricas apresentadas anteriormente:

1. Configuração Padrão. (Seção 5.2.1)
2. Utilizando *timesteps* diferentes. (Seção 5.2.2)
3. Utilizando *epochs* diferentes. (Seção 5.2.3)
4. Utilizando dados da macroeconomia. (Seção 5.2.4)
5. Comparação dos resultados com outras estratégias. (Seção 5.2.5)

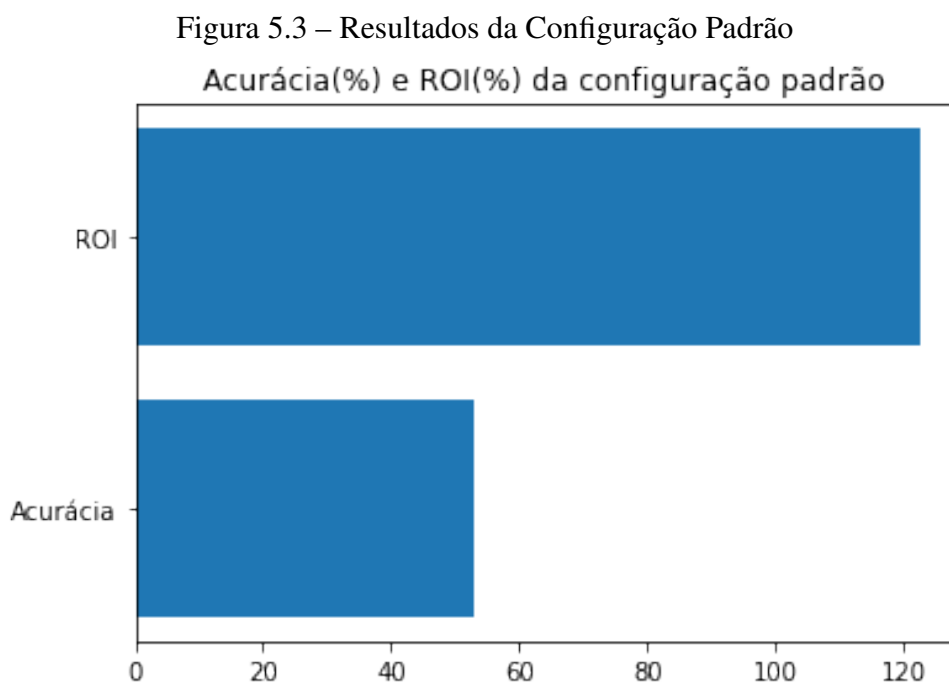
5.2.1 Configuração Padrão

Como serão feitos diversos experimentos para verificar qual configuração tem os melhores resultados tanto em relação à acurácia quanto em relação ao ROI, será proposta nesta seção uma configuração inicial que servirá como base de comparação quando se altera algum parâmetro do modelo ou a base de dados com que ele é treinado.

Todos os experimentos realizados têm em comum a fonte de dados do preço do Bitcoin, já que é o objeto de estudo que o modelo está tentando prever é essencial que ele

esteja na base de dados de todos os experimentos. Os dados da configuração padrão terão 10 *timesteps*, ou seja, será um conjunto dos 10 registros com as *features* correspondentes a cada período e o valor da tendência do Bitcoin no intervalo posterior ao último registro, 1 caso o preço tenha subido e 0 caso o preço tenha caído. O modelo padrão será treinado 25 vezes, ou seja, seu valor de *epochs* será 25.

Os resultados obtidos com essa configuração foram uma acurácia média de 52,90% e um ROI de 122,82% e estão representados na Figura 5.3 . Vale ressaltar que os dados de teste compreendem um intervalo de tempo total de 1 ano, 4 meses, e 22 dias o que torna o ROI obtido interessante já que o capital investido no experimento mais que dobrou em menos de um ano e meio o que é considerado um bom resultado no mercado financeiro.



Fonte: Próprio Autor

5.2.2 *Timesteps*

O objetivo desse experimento é avaliar qual a influência do valor de *timesteps* nos resultados do modelo e qual valor obtém os melhores resultados. O número de *timesteps* utilizado em cada experimento varia a quantidade de amostras de cada conjunto de registros que servirão de entrada para o modelo. Consequentemente isso altera o formato de entrada do modelo e também diminui a quantidade de amostras do modelo. A Tabela 5.1

apresenta os resultados variando-se o número de *timesteps*

Tabela 5.1 – Tabela resultados variando *timesteps*

Número de <i>timesteps</i>	Acurácia	ROI	Tempo de Processamento
5	52,27%	150,01%	4m43s
10	52,9%	122,82 %	5m35s
20	55,00%	25,00%	6m55s
40	58,86 %	150,94 %	9m43s

5.2.3 Epochs

O objetivo desse experimento é avaliar qual a influência do valor de *epochs* nos resultados do modelo e qual valor obtém os melhores resultados. Nessa etapa do experimento, são testados os resultados obtidos variando-se o número de *epochs* do modelo, ou seja, o número de iterações que o modelo é treinado com os dados fornecidos. Com o aumento do número de *epochs* é de se esperar que o modelo demande mais tempo de processamento, já que ele faz mais iterações de treino. A Tabela 5.2 apresenta os resultados obtidos:

Tabela 5.2 – Tabela resultados variando *epochs*

Número de <i>epochs</i>	Acurácia	ROI	Tempo de Processamento
12	50,67%	98,96%	4m38s
25	52,9%	122,82 %	5m35s
50	54,18%	208,23%	7m29s
100	55,13 %	238,25 %	12m8s
200	53,64 %	156,78 %	22m

Nota-se com base nos resultados apresentados que os melhores resultados obtidos tanto em relação a acurácia quanto ao retorno do investimento foram oriundos do número de *epochs* igual a 100.

5.2.4 Adicionando dados da Macroeconomia

Esse experimento tem como objetivo avaliar os efeitos da adição de dados que não estão diretamente ligados ao preço do Bitcoin tem nos resultados do modelo. Como apresentado anteriormente na Seção 2.1.5, a macroeconomia tem alta correlação com o preço do Bitcoin. Nesse experimento será analisada qual a influência da adição de dois

tipos de dados macroeconômicos: índices da bolsa americana e preço das *commodities*. A tabela a seguir mostra os resultados obtidos utilizando esses dados junto com o Bitcoin.

Tabela 5.3 – Tabela resultados variando dados de entradas

Dados	Acurácia	ROI	Tempo de Processamento
Bitcoin	52,9%	122,82 %	5m35s
Bitcoin + Bolsa Americana	51,82%	70,75 %	5m30s
Bitcoin + <i>commodities</i>	47,97 %	84,83 %	5m46

Os dados usados junto ao Bitcoin foram os índices da NASDAQ, S&P500 e NYA e para as *commodities* foram o preço histórico do ouro e da prata. Foram escolhidos esses dados por serem os mais relevantes dentro de suas categorias. Os gráficos dos dados dos índices da bolsa americana encontram-se no Apêndice A e os gráficos das *commodities* encontram-se no Apêndice B.

Percebe-se que a adição de novas *features* ao modelo acabam por não favorecerem os resultados apesar dos dados utilizados terem alta correlação com o preço do Bitcoin.

5.2.5 Comparação com outras estratégias

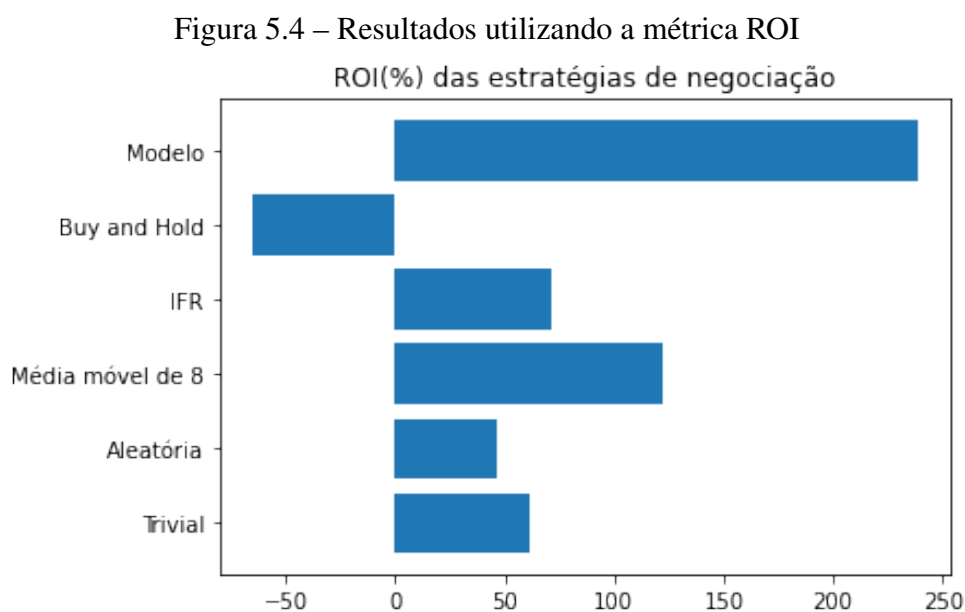
Com os resultados obtidos nos experimentos anteriores, notou-se que os hiperparâmetros do modelo que fornecem os melhores resultados separadamente são utilizando 40 *timesteps*, 100 *epochs* de treinamento e sem a adição de novos dados além dos vinculados ao preço do Bitcoin. Porém, quando esses hiperparâmetros foram utilizados juntos em um mesmo teste, os resultados obtidos foram: acurácia de 56,22% e um ROI de 207,73% e apesar de obter uma acurácia maior que a configuração com 10 *timesteps*, 100 *epochs* e sem adição de novos dados essa segunda configuração obteve melhores resultados em relação ao ROI e por esse motivo é a utilizada para comparação. Nesta subseção, são apresentadas outras estratégias de negociação para comparar com os resultados obtidos pelo modelo proposto.

Como o modelo foi treinado para indicar se no próximo período de tempo o preço do Bitcoin irá subir ou descer, as estratégias comparadas também seguem esse padrão. Algumas delas foram adaptadas para que sempre haja um resultado indicando se o ativo deve ser comprado ou vendido. Por exemplo, o indicador IFR comumente é utilizado com os valores de 70 e 20 como parâmetros de sobre-compra e sobre-venda, respectivamente, mas neste trabalho usamos o valor de 50 isoladamente onde valores acima desse limiar indicam venda e abaixo indicam compra.

As estratégias que serão comparadas ao modelo são as seguintes:

1. *Buy and Hold* - consiste em comprar o ativo no primeiro intervalo de tempo e vendê-lo no último;
2. Estratégia Trivial - consiste em comprar ou vender um ativo de acordo com o período anterior ao atual, comprar caso a variação de preço tenha sido positiva e vender caso contrário;
3. Estratégia Aleatória - consiste em gerar um resultado de compra ou venda na proporção de 50%;
4. Usando médias móveis - consiste em usar médias móveis como parâmetro de negociação;
5. Usando IFR - consiste em usar o índice de força relativa como parâmetro de negociação.

A única estratégia de compra e venda que não possui acurácia é a *Buy and Hold*, já que ela não consiste em uma série de negociações, só comprar o ativo no primeiro período e vendê-lo no último período. A média móvel utilizada na análise é a de 8 períodos semanais já que obteve os melhores resultados testando todos os períodos no intervalo de 1 a 30, o mesmo vale para o IFR. A Figura 5.4 ilustra os resultados de cada estratégia em relação ao ROI:



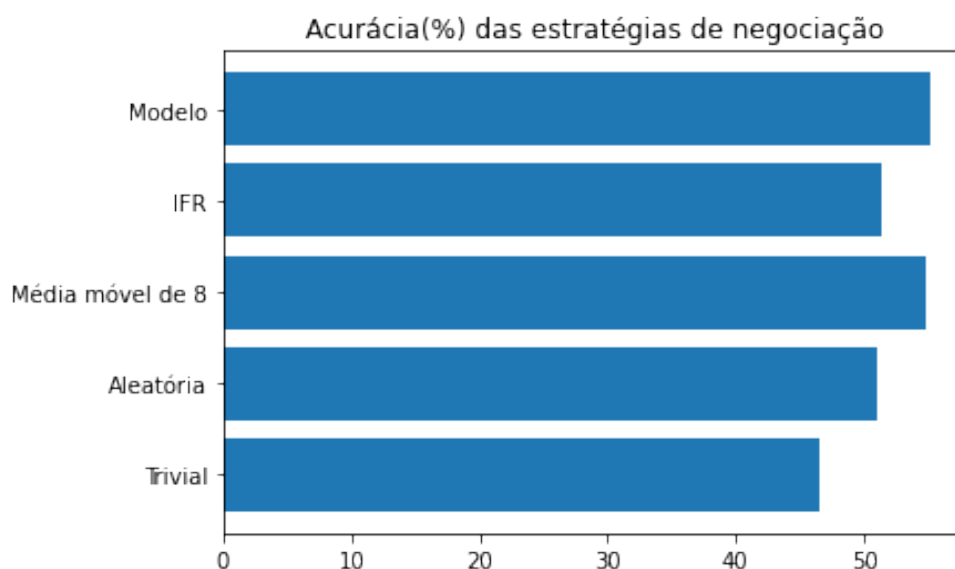
Fonte: Próprio Autor

O primeiro resultado que chama atenção é o do *Buy and Hold*, que foi a única estratégia que obteve retorno negativo nesses experimentos devido a ser uma estratégia de

longo prazo, logo, *Buy and Hold* não é a melhor para o intervalo de tempo relativamente curto que foi utilizado como teste neste estudo. Outro resultado interessante na métrica ROI foi a da estratégia aleatória, que apesar de sua acurácia de aproximadamente 50%, obteve um retorno positivo.

O modelo proposto foi o que obteve os melhores resultados em relação ao ROI, mostrando que os lucros obtidos por ele foram maiores do que as outras estratégias analisadas nessa etapa. Ele foi consideravelmente melhor que a estratégia de utilizar as médias móveis que obteve o segundo maior ROI das estratégias analisadas, um ROI de 126%, porém não necessitou de muito processamento já que ela é simplesmente uma estratégia baseada em médias dos últimos 8 períodos. A Figura 5.5 ilustra os resultados de cada estratégia em relação à acurácia:

Figura 5.5 – Resultados utilizando a métrica acurácia



Fonte: Próprio Autor

Os resultados da acurácia para todas as estratégias ficaram entre o intervalo de 45% e 55%. O modelo proposto, apesar de ter obtido o melhor ROI por uma grande margem de diferença, não obteve o mesmo resultado em relação à acurácia, ficando com acurácia muito próxima à da estratégia de média móvel, que obteve uma acurácia de 54,79%. Possivelmente, essa grande diferença no ROI obtida pelo modelo é devida às negociações que obtiveram maior lucro que as negociações da estratégia de média móvel apesar da acurácia próxima. A estratégia trivial obteve a pior acurácia das analisadas nesse experimento, obtendo uma acurácia de aproximadamente de 45%. Apesar do resultado ruim, por ter sido pior que a estratégia aleatória no quesito acurácia, obteve um ROI positivo.

5.3 Considerações Finais

Neste capítulo, foram apresentados todos os experimentos realizados durante o desenvolvimento deste trabalho e a análise dos seus resultados. O primeiro conjunto de experimentos teve como objetivo analisar qual impacto nas métricas obtém-se variando alguns hiperparâmetros de entrada do modelo. Após encontrar a melhor configuração do modelo com base nos resultados obtidos, foi feita uma comparação com os resultados de outras estratégias de negociação.

6 CONCLUSÃO

Esse trabalho apresentou um modelo de rede neural de classificação onde foram utilizadas células LSTM para prever a tendência do preço do Bitcoin. A maioria dos estudos em relação a predição de preço de ativos econômicos tenta prever os valores exatos do preço do ativo no futuro, o que torna o problema de predição muito mais complexo. A classificação nesse problema faz com que ele fique mais simples para a aplicação em estratégias de negociação. O modelo foi submetido a diversos testes com configurações diferentes e foi encontrada uma configuração que obteve resultados satisfatórios e melhores em comparação a outras estratégias de negociação. Os resultados obtidos mostraram que o modelo proposto é robusto suficiente para ser utilizado no mercado de criptomoedas, o qual tem alta volatilidade.

O modelo proposto, porém, tem uma grande variância nos seus resultados, sendo que seu uso em um mercado real possui alto risco e não foi feito para grandes aportes. É possível a implementação de um algoritmo de compra e venda que esteja vinculado a uma corretora que utilize o modelo como classificador de tendência para decidir qual seria a negociação que iria obter lucro.

Uma grande contribuição desse trabalho para o campo da inteligência artificial foram os resultados positivos utilizando redes neurais como classificador da tendência do Bitcoin, que é um ativo relativamente novo e que foi poucas vezes usado na literatura acadêmica como base de dados de estudos.

Com relação a trabalhos futuros, o modelo proposto pode ser avaliado em operações reais através de alguma ferramenta que faça as negociações automaticamente utilizando a classificação do modelo. Uma maneira de fazer esse algoritmo seria fazê-lo utilizando a API de alguma corretora. Outra possibilidade de extensão para o trabalho é utilizar o modelo para predição de outras criptomoedas, como Ethereum (Ethereum, 2022). Além disso, pode ser feita uma comparação dos resultados utilizando outros tipos de redes neurais como GRU e rede neural direta.

REFERÊNCIAS

- bussoladoinvestidor. **Figura de um Candlestick**. 2022. Disponível em: <<https://www.bussoladoinvestidor.com.br/grafico-de-candlestick/>>. Acesso em: 12 de setembro 2022.
- coinmarketcap.com. **Cryptocurrency Market Cap**. 2022. Disponível em: <<https://coinmarketcap.com/>>. Acesso em: 22 de março 2022.
- Companies Marketcap. **Top Assets by Market Cap**. 2022. Disponível em: <<https://companiesmarketcap.com/assets-by-market-cap/>>. Acesso em: 22 de março 2022.
- deeplearningbook. **Figura de uma LSTM**. 2022. Disponível em: <<https://www.deeplearningbook.com.br/arquitetura-de-redes-neurais-gated-recurrent-unit-gru/>>. Acesso em: 12 de setembro 2022.
- ELDER, A. **Trading for a living: psychology, trading tactics, money management**. [S.l.]: John Wiley & Sons, 1993.
- ELDER, A. **Come into my trading room: A complete guide to trading**. [S.l.]: John Wiley & Sons, 2002.
- Ethereum. **Ethereum White Paper**. 2022. Disponível em: <<https://ethereum.org/en/whitepaper/>>. Acesso em: 12 de setembro 2022.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, v. 9, p. 1735–80, 12 1997.
- HSU, M.-W. et al. Bridging the divide in financial market forecasting: Machine learners vs. financial economists. **Expert Systems with Applications**, v. 61, 05 2016.
- HUANG, J.-Z.; HUANG, W.; NI, J. Predicting bitcoin returns using high-dimensional technical indicators. **The Journal of Finance and Data Science**, v. 5, 11 2018.
- Intel. **Processador Intel® Core™ i5-7400**. 2022. Disponível em: <<https://www.intel.com.br/content/www/br/pt/products/sku/97147/intel-core-i57400-processor-6m-cache-up-to-3-50-ghz/specifications.html>>. Acesso em: 12 de setembro 2022.
- JAQUART, P.; DANN, D.; WEINHARDT, C. Short-term bitcoin market prediction via machine learning. **The Journal of Finance and Data Science**, v. 7, 03 2021.
- Keras. **Keras API reference**. 2022. Disponível em: <<https://keras.io/api/>>. Acesso em: 12 de setembro 2022.
- KIM, J.; CHOI, J.; KANG, C. Time series prediction using recurrent neural network. **The Korean Data Analysis Society**, v. 21, p. 1771–1779, 08 2019.
- LIVIERIS, I. et al. An advanced cnn-lstm model for cryptocurrency forecasting. **Electronics**, 01 2021.

Matplotlib. **Matplotlib documentation**. 2022. Disponível em: <<https://matplotlib.org/stable/index.html>>. Acesso em: 12 de setembro 2022.

MOGHAR, A.; HAMICHE, M. Stock market prediction using lstm recurrent neural network. **Procedia Computer Science**, v. 170, p. 1168–1173, 01 2020.

NAKAMOTO, S. Bitcoin: A peer-to-peer electronic cash system. **Cryptography Mailing list at <https://metzdowd.com>**, 03 2009.

Numpy. **Numpy documentation**. 2022. Disponível em: <<https://numpy.org/doc/>>. Acesso em: 12 de setembro 2022.

Pandas. **Pandas documentation**. 2022. Disponível em: <<https://pandas.pydata.org/docs/index.html>>. Acesso em: 12 de setembro 2022.

PANG, X. et al. An innovative neural network approach for stock market prediction. **The Journal of Supercomputing**, v. 76, 03 2020.

Pypi. **Yfinance documentation**. 2022. Disponível em: <<https://pypi.org/project/yfinance/>>. Acesso em: 12 de setembro 2022.

SAAD, M.; MOHAISEN, D. Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions. In: . [S.l.: s.n.], 2018. p. 704–709.

SIN, E.; WANG, L. Bitcoin price prediction using ensembles of neural networks. In: . [S.l.: s.n.], 2017. p. 666–671.

YAGA, D. et al. Blockchain technology overview. 06 2019.

yahoo. **Yahoo Finance - Stock Market Live, Quotes, Business and more**. 2022. Disponível em: <<https://finance.yahoo.com/>>. Acesso em: 22 de março 2022.

ZHENGYANG, W. et al. Prediction of cryptocurrency price dynamics with multiple machine learning techniques. In: . [S.l.: s.n.], 2019. p. 15–19. ISBN 978-1-4503-6323-5.

APÊNDICE A — DADOS DA BOLSA AMERICANA

Figura A.1 – Índice histórico da NASDAQ



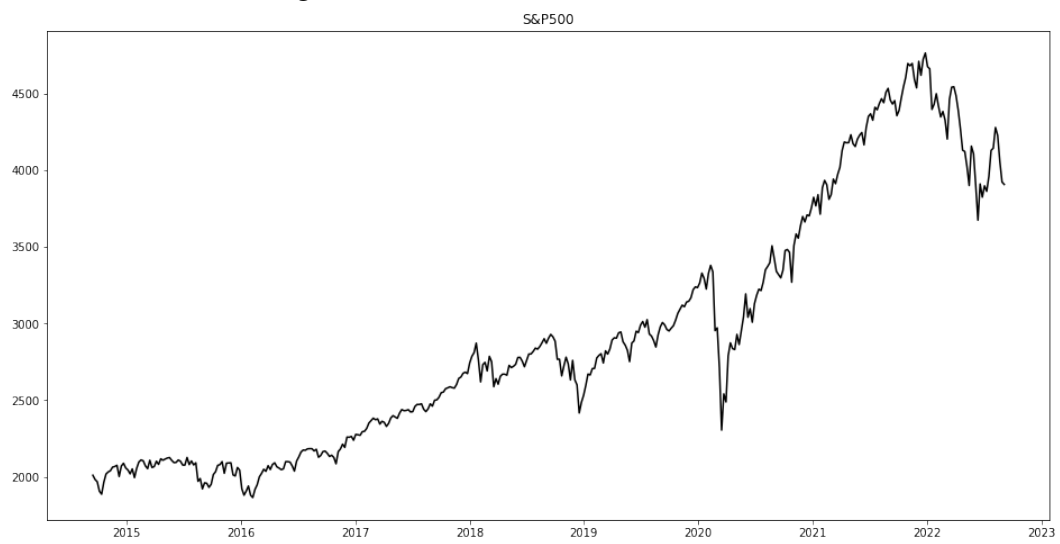
Fonte: Próprio Autor

Figura A.2 – Índice histórico da NYA



Fonte: Próprio Autor

Figura A.3 – Índice histórico do S&P500



Fonte: Próprio Autor

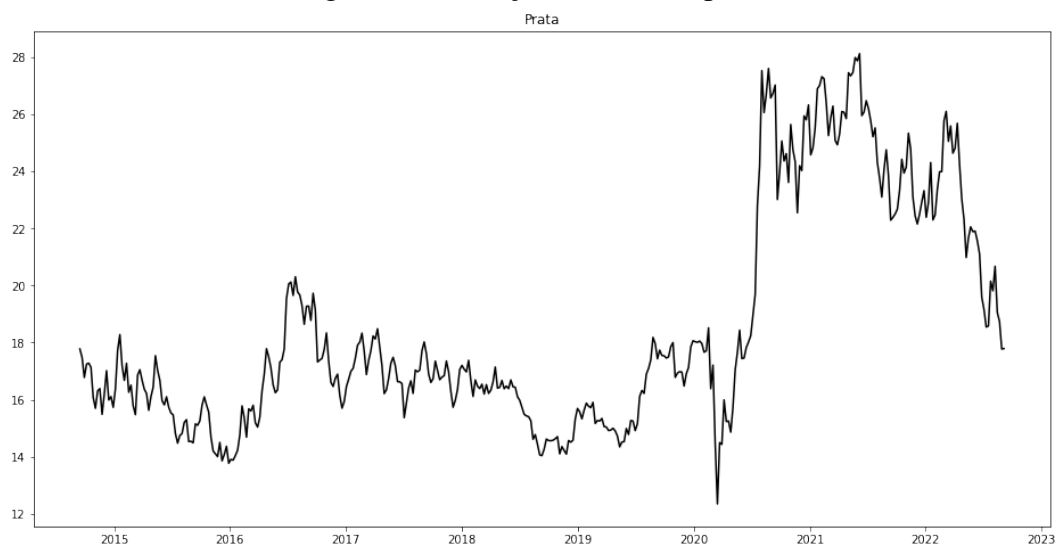
APÊNDICE B — *COMMODITIES*

Figura B.1 – Preço histórico do ouro



Fonte: Próprio Autor

Figura B.2 – Preço histórico da prata



Fonte: Próprio Autor