



COMPARAÇÃO DE DOIS MODELOS DE PREDIÇÃO DE PREÇO DE BITCOINS UTILIZANDO ANÁLISE TÉCNICA E LSTM

COMPARISON OF TWO BITCOIN PRICE PREDICTION MODELS USING TECHNICAL ANALYSIS AND LSTM

7. Inovação, Tecnologia E Empreendedorismo

Mateus Tenório Dos Santos, UEM, Brasil, ra99829@uem.br

Wagner Igarashi, UEM, Brasil, wigarashi@uem.br

Deisy Cristina Corrêa Igarashi, UEM, Brasil, dccigarashi@uem.br

Resumo

As formas de investimento e acúmulo de bens vêm se diversificando durante as últimas décadas. Entre as inúmeras possibilidades de investimento, algumas recebem destaque devido à grande volatilidade, grande fluxo de transações diárias e possibilidades de ganho por meio de especulações, como é o caso do Bitcoin, uma criptomoeda que atingiu uma valorização de 1300% no ano de 2017. O Bitcoin trata-se de um ativo passível de especulação, e pode ser analisado por várias técnicas aplicadas no mercado de ações, como é o caso da análise técnica e das redes neurais artificiais LSTM. Vários estudos têm sido desenvolvidos na predição do movimento de preços do Bitcoin utilizando estas técnicas, assim, este estudo tem como objetivo analisar o desempenho dos trabalhos desenvolvidos por Kikuti (2020) e Machado (2020), de modo a analisar o porquê eles apresentaram desempenhos diferentes. Durante o processo de análise dos resultados, verificou-se que o modelo desenvolvido por Machado (2020) teve um desempenho melhor do que o modelo desenvolvido por Kikuti. Com alterações em alguns parâmetros do modelo de Kikuti (2020) foi possível equipará-lo ao modelo proposto por Machado.

Palavras-chave: Deep Learning; Bitcoin; Análise Técnica; Inteligência Artificial.

Abstract

The forms of investment and accumulation of goods have been diversifying during the last decades. Among the numerous investment possibilities, some are highlighted due to their high volatility, large flow of daily transactions and possibilities of gain through speculation, such as Bitcoin, a cryptocurrency that reached a valuation of 1300% in 2017. Bitcoin is an asset subject to speculation, and can be analyzed by various techniques applied in the stock market, such as technical analysis and neural networks LSTM. Several studies have been developed in the prediction of Bitcoin price movement using these techniques, so this study aims to analyze the performance of the works developed by Kikuti (2020) and Machado (2020), in order to analyze why they presented performances many different. During the process of analyzing the results, it was found that the model developed by Machado (2020) performed better than the model developed by Kikuti. With changes in some parameters of the Kikuti model (2020), it was possible to equate it to the model proposed by Machado.

Keywords: Deep Learning; Bitcoin; Technical Analysis; Artificial Intelligence.

1. INTRODUÇÃO

A análise e predição de tendências e de risco financeiro trata-se de um assunto relevante na área da contabilidade e finanças desde o século passado, e apresenta crescimento nas análises e pesquisas realizadas nos últimos anos. Isso se ocorre devido ao crescimento dos mercados

financeiros juntamente ao aumento do número de fraudes e crises financeiras que ocorrem concomitantemente. Dessa forma, se observa é que diversos algoritmos, métodos, simulações e técnicas computacionais vêm sendo desenvolvidos e empregados na área financeira nos últimos anos (Grzelak & Oosterlee, 2011, Zhang & Yang (2019). De forma semelhante crescem os mercados financeiros de *criptomoedas*, seja para especulação ou como reserva de valor.

O Bitcoin surge em 2008 como uma resposta à crise financeira da época, tendo como principal ideia substituir a moeda física e retirar a dependência de intermediação de bancos nas operações financeiras. A primeira menção ao Bitcoin ocorreu em um artigo publicado por Satoshi Nakamoto (Nakamoto, 2008), descrevendo como funcionaria a moeda digital e o *blockchain*.

Contudo, o mercado de criptomoedas e em específico do Bitcoin é marcado pela volatilidade de preços em curtos períodos de tempo. Alguns exemplos dessa volatilidade podem ser percebidos em 2017 quando o Bitcoin valorizou 1.300% e uma queda abrupta do preço em maio de 2021, quando caiu quase 50% em relação a maio do ano anterior. Neste sentido, a utilização de técnicas estatísticas e modelos computacionais como redes neurais auxiliam a identificar momentos adequados para compra, venda e a manutenção de Bitcoins.

Segundo Murphy (1999), Vidotto, Migliato e Zambon (2009) a análise técnica é baseada no princípio de que os preços de ativos em um determinado momento futuro, são influenciados por seus preços passados, de modo que, é possível prever a tendência de movimentação desses preços em determinado período de tempo. Isto pode ser feito utilizando indicadores técnicos para indicar momentos oportunos para a realização de transações.

De modo complementar, diversos trabalhos na literatura têm destacado a utilização de técnicas computacionais na área de finanças, em especial o Bitcoin. Rane e Dhage (2019) realizam uma comparação entre sete modelos para a predição do preço do Bitcoin. Em seu trabalho, o modelo implementado conhecido como NARXS (*Non-linear Auto-Regressive with Exogenous Input Model*) demonstrou uma melhor obtenção de resultados. Em Aggarwal, Gupta, Garg e Goel (2019) estudos foram realizados com relação aos impactos dos fatores socioeconômicos na predição do preço do Bitcoin. No estudo, três diferentes abordagens foram utilizadas: CNN (*Convolutional Neural Network*), LSTM (*Long short Term Memory*) e GRU (*Gated Recurrent*), onde o modelo LSTM obteve melhores resultados na predição dos valores do Bitcoin.

Considerando o cenário apresentado em que se aplica indicadores técnicos e modelos computacionais utilizando redes neurais ou mesmo aprendizagem profunda no mercado de criptomoedas, este estudo tem como objetivo analisar e comparar os modelos de Kikuti (2020) e Machado (2020) para a predição do movimento de preços do Bitcoin, observando suas diferenças, similaridades e possíveis potenciais.

Uma contribuição deste estudo está no âmbito de demonstrar a interdisciplinaridade do processo de predição do movimento de preços do Bitcoin, que abrange concomitantemente os campos do mercado financeiro, da análise técnica e de modelos computacionais. O estudo, também, se justifica, considerando que o mercado de criptomoedas é recente, e que demanda de estudos aplicados previamente e da comparação das diversas técnicas de modelagem para a predição do movimento de preço.

2. BITCOIN

Em 2008, Nakamoto apresentou o Bitcoin, uma criptomoeda a fim de solucionar problemas como custo de mediação, possibilidade de transações irreversíveis, valor mínimo de transação, anonimidade da transação e garantia de impossibilidade de fraudes. Essa criptomoeda conseguiu atrair a atenção, principalmente na área de investimentos, por sua descentralização e autorregulagem, por meio de um sistema conhecido como *peer to peer*, ou P2P.

A utilização da arquitetura *peer to peer* permite que as transações envolvendo a *criptomoeda* sejam de acesso comum a todos os usuários, que podem legalmente verificar todas as transações realizadas utilizando o Bitcoin. Isso ocorre devido a tecnologia que funciona por trás do Bitcoin, o *Blockchain*, que também foi apresentado por Nakamoto (2008).

Existe uma infinidade de possibilidades e fins para a utilização do Bitcoin, porém ele passou a ser uma boa alternativa para investidores quando se trata de reserva de valor. Reserva de valor trata-se de um bem cuja característica é a de preservar seu poder de compra com o passar dos anos, e o Bitcoin possui esta característica uma vez que sua quantidade é limitada em 21 milhões. Considerando os fatores de que o Bitcoin é uma moeda autorregulada, descentralizada e que possui circulação limitada, sua valorização a longo prazo passa a ser positiva.

Sob outra ótica, muitos investidores consideram a moeda como uma boa maneira para lucrar por meio da especulação, uma vez que o mercado da *criptomoeda* é volátil e sofre inúmeras variações em um curto período de tempo. A especulação em ativos financeiros é a ação de comprar e vender ativos em curto ou médio prazo, visando o lucro. Para isso, esta abordagem costuma envolver a análise técnica de gráficos e de indicadores estatísticos.

3. ANÁLISE TÉCNICA

De acordo com Murphy (1999) a análise técnica consiste no estudo dos movimentos do mercado e tem como propósito a previsão de tendências de preços futuros. Portanto, busca-se por meio desta análise a identificação de padrões e tendências possíveis de serem exploradas em busca de melhores rentabilidades do que as da média do mercado. A análise técnica utiliza diversas ferramentas, entre as quais há o grupo dos indicadores técnicos.

Os indicadores técnicos permitem identificar pontos de entrada e saída do mercado, auxiliando os investidores na maximização do retorno de seu capital. Um aspecto importante dos indicadores técnicos é a objetividade por trás desses sistemas, pois é eliminada a possibilidade de interferência emocional na tomada de decisões, fazendo com que o operador aja conforme os sinais indicam. Murphy (1999) destaca três principais fontes de informações disponíveis e relevantes para esse tipo de análise: o preço, o volume e o interesse dos investidores.

Para Murphy (1999) a frase "*market actions discounts everything*" forma a base para a aceitação e compreensão de toda a teoria por trás da análise técnica. Dessa forma, o estudo do preço de determinado ativo se faz suficiente para predizer seu comportamento. A análise técnica se baseia na análise gráfica dos preços de um ativo e na utilização de indicadores estatísticos como ferramenta para a predição de momentos de compra, venda e manutenção desse ativo.

3.1 Indicadores Técnicos

Os analistas técnicos costumam utilizar dos indicadores estatísticos e financeiros com o objetivo de facilitar a identificação de determinados padrões nos ativos. Esses indicadores foram desenvolvidos para auxiliar na identificação de certas tendências de mercado em manter preços em alta, em baixa ou neutro.

São diversos tipos de indicadores que existem na análise técnica, sendo eles: indicadores de tendência, que buscam por padrões que indiquem a direção do mercado bem como momentos de reversão; os osciladores, que auxiliam na identificação de pontos de mudança com antecedência; e por fim, os indicadores de volume, que representam o volume transacionado.

3.1.1 Média móvel simples

Uma média móvel simples (MMS ou *simple moving average* – SMA) é uma série temporal composta de valores obtidos ao calcular a média de outra sequência de valores, sendo definida como uma forma de convolução matemática (Hyndman, 2011). Murphy (1999) destaca que, a MMS está entre os indicadores mais versáteis e mais utilizados. Isso se deve principalmente à sua facilidade de uso. Murphy (1999) cita também que enquanto a análise de gráficos é subjetiva e difícil de ser feita, MMS podem facilmente ser programadas em um computador. Técnicos podem discordar na leitura de um gráfico, mas não do sinal emitido por uma MMS.

O tamanho da MMS (a quantidade de períodos considerados) deve variar de acordo com o mercado. Uma MMS maior irá produzir um atraso para mudanças bruscas na tendência, mas em contrapartida produzirá um sinal mais conservador. Quanto maior a média, maior o atraso para se verificar uma nova tendência (Murphy, 1999).

A partir da MMS surgem outros indicadores. Por exemplo, a média móvel exponencial (MME ou *Exponential Moving Average* - EMA) que surgiu como alternativa para algumas defasagens deixadas pelas MMS. O que esse indicador faz é a atribuição de um peso maior aos preços mais recentes. O cálculo de uma MME ocorre da seguinte maneira:

$$\text{MME: } \{ \text{Fechamento} - \text{MME (dia anterior)} \} \times \text{fator} + \text{MME (dia anterior)}$$

onde: $\text{fator} = \frac{2}{n+1}$, e n é o número de períodos.

3.1.2 MACD (Moving average convergence divergence)

Outro indicador técnico que faz uso das médias móveis é o MACD (*Moving average convergence divergence*) que é amplamente utilizado por analistas e *traders*. De forma geral, o MACD transforma duas médias móveis em um indicador de momento, a partir da subtração da média móvel que possui maior período pela média móvel que possui menor período.

O MACD trabalha com três parâmetros: a linha do MACD, a linha do sinal e o histograma do MACD. A linha do MACD trata-se de uma MME de período de 12 dias subtraída de uma MME de período de 26 dias, para os preços de fechamento. A linha de sinal é definida por uma MME da linha MACD (geralmente de 9 períodos) e é utilizada para identificar pontos de virada. O histograma, é o resultado da subtração da linha do MACD e a linha de sinal. Quando a linha do MACD está acima da linha de sinal o histograma é positivo, do contrário é negativo.

3.1.3 Relative Strength Index

O *Relative Strength Index* (RSI) foi desenvolvido por Welles (1978) e tem como função avaliar a tendência de força de um ativo dada sua variação de preço. O RSI é um indicador que geralmente é utilizado em conjunto com outros indicadores, e pode indicar possíveis zonas de sobrecompra e sobrevenda. Costumeiramente o RSI é exibido como um gráfico que pode variar de 0 a 100. O cálculo do RSI é realizado por meio da seguinte fórmula:

$$RSI = 100 - \frac{100}{1+RS} \text{ onde: } RS = \frac{\text{Média de } n \text{ (fechamentosAlta)}}{\text{Média de } n \text{ (fechamentosBaixa)}}$$

Obtendo o valor do indicador, se faz possível a realização de algumas inferências a respeito do valor calculado: Entre 0 e 30 o RSI indica uma zona de sobrevenda. Indica boa oportunidade para compra, que deve ser evitada a realização de vendas no período; Entre 30 e 70 não há informação específica; Entre 70 e 100 o RSI indica uma zona de sobrecompra. Indica boa oportunidade para venda, que deve ser evitada a realização de compras no período.

3.1.4 Oscilador estocástico

Outro indicador utilizado nas operações de compra e venda de ativos é o oscilador estocástico. Trata-se de um indicador de volume idealizado por Lane (1984), o oscilador estocástico avalia a posição relativa do preço de fechamento comparada ao ponto mais baixo, para um número selecionado de dias, em relação à diferença entre o ponto mais alto e o ponto mais baixo (Migliato, Storani & Masili, 2008).

Sua utilização é caracterizada pelo uso de duas linhas que se movimentam de forma simultânea, a linha %K (estocástico bruto) e a %D (Migliato, Storani & Masili, 2008). A linha %K é dada pela divisão entre a diferença do preço de fechamento corrente e a menor mínima do período definido e a diferença do fechamento corrente e a maior máxima do período definido. A %D é dada pela média móvel de %K em dado período.

$$\%K = C - Ll(n) / Hh - Ll(n)$$

onde C = preço de fechamento, $Ll(n)$ = mínimo dos mínimos dos últimos n períodos,

$Hh(n)$ = máximo dos máximos dos últimos n períodos e n = período no qual serão definidos os valores de Ll e Hh .

Similar ao RSI, o oscilador estocástico foi concebido para variar entre 0 e 100%. Porém, são demarcadas duas linhas de referência em níveis de 20% e 80%. Se o valor do oscilador estiver abaixo da linha de 20%, pode-se inferir que o mercado está sobrevendido. Assim como, se o oscilador estiver acima da linha de 80%, o mercado está sobrecomprado.

4. REDES NEURAIAS

Mitchell (1997) afirma que, com um maior entendimento de como as máquinas podem aprender, é mais simples de perceber como funcionam as habilidades humanas de aprendizado. Para Goodfellow, Bengio e Courville (2016) o maior desafio é o desenvolvimento de soluções para problemas que apesar de intuitivos e relativamente simples para humanos, são complexos para uma máquina, como reconhecimento de imagem e voz.

Para solucionar esses problemas mais complexos, a adoção de redes neurais artificiais (RNA) tem-se mostrado útil. Uma RNA aprende a partir de um grande conjunto de dados que serve como exemplos para treino (Nielsen, 2015). Nielsen (2015) ainda destaca que há dois tipos de neurônios artificiais, os *perceptrons*, criados na década de 50 e 60 pelo cientista Frank Rosenblatt, e sua versão mais recente, chamada de *sigmoid neuron*.

O *Perceptron* é um modelo capaz de separar dados entre classes linearmente separáveis, configurado a partir de um algoritmo de aprendizagem capaz de adaptar os pesos internos do neurônio (Rauber, 2005). Os *sigmoid neurons* recebem uma singela mudança no que é dado de entrada e saída, onde os valores de entrada e saída deixam de ser valores binários para ser valores decimais entre 0 e 1. Dessa forma, os valores de saída apresentados pelos *sigmoid neurons* podem ser interpretados como uma probabilidade.

4.1 Arquitetura de redes

De modo geral há três tipos de arquiteturas de RNAs: redes alimentadas à frente com camada única, redes alimentadas à frente com camadas múltiplas e redes recorrentes (Haykin, 2007). Dentre as diversas arquiteturas, as redes neurais recorrentes (RNR) se distinguem das demais por possuírem ao menos um laço de repetição. Trata-se de uma rede em que a saída de cada neurônio serve como a entrada dos demais. Esse tipo de rede pode ou não possuir a auto realimentação, ou seja, o neurônio realimenta sua entrada com sua própria saída. Também podem existir camadas ocultas na estrutura da rede. A presença de laços de realimentação na rede tem um impacto significativo na capacidade de aprendizagem e desempenho da rede (Haykin, 2007).

4.2 Long short term memory

Um dos modelos de aprendizado profundo mais utilizados quando se trata de previsão de preços de qualquer produto ou ativo financeiro é a RNR do tipo LSTM (*Long Short Term Memory*) (Kikuti, 2020). Wang e Gao (2018) citam que o LSTM, modelo desenvolvido por Hochreiter e Schmidhuber no ano de 1977 com o intuito de resolver problemas de explosão e desaparecimento de gradientes. Sua estrutura interna é capaz de armazenar padrões que são repetidos inúmeras vezes e ao mesmo tempo permite a realização do truncamento de informações quando essas são irrelevantes ou desnecessárias.

O modelo LSTM processa as variáveis de entrada de forma incremental, adicionando novos dados em um único compartimento de memória, possuindo portões cuja função é controlar o alcance de qual novo dado deve ser armazenado e quais dados antigos devem ser apagados. Diferentemente das redes neurais recorrentes, em que é produzida uma lista de estados durante o processamento, no LSTM, o próximo estado computado depende exclusivamente dos estados $h_t \dots h_{t-1}$. Portanto, a rede realiza a sua atualização de forma Markoviana, onde um estado depende exclusivamente do estado anterior.

Nas subseções a seguir serão apresentados os desenvolvimentos dos trabalhos desenvolvidos por Kikuti (2020) e Machado (2020), a modo de entender melhor os modelos aplicados para realizar a previsão do preço do Bitcoin.

4.5 Trabalho Desenvolvido por Kikuti (2020)

Kikuti (2020) desenvolveu três modelos para a predição do preço do Bitcoin. Foi desenvolvido um modelo que utiliza LSTM, um que utiliza apenas indicadores técnicos como forma de predição e um modelo híbrido, que se trata de uma mescla das duas abordagens anteriores. Serão analisados os modelos LSTM e o híbrido.

Para o desenvolvimento do modelo LSTM, foi utilizado o ambiente de desenvolvimento *Google Colaboratory*, linguagem *python* e as bibliotecas TensorFlow, Keras, Python t.a., Numpy e Pandas. Além disso, Kikuti (2020) utilizou da base de dados fornecida pela API do Yahoo Finance, que contém dados referentes a cotações diárias do Bitcoin, do período de 16 de setembro de 2014 a 31 de dezembro de 2019. Dentre as informações que a base de dados fornece estão: Preço de abertura, preço de fechamento, maior preço atingido no dia, menor preço atingido no dia e o volume de transações realizadas.

Para a criação do modelo LSTM foi utilizado de um modelo sequencial fornecido pela biblioteca *Keras*, modelo que é apropriado para uma pilha simples de camadas. Sua primeira camada é composta por 128 camadas (KIKUTI, 2020) do tipo *LSTM* que são fornecidos pela biblioteca do *TensorFlow*, padrão que se repete para a segunda e terceira camada. Diferentemente das camadas anteriores, a quarta camada é composta por nós do tipo *Dense*, que também são fornecidos pela biblioteca *Keras*, e funciona como um nó comum em uma rede neural, onde como parâmetro é passada uma função de ativação do tipo linear. Por fim, a última camada possui apenas 2 nós e recebe como parâmetro uma função de ativação do tipo *softmax*. Essa função de ativação é comumente utilizada na última camada, pois seu resultado pode ser interpretado como um vetor de distribuição de probabilidade.

O modelo híbrido mescla a utilização dos indicadores técnicos com os resultados obtidos pela predição do modelo LSTM. A tomada de decisão é realizada da seguinte maneira: O modelo híbrido realiza a verificação da saída do modelo de aprendizado de máquina para realizar uma tomada de decisão de compra. Para a tomada de decisão de venda, o modelo verifica o indicador técnico MACD, onde caso esse aponte para um período de alta o ativo é vendido, caso contrário ele é mantido na carteira.

Essa decisão de utilização de somente um indicador técnico se justificou pelos resultados apresentados por Kikuti (2020), onde foi verificado que, à medida que o uso de indicadores aumenta, o modelo torna-se mais restritivo em relação a compras e a vendas, impactando na rentabilidade do modelo.

4.6 Trabalho Desenvolvido por Machado

Machado (2020) utilizou dados extraídos da base de dados *Yahoo Finance* para coletar dados para a utilização de indicadores de análise técnica e para a rede Neural. Machado (2020) realizou a implementação de dois modelos, um utiliza da implementação do indicador técnico de médias móveis, e o segundo modelo utiliza da implementação de uma LSTM que utiliza de médias móveis para seu treinamento. Neste estudo analisa o modelo LSTM.

Machado (2020) optou por utilizar de uma rede neural sequencial multicamadas LSTM, em que cada neurônio de uma camada envia seu sinal para cada neurônio da camada posterior, tratado

por um peso. Os modelos utilizados em seu trabalho possuem 12 entradas (cada entrada sendo o sinal emitido por uma das combinações de médias móveis utilizadas anteriormente), três camadas internas e um único neurônio de saída.

Com relação ao tamanho das camadas internas, duas configurações distintas foram utilizadas a fim de verificar o quão válido é para um usuário construir uma rede neural muito grande buscando um melhor resultado, sendo elas: (I) 144, 72, 36; e (II) 9, 9, 9.

Outras características da rede neural foram parametrizadas, sendo elas: Épocas de treino (quantidade de vezes em que a rede é treinada pela mesma massa de dados); Taxa de aprendizado (a taxa de aprendizado do otimizador aumenta ou diminui o impacto que a função de erro tem sobre o processo de aprendizado); Função de perda (testes utilizaram três funções de perda diferentes – Erro quadrático médio (MSE); Regressão de Poisson e Divergência de *Kullback-leibler* (KLD)); Dropout (a utilização de Dropout em cada camada pode influenciar na forma em que ela se ajusta durante os testes para evitar o sobre ajuste do modelo).

5. METODOLOGIA

Esta pesquisa teve como primeira etapa o levantamento de referencial teórico necessário ao entendimento de conceitos de *criptomoedas*, indicadores técnicos e redes neurais. Também foi necessária uma análise de estudos aplicados de predição de preços de ativos em *criptomoedas*, onde foram pesquisadas as tecnologias utilizadas nas aplicações desenvolvidas por Kikuti (2020) e Machado (2020), a fim de reproduzir as aplicações propostas pelos autores e comparar os resultados encontrados, buscando delinear combinações de indicadores, parametrizações e modelos mais adequados para a predição do preço de *criptomoedas*.

Na segunda etapa foram realizadas simulações preliminares para o entendimento dos modelos criados por Kikuti (2020) e Machado (2020). Para tanto, foi utilizada a linguagem de programação *Python*, e bibliotecas como: Interface de Programação de Aplicações (IPA ou *Application Programming Interface* – API) do Yahoo Finance – para acesso a dados históricos da criptomoeda Bitcoin; API Python Technical Analysis – para cálculo dos indicadores técnicos; a biblioteca Pandas, para a importação dos dados de cotações referentes ao Bitcoin em formato .csv; Tensorflow e Keras - frameworks de aprendizado profundo para treinamento de redes neurais profundas.

Os experimentos foram executados em uma máquina equipada com processador i5 11400f (6 núcleos, 12 threads, 4.4Ghz e 12M de cache), 32Gb de Memória RAM e 1Tb SSD m2 para armazenamento que utiliza como sistema operacional o Windows 11 na versão 21H2.

Na terceira etapa foram feitos ajustes aos modelos de Kikuti para identificar quais parâmetros poderiam melhorar seu desempenho. Na quarta etapa foram realizados experimentos a partir dos modelos e feita a análise dos resultados de acordo com as métricas de *F1-score*, acurácia, e simulação de retorno financeiro.

Com relação as métricas estudadas e analisadas nesse estudo, foram utilizadas métricas cujo foco está nos ganhos e lucros gerados pelos diferentes tipos de modelos implementados. Não somente, métricas específicas para a aprendizagem de máquina como quantidade de compras e

vendas, *F1-Score* e acurácia (*accuracy*) também tiveram foco nesse estudo, uma vez que as mesmas podem nos dar dicas de que o treinamento foi bom ou não.

A quantidade de compra e vendas realizadas no período de um ano foi escolhida como uma métrica de análise pelo fato de que esses valores transparecem se o treinamento foi suficiente, uma vez que valores altos podem nos indicar de que a rede neural obteve sucesso em prever certos momentos de compra e venda, enquanto que valores próximos ou iguais a 1, ou então valores iguais a 0 nos mostram que a LSTM aprendeu pouco (*underfitting*) ou se ajustou demais aos dados (*overfitting*).

A métrica de lucro bruto final foi utilizada de duas formas, tanto em porcentagem quanto em dólar. Ela foi utilizada para verificar o retorno dos modelos desenvolvidos, sendo ela a principal métrica para a comparação de qual modelo implementado teve o melhor desempenho.

Como métricas específicas para o trabalho de Kikuti (2020), foram analisadas a *F1Score* que é calculada por meio da biblioteca *sklearn* e a acurácia fornecida e calculada pela biblioteca *Keras*. A métrica acurácia cria duas variáveis locais, uma sendo o total e a outra um contador, que são utilizadas para calcular a frequência que a predição corresponde ao esperado. Dessa forma, seu cálculo se dá pela divisão do contador pelo total, sendo que o contador representa a quantidade de vezes que o valor predito foi igual ao valor real.

A métrica *F1-Score* possui seu cálculo com base em duas outras métricas. A precisão e o *recall*. A precisão é utilizada para determinar dentre os casos analisados pelo modelo como positivos, quantos deles realmente foram positivos. O *recall* é utilizado para calcular a proporção de casos positivos que foram identificados corretamente. O valor ótimo para o *F1-Score* é alcançado quando tanto a precisão quanto o *recall* são 100%. Além disso, caso um dos seja igual a 0, o *F1-Score* também será 0.

Essas duas últimas métricas foram utilizadas para assegurar que o treinamento do modelo obtivesse bons resultados. Durante os experimentos, foi observado que valores muito baixos de alguma dessas métricas nos garantiam que o teste não obteria bons resultados, enquanto que valores um pouco acima de 0,5 de ambos garantiriam resultados onde ao menos o comportamento da rede neural não foi um *buy and hold*.

6. ANÁLISE DE RESULTADOS

Os experimentos desta pesquisa buscaram compreender melhor os trabalhos desenvolvidos por Kikuti (2020) e Machado (2020). Por isso, foram realizadas as mesmas simulações que os autores originalmente realizaram, a fim de coletar informações sobre os simuladores, verificar as similaridades e apontar diferenças, bem como entender o que causou a diferença nos resultados. Os experimentos foram similares aos aplicados por Machado (2020) para testar seu modelo LSTM. Nota-se que a LSTM de Machado (2020) é relativamente mais simples. Dessa forma, os testes realizados buscaram manter alguns parâmetros da LSTM desenvolvida por Kikuti (2020) intactos, tentando obter melhores resultados a partir do ajuste de outros parâmetros. Entre os parâmetros escolhidos para tentar otimizar os modelos de Kikuti (2020), foram escolhidos os seguintes: taxa de aprendizado, Decay, épocas, tamanho do lote, função de ativação de camadas densas.

Os dois tipos de modelos de Kikuti (2020) analisados são o de rede neural LSTM, e o modelo híbrido, uma LSTM para compra e o indicador técnico MACD para venda. O modelo analisado de Machado (2020) é um modelo de rede neural LSTM.

Na sequência, são apresentados os resultados obtidos pelos experimentos nos períodos avaliados tanto pelo trabalho desenvolvido por Kikuti (2020), quanto pelo trabalho desenvolvido por Machado (2020). Para o primeiro experimento (utiliza-se o período avaliado por Kikuti (2020)) foram realizados diversos testes de parâmetros, para os experimentos nos períodos avaliados por Machado (2020) foram utilizados os modelos que obtiveram melhores resultados no primeiro experimento. Os períodos trabalhados em todas as análises foram os seguintes: Período 1 - 01/01/2020 à 31/12/2020 (Kikuti, 2020); Período 2 - 01/07/2019 à 30/06/2020 (Machado, 2020); Período 3 - 01/07/2018 à 30/06/2019 (Machado, 2020).

A seguir são especificados os nomes dos modelos e suas características. Esses modelos foram os melhores para essa simulação, não necessariamente em montante final, mas também na quantidade de compra e venda realizado pela rede neural, o que pode ser um indício de melhor aprendizado. Modelos que realizaram um número pequeno de compras e vendas ou obtiveram valores de montante final muito baixos não entraram nesse primeiro período.

Seguem os modelos modificados de Kikuti (de 1 a 5) e seu modelo original (6);

1. Modelo 1: *Epochs* = 16, *Batch size* = 32, sem função de ativação, 0.3LR 0.03 *Decay*;
2. Modelo 2: *Epochs* = 16, *Batch size* = 32, sem função de ativação, 0.3LR 0.3 *Decay*;
3. Modelo 3: Sem *Batch Normalization* e sem *Decay*;
4. Modelo 4: *Batch size* = 32 e sem *Decay*;
5. Modelo 5: Sem *Decay*;
6. Original: Igual ao modelo desenvolvido por (KIKUTI, 2020).

Importante deixar explicar que os modelos em que algumas características não foram mudadas, seguem o padrão implementado por Kikuti (2020), sendo eles os seguintes: *Epochs* = 100; *Batch_size* = 128; Função de ativação nas duas camadas densas; *Learning rate* = 0.001; *Decay* = 1e-6; *Batch Normalization* após cada camada simples.

6.1 Resultados obtidos no Período 1

A tabela 1 apresenta o resultado obtido pelos cinco modelos híbridos juntamente com os resultados obtidos pelo modelo original desenvolvido por Kikuti (2020) no período 1.

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	425640,69	7	7	325,64%
2	100000	304165,60	62	62	204,00%
3	100000	246877,33	69	69	146,88%
4	100000	338724,73	44	44	238,73%
5	100000	369717,09	46	46	269,72%

6	100000	337086,55	32	32	237,00%
---	--------	-----------	----	----	---------

Tabela 1 – Resultado dos modelos de LSTM de Kikuti (2020) no período 1.

Na tabela 2 são comparados os resultados obtidos pelos cinco modelos híbridos juntamente com os resultados obtidos pelo modelo original.

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	435.412,07	2	2	335,41%
2	100000	385.027,13	28	28	285,00%
3	100000	356.765,41	30	30	256,77%
4	100000	493.036,98	24	24	393,04%
5	100000	584.792,34	23	23	484,79%
6	100000	401.726,26	14	14	301,73%

Tabela 2 – Resultado dos modelos Híbridos de Kikuti (2020) no período 1.

Como observado a partir das tabelas 1 e 2 foi possível obter bons ganhos, a partir de um volume suficiente de compras e vendas. O que significa que não somente a LSTM foi capaz de aprender bem o comportamento dos valores do Bitcoin, como também obter um bom ganho. A Tabela 3 representa os modelos da rede neural desenvolvidas por Machado (2020) que obtiveram melhores resultados nesse período analisado.

Modelo	Lucro
16 0.2 kl_divergence In20_10_5	541,62%
2 0.01 kl_divergence In20_10_5	408,46%
4 0.1 kl_divergence In20_10_5	370,09%
4 0.01 kl_divergence In20_20_None	369,88%
4 0.1 kl_divergence In20_None	367,01%

Tabela 3 – Resultado dos melhores modelos de RNA de Kikuti (2020) no período 1.

Ao analisar a rede neural desenvolvida por Machado (2020), tem-se uma rede neural mais simples, sem função de ativação e apenas uma camada densa para saída dos dados, enquanto que a rede neural desenvolvida por Kikuti (2020) possui em sua composição uma camada densa “escondida” e uma para saída de dados, ambas com função de ativação. Isso torna a rede neural mais complexa e que poderia obter melhores resultados. Entretanto, o modelo de rede neural desenvolvido por Machado (2020) funciona como um modelo híbrido, uma vez que o conjunto de dados utilizado para treinar a rede trata-se de um modelo que possui 12 entradas, onde cada entrada é o sinal emitido por uma das combinações de médias móveis utilizadas para realizar a análise técnica. Isso significa que, para meios de comparação, deve-se comparar valores obtidos pelo modelo híbrido, e não somente pelo modelo LSTM.

Dessa forma, ao analisar os resultados, observa-se que somente com o ajuste dos parâmetros foi possível alcançar valores próximos aos obtidos pelos modelos treinados por Machado

(2020) em seu trabalho, sem a necessidade de realizar muitas alterações na estrutura do modelo original de Kikuti (2020).

Ao analisar *FI-Score* e *accuracy*, notou-se que um valor pequeno para *accuracy* pode significar que o modelo de rede neural obterá um número pequeno de compras e vendas, como é o caso do Modelo 1, que obteve o menor número de compras e vendas e a menor *accuracy* no treinamento. Também se verificou que um maior *FI-Score* não necessariamente produz um lucro maior, que foi o caso do modelo 1.

6.2 Resultados obtidos no Período 2

Na tabela 4 apresenta o resultado obtido pelos cinco modelos apresentados anteriormente juntamente com os resultados obtidos pelo modelo original, que é o modelo desenvolvido por Kikuti (2020).

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	100.000,00	0	0	0,00%
2	100000	108.237,83	2	2	8,24%
3	100000	61.378,41	55	55	-38,60%
4	100000	80.312,57	28	28	-19,68%
5	100000	133.456,89	61	61	33,46%
6	100000	182.776,47	50	50	82,78%

Tabela 4 – Resultado dos modelos LSTM de Kikuti (2020) no período 2.

Na tabela 5 são comparados os resultados obtidos pelos cinco modelos híbridos juntamente com os resultados obtidos pelo modelo original.

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	84.597,91	1	1	-15,40%
2	100000	133.905,26	2	2	33,91%
3	100000	75.185,80	28	28	-24,82%
4	100000	88.387,63	13	13	-11,70%
5	100000	161.878,44	36	36	61,88%
6	100000	17.273,75	26	26	72,74%

Tabela 5 – Resultado dos modelos Híbridos Kikuti (2020) no período 2.

Diferente do primeiro experimento, nesse nem todos os modelos obtiveram bons resultados. As tabelas 4 e 5 mostram que muitos modelos obtiveram valores de ganho abaixo de 100%, ou seja, o valor de montante final representa um valor menor do montante inicial. A Tabela 6 representa os modelos da rede neural desenvolvida por Machado (2020) que obtiveram melhores resultados nesse período analisado.

Modelo	Lucro
--------	-------

IIP-2-0.001	87.39%
IP-64-0.01	85.48%
IIM-1-0.001	73.94%
IP-64-0.25	64.17%
IK-1-0.25	64.15%
IIP-32-0.001	63.37%

Tabela 6 – Resultado dos melhores modelos LSTM de Machado (2020) no período 2.

Entretanto, tanto o trabalho original desenvolvido por Kikuti (2020), quanto a rede neural que não possuía *decay* na taxa de aprendizado conseguiram os melhores resultados. Isso mostra que a técnica de *Batch normalization* combinada com um número maior de *Batch size* e um número maior de *epochs* tornou o aprendizado mais efetivo. Não somente, a taxa de aprendizado ser maior não auxiliou a ter melhores resultados, algo que pode ser visto, também, nos resultados obtidos por Machado (2020), modelo no qual as menores taxas de aprendizado possuíram melhores resultados.

Analisando os valores calculados de *F1-score* e *Accuracy* dos experimentos verificou-se que quanto menor o valor de *accuracy*, menor é a quantidade de compras e vendas. Outro fato observado é que, quando o valor de *F1-score* é 0 o modelo 1 não realizou nenhuma compra e venda.

6.3 Resultados obtidos no Período 3

A tabela 7 apresenta o resultado obtido pelos cinco modelos apresentados anteriormente juntamente com os resultados obtidos pelo modelo original desenvolvido por Kikuti (2020). Da mesma forma, a tabela 8 compara os resultados obtidos pelos cinco modelos híbridos juntamente com os resultados obtidos pelo modelo original. A tabela 9 apresenta os resultados dos modelos LSTM desenvolvidos por Machado (2020) que obtiveram melhores resultados nesse período analisado.

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	100.000,00	0	0	0,00%
2	100000	95.787,80	70	70	-4,21%
3	100000	29.527,58	38	38	-70,47%
4	100000	76.384,93	79	79	-23,62%
5	100000	29.488,94	48	48	-70,51%
6	100000	27.435,57	43	43	-72,57%

Tabela 7 – Resultado dos modelos LSTM de Kikuti (2020) no período 3.

É possível observar por meio das Tabelas 7 e 8 que, nesse caso, apresentaram resultados parecidos com o primeiro experimento, em que os modelos com maior taxa de aprendizado obtiveram os melhores resultados. Apesar disso, o modelo 1 da Tabela 8 não mostrou aprendizado, pois não realizou novas transações de compra e venda.

Modelo	Montante Inicial	Montante Final	Compras	Vendas	Lucro
1	100000	163.544,91	1	1	63,54%
2	100000	109.936,52	39	39	9,94%
3	100000	95.500,97	21	21	-4,50%
4	100000	114726,18	40	40	14,73%
5	100000	40173,62	25	25	-59,83%
6	100000	59750,23	24	24	-40,25%

Tabela 8 – Resultado dos modelos Híbridos no período 3.

Nota-se que nenhum dos modelos chegou próximo dos valores obtidos pelos modelos desenvolvidos por Machado (2020). Nesse caso, pode-se atribuir essa diferença a abordagem utilizada pelo autor, uma vez que os valores de entrada são os sinais das 12 médias móveis que o mesmo utilizou para realizar a análise técnica. Dessa forma, os modelos treinados por Machado (2020) possuem certa vantagem, uma vez que a mesma possui um conhecimento a mais advindo dos indicadores utilizados pelo autor para realizar a análise técnica.

Modelo	Lucro
IK-32-0.3	243.25%
IK-128-0.25	236.97%
IP-128-0.05	233.63%
IK-32-0.25	232.65%
IM-32-0.15	229.30%
IK-32-0.15	227.94%

Tabela 9 – Resultado dos melhores modelos de Redes Neurais de Machado (2020) no período 3.

Conforme discutido nos experimentos anteriores, foi observado novamente a correlação entre *accuracy* e o número de compras e vendas realizadas. O modelo cujo valor de *accuracy* está próximo ou menor que 0,4, é responsável por obter um pequeno número de compras e vendas.

7. CONCLUSÃO

Esta pesquisa propôs a realização de experimentos com dois trabalhos de predição de preço do Bitcoin previamente realizados por Kikuti (2020) e Machado (2020), de modo a identificar quais foram os indicadores, parâmetros e modelos de aprendizado de máquinas utilizados, de modo a identificar diferenças no desempenho de cada um.

Para que os modelos pudessem ser analisados da melhor forma possível, uma revisão bibliográfica dos tópicos relacionados a cada modelo desenvolvido pelos alunos foi realizada: Bitcoin, análise técnica e redes neurais artificiais. Com relação a análise técnica, os indicadores identificados foram a média móvel, MACD, RSI e Oscilador estocástico. Em relação às métricas utilizadas para a análise dos modelos foram *FI-Score*, *Accuracy* e resultado financeiro a partir de simulações de investimento com um Montante inicial e um final.

Referente aos resultados obtidos nesta pesquisa, conclui-se que os resultados foram satisfatórios, considerando que os modelos apresentaram certa otimização dos resultados. Não é possível garantir que determinado modelo obtenha os melhores resultados em todos os experimentos realizados, porém em todos os experimentos foi possível observar ganho nos rendimentos. Durante os experimentos, foi possível concluir que uma mudança no conjunto de dados pode trazer benefícios para o treinamento de redes neurais, sendo o trabalho desenvolvido por Machado (2020) um exemplo disso.

Pode-se concluir, também, que a rede neural consideravelmente mais simples do modelo de Machado (2020) do que a desenvolvida por Kikuti (2020) conseguiu obter melhores resultados. Dessa forma, otimizações foram aplicadas no trabalho de Kikuti (2020) para obter resultados próximos, assim o modelo que retirava o parâmetro de *Decay* da taxa de aprendizado se sobressaiu em vários experimentos. Foi possível observar que modelos que com baixa acurácia tinham maior probabilidade de efetuar um menor número de compras e vendas, e que modelos com *F1-Score* igual à zero podem ser descartados, uma vez que isso representa que eles não realizam nenhuma compra ou venda durante a simulação.

Outra conclusão é de que o indicador MACD é um bom indicador por si só, e que a combinação de dois MACD um para compra e um para venda pode ser uma alternativa, mas que nem todos os cenários possíveis podem vir a trazer os melhores resultados, uma vez que a combinação com os indicadores RSI e Oscilador estocástico juntamente com os dois MACD's obteve melhores resultados. Entretanto, o indicador de médias móveis utilizado por Machado (2020) se mostrou mais versátil e obteve melhores resultados em todos os cenários dos experimentos.

Como possíveis trabalhos futuros tem-se a exploração de novas técnicas para aprendizado de máquina ou a utilização de diferentes indicadores técnicos juntamente com os modelos de aprendizado de máquina, que se mostraram superiores em resultado quando comparados aos modelos de LSTM e análise técnica.

REFERÊNCIAS

- Aggarwal, A., Gupta, I., Garg, N. & Goel, A. (2019). Deep learning approach to determine the impact of socio economic factors on Bitcoin price prediction. *2019 Twelfth International Conference on Contemporary Computing (IC3)*, p. 1–5.
- Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT press.
- Grzelak, L. A & Oosterlee, C. W. (2011). On the heston model with stochastic interest rates. *SIAM Journal on Financial Mathematics*, SIAM, v. 2, n. 1, p. 255–286.
- Haykin, S. *Redes neurais: princípios e prática*. São Paulo: Bookman Editora, 2007.
- Hyndman, R. J. (2011). Moving Averages. In: Lovric, M. (eds) *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_380.
- Kikuti, C. (2020). *Aplicação de uma técnica de aprendizagem profunda e indicadores técnicos na previsão do momento de compra e venda do Bitcoin*. Monografia de Trabalho de Conclusão de Curso, Universidade Estadual de Maringá.

- Lane, G. C. Lane's stochastics. (1984). *Technical Analysis of Stocks and Commodities*, v. 2, n. 3, p. 80.
- Machado, H. M. *Análise de técnicas de previsão de movimento do mercado de Bitcoin utilizando uma técnica de Aprendizado de Máquina*. Monografia de Trabalho de Conclusão de Curso, Universidade Estadual de Maringá, Maringá, 2020.
- Migliato, A. L. T., Storani, K. & Masili, M. Osciladores, padrões de candlesticks e variações de preços de ativos financeiros: Um estudo preliminar sobre a relação entre esses fatores no mercado de ações. *Multiciência*, São Carlos, v. 9, p. 291–301, 2008.
- Murphy, J. J. *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. EUA: Editora Prentice Hall Press, 1999.
- Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, p. 21260.
- Nielsen, M. A. (2015). *Neural networks and deep learning*. San Francisco: Determination press San Francisco, CA.
- Rane, P. V. & Dhage, S. N. (2019). Systematic erudition of Bitcoin price prediction using machine learning techniques. In: IEEE. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, p. 594–598.
- Rauber, T. W. (2005). Redes neurais artificiais. *Universidade Federal do Espírito Santo*, v. 29, 2005.
- Mitchell, T. M. (1997). *Machine Learning*. Alemanha: McGraw-Hill.
- Vidotto, R. S., Migliato, A. L. T. & Zambon, A. C. (2009). O moving average convergence-divergence como ferramenta para a decisão de investimentos no mercado de ações. *Revista de Administração Contemporânea*, SciELO Brasil, v. 13, p. 291–309.
- Wang, C. & Gao, Q. (2018). High and low prices prediction of soybean futures with lstm neural network. In: IEEE. *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, China, p. 140–143.
- Welles, J. W. (1978). *New concepts in technical trading systems*. Indiana: Editora Trend Research.
- Zhang, Y. & Yang, S. (2019). Prediction on the highest price of the stock based on pso-lstm neural network. In: *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, pp. 1565-1569, doi: 10.1109/EITCE47263.2019.9094982.