



ANÁLISE TÉCNICA E REDES NEURAIAS NA PREDIÇÃO DE PREÇO DE BITCOINS

TECHNICAL ANALYSIS VERSUS NEURAL NETWORKS TO BITCOIN PRICE PREDICTION

ÁREA TEMÁTICA: 6 FINANÇAS

Wagner Igarashi, Universidade Estadual de Maringá, Brasil, wigarashi@uem.br

Henrique Misael Machado, Universidade Estadual de Maringá, Brasil, ra99660@uem.br

Deisy Cristina Corrêa Igarashi, Universidade Estadual de Maringá, Brasil, dccigarashi@uem.br

Resumo

Em 2007, o Bitcoin (BTC) foi introduzido como a primeira moeda digital mundial descentralizada. O BTC vem apresentando um crescimento em seu valor de mercado, vem sendo utilizado em investimentos especulativos, e, portanto, seria possível utilizar as mesmas técnicas já aplicadas no mercado de ações. Foram identificadas algumas estratégias que obtiveram um resultado positivo, inclusive em momentos de queda como o período entre julho de 2019 e junho de 2020. Neste cenário, este estudo visa analisar a aplicação de uma rede neural para realizar predições do movimento de preço do Bitcoin em contraponto e uso da análise técnica. O primeiro modelo construído consistiu em um agente que verificava o sinal emitido por determinados indicadores técnicos, e simulava a compra ou venda de BTCs. Estes experimentos obtiveram bons resultados, e também forneceram insumos para a segunda parte do trabalho: uma rede neural construída para tomar decisões com base em indicadores técnicos testados na etapa anterior. Os resultados das redes neurais foram mais positivos do que os resultados de cada um dos agentes de indicadores técnicos, obtendo até 80% de aumento do montante em um período em que o BTC perdeu 15% de seu valor de mercado.

Palavras chave: análise técnica, inteligência artificial; Bitcoin; predição financeira

Abstract

In 2007, Bitcoin (BTC) was introduced as the world's first decentralized digital currency. BTC has shown a growth in its market value, has been used in speculative investments, and therefore it would be possible to use the same techniques already applied in the stock market. Some strategies that obtained a positive result were identified, including in moments of decline such as the period between July 2019 and June 2020. In this scenario, this study aims to analyze the application of a neural network to predict the price movement of Bitcoin in counterpoint and use of technical analysis. The first model built consisted of an agent that verified the signal emitted by certain technical indicators, and simulated the purchase or sale of BTCs. These experiments obtained good results, and also provided inputs for the second part of the work: a neural network built to make decisions based on technical indicators tested in the previous stage. The results of the neural networks were more positive than the results of each of the agents of technical indicators, obtaining up to 80% increase in the amount in a period when BTC lost 15% of its market value.

Keywords: technical analysis; artificial intelligence; Bitcoin, financial prediction

1. INTRODUÇÃO

A especulação de preços tem tornado a bolsa de valores um mercado lucrativo e ao mesmo tempo carente de profissionais capacitados. A previsibilidade do valor de uma ação é possível com uma margem razoável de segurança levando em consideração a utilização de métodos próprios para especulação. Elder (1993) explica que especular no mercado financeiro a partir da previsão do movimento do mercado é possível a partir de análise técnica ou fundamentalista. Murphy (1999), entretanto, considera a análise técnica como uma função de probabilidade, uma vez que a reação da massa segue padrões (desconsiderando a negociação individual, e considerando apenas a ação coletiva da população que o fará). Para tanto, a análise técnica utiliza diversos indicadores para verificar a variação de preço do mercado (Murphy, 1999).

Dada esta natureza matemática e estatística da análise técnica, que prevê padrões nos indicadores para tentar prever o futuro preço das ações, tem-se modelos de redes neurais como um possível candidato a automação da tarefa, visto que são sistemas eficientes em aprender padrões. Redes neurais são sistemas de inteligência artificial inspirados na biologia dos cérebros animais. Estes sistemas são capazes de aprender tarefas para as quais não foram programados, mas sim ensinados por meio da apresentação de padrões e seus resultados. Para tal aprendizado são utilizadas técnicas de retropropagação, desenvolvidas por meio da aplicação de programação dinâmica ao problema de aprendizado artificial.

Todavia, sob o ponto de vista especulativo, há várias possibilidades de ativos que poderiam ser escolhidos para se estudar e aplicar a análise técnica e redes neurais. Uma delas seria o estudo da variação de preços do Bitcoin (BTC) no mercado de criptomoedas. Deste modo, esta pesquisa visa analisar a aplicação de uma rede neural para realizar previsões do movimento de preço do Bitcoin em contraponto e uso da análise técnica. Destaca-se que o mercado de criptomoedas é mais volátil se comparado ao mercado acionário, assim, cabe ressaltar que a maioria das técnicas de análise técnica e de inteligência artificial necessitam de estudos e avaliações, para identificar diferentes parametrizações de modelos causadas pela diferença de volatilidade. Bem como há menos estudos destas técnicas em criptomoedas devido a este ativo ser mais recente e com falta de regulamentação em muitos países.

2. ANÁLISE TÉCNICA

A análise técnica é "...o estudo da ação do mercado, primariamente por meio do uso de gráficos, com o propósito de prever tendências de preço futuro" (Murphy, 1999, p. 2). Há duas abordagens para avaliar o preço de ações: utilizando indicadores e osciladores; ou utilizando gráficos (Chavarnakul & Enke, 2008). Murphy (1999) pontua que o técnico possui três principais fontes de informação: o preço, volume e posições em aberto.

Há três premissas que jamais devem ser desprezadas, a saber: (i) "A ação do mercado representa tudo" (Murphy, 1999, p. 2) – considera que tudo o que poderia causar um efeito no preço de mercado, já causou e seu efeito já está refletido no preço observado; (ii) "Os preços seguem tendências" (Murphy, 1999, p. 3) – considera que frente a uma tendência estabelecida, há maior probabilidade de esta se manter do que se reverter; e (iii) "a história se repete" (Murphy, 1999, p. 3-4) – considera que é possível analisar os gráficos dos últimos cem anos, e averiguar que

por diversas vezes as reações a determinados padrões nos gráficos foram semelhantes, fazendo com que a história se repetisse.

Murphy (1999) define a análise fundamentalista num contraponto direto à primeira premissa supracitada. Enquanto, para um técnico, tudo já está expresso no preço, o analista fundamentalista busca a causa da ação do mercado e do preço. Enquanto busca tal causa, entretanto, perde a tendência que poderia ter aproveitado, pois os fundamentos básicos já estão expressos e todos conhecem; e quando os fundamentos estudados pelo analista forem descobertos seus efeitos já terão há muito passado (Murphy, 1999, p. 5). Visto que este trabalho tem como objetivo realizar a descoberta de tendências a partir do preço, não serão discutidos conceitos mais aprofundados sobre a análise fundamentalista.

No fim do milênio passado, com a maior democratização do acesso a computadores, a utilização destes para desenvolvimento de novos métodos para predição de preço começou a ser mais difundida, dando origem a técnicas que muito se assemelham à análise técnica, mas que dispensam a utilização de gráficos. Estes analistas estatísticos buscam quantificar conceitos subjetivos que os analistas técnicos tradicionais precisam compreender para produzir bons gráficos. O objetivo é produzir sinais de “compra” ou “venda” baseado probabilística e estatisticamente em valores e números objetivos, procurando afastar-se da falha humana e relatividade do julgamento (Murphy, 1999, p. 10). Murphy (1999) pondera que, sim, existe um grande fator de experiência e habilidade por parte do analista técnico, mesmo que este trabalhe com indicadores, afinal a produção de indicadores deve ser feita com exímia qualidade para que estes expressem a reação do mercado.

3. REDES NEURAIS

Inicialmente, redes neurais tinham por objetivo solucionar problemas os quais algoritmos tradicionais não conseguiam em tempos razoáveis. O modelo, inicialmente, seguia o padrão biológico, mas logo se tornou mais flexível a fim de responder às necessidades de mais e mais otimizações para solução de problemas diversos (Mackay, 2003). Redes neurais artificiais são um excelente método para se aplicar em problemas em que os dados de entrada possuem muitos ruídos e informações que podem ser descartáveis, como dados gerados por câmeras e microfones (Mitchell, 1997). Neurônios são conectados entre si por arestas com direção e peso. Regras de ativação e aprendizado podem variar a depender da arquitetura optada pelo autor (Mackay, 2003). Segundo Hornik, Stinchcombe e White (1989), não há restrições teóricas para o sucesso de redes neurais *feedforward*, e a falta de sucesso é mais provável de ser causada por treinamento inadequado, complexidade insuficiente da rede (quantidade de neurônios e quantidade de camadas ocultas/internas) ou relacionamento não determinístico entre a entrada e a resposta esperada.

2.5.1 Neurônio

Diversas redes neurais são compostas de apenas um neurônio, mas redes mais complexas podem ser compostas por diversas camadas de neurônios que se comunicam por meio dos impulsos transmitidos por conexões entre os neurônios.

Um neurônio possui um conjunto I de tamanho n de entradas (x_1, x_2, \dots, x_n) e uma saída y. Também possui um conjunto W de pesos, também de tamanho n. O valor de cada entrada é filtrado pelo peso de mesmo índice no seu próprio conjunto. Alguns neurônios possuem uma entrada adicional chamada de “viés”, que tem sempre valor 1 e é filtrada por um “peso de viés”; ambos valores de viés ocupam a posição 0 dos seus respectivos conjuntos. (Mackay, 2003)

A atividade do neurônio é feita em passos:

(i) dado um conjunto I de entradas x_i e um conjunto W de pesos w_i , ambos conjuntos de tamanho n, o valor de ativação é dado pela soma:

$$a = \sum_{i=m}^n x_i w_i, \text{ onde } m = 0 \text{ se há viés, ou } m = 1 \text{ se não}$$

(ii) A saída é o resultado do valor de ativação aplicado a uma função chamada de função de ativação. A saída pode ser chamada também de atividade do neurônio. A função de ativação em si pode variar de acordo com a proposta, tais como (Mackay, 2003):

a) Determinística

i. $y(a) = a$

ii. $y(a) = \frac{1}{1+e^{-a}}$, onde o resultado será um sinal percentual

iii. $y(a) = \tanh(a)$, onde o resultado será um sinal percentual positivo ou negativo

iv. $y(a) = \theta(a)$, onde o resultado será um sinal binário (1 ou -1)

b) Aleatória

i. $y = 1$ com probabilidade $\frac{1}{1+e^{-a}}$; senão $y = -1$

ii. Considerando o valor anterior da saída (y), se $\Delta = ay \leq 0$, alterna o estado de y; senão, alterna com uma probabilidade de $e^{-\Delta}$

2.5.2 Perceptron

Os perceptrons estabeleceram a fundação de redes neurais na década de 80 e 90 (Hastie, Tibshirani & Friedman, 2009). O Perceptron é um tipo específico de neurônio cuja saída é binária, emitindo sinal negativo ou positivo. A função de ativação que emite o sinal é dada pela soma de todos os elementos da entrada multiplicados pelo seu peso correspondente; esta soma deve ser positiva ou negativa, e o resultado será simplificado para apenas um sinal. (Mitchell, 1997, p. 86)

Esta estrutura pode ser utilizada individualmente para representar qualquer função booleana primitiva (AND, OR, NAND, NOR), mas é necessário mais de um perceptron para representar funções como XOR. Basicamente qualquer função booleana pode ser representada por uma rede de perceptrons de até dois níveis de profundidade (Mitchell, 1997, p. 87).

No quadro 1 estão descritos os elementos necessários a parametrização de um perceptron.

Requisitos	Descrição dos elementos
Regra de treinamento do perceptron	Refere-se a inicialização de pesos. Uma alternativa comum é iniciar todos com valores aleatórios, e ajustá-los conforme o perceptron der sinais errados. O ajuste de cada peso é dado por uma função da sua respectiva entrada (Mitchell, 1997, p. 88)
Gradiente descendente e regra do diferencial	Esta regra de treinamento, também chamada de regra delta, é desenhada especificamente para um conjunto de entradas (não linearmente separáveis). “Se os exemplos de treinamento não são linearmente separáveis, a regra delta irá convergir em direção à aproximação que melhor couber no conceito alvo” (Mitchell, 1997, p. 89). Para definir o aprendizado, primeiramente é preciso definir um cálculo do erro, a fim de identificar quão longe da resposta ideal o perceptron está chegando.
Sobreajuste	Sobreajuste é um termo da estatística que descreve quando um modelo estatístico — ou seja, um modelo que utiliza valores e funções matemáticas estatísticas para prever um resultado — se adaptou demais a uma massa de dados previamente apresentada, de forma que se mostra ineficiente para uma massa de dados nova, com as mesmas características (Cárdenas-Monte, 2015). sobre ajuste, redes neurais grandes costumam ser caras computacionalmente, e isto escalona cada vez mais quanto mais complexo os cálculos estatísticos que as unidades fazem se tornam (Srivastava, Hinton, Krizhevsky, Sutskever & Salakhutdinov, 2014). Dropout é uma técnica que soluciona este problema
Dropout	Reduz o custo computacional de redes neurais muito complexas, e também ajuda a evitar que o modelo se adapte em demasia aos dados de treino fornecidos (Srivastava et al., 2014). O termo se refere a, literalmente, remover unidades da rede temporariamente para a interação, de maneira que suas conexões sejam ignoradas e esta não seja afetada por algum ajuste que os demais pesos terão Importante ressaltar que esta técnica é somente aplicada durante a etapa de treino, e as configurações de dropout são ignoradas na tomada de decisão (Gal; Ghahramani, 2016).

Quadro 1 – Parametrização de um perceptron

Cabe observar que para cada iteração do treino um determinado percentual de neurônios é ocultado e isolado para fora do cálculo, pode-se ter que a sessão de treino é feita para várias redes enxugadas da original, porém a compondo; estas redes enxugadas são treinadas, cada uma, muito raramente, se é que o são alguma vez (Srivastava et al., 2014).

Durante o desenvolvimento desta pesquisa foram identificados os estudos de Velankar, Valecha e Maji (2018), Oliveira, Zárate, Reis e Nobre (2011) e Jasemi, Kimiagari e Memariani (2011), os quais têm características semelhantes a esta pesquisa.

Velankar, Valecha e Maji (2018) utilizaram o método de regressão Bayesiano para prever o preço do Bitcoin. Ao obter as informações e quebrá-las em intervalos de 180, 360 e 720 segundos. Velankar, Valecha e Maji (2018) aplicaram k-means para clusterizar as amostras de cada intervalo em 100 grupos, totalizando 300. Em uma segunda etapa, construindo três séries de dados para 30, 60 e 120 minutos (respectivamente relacionados aos intervalos anteriores de 180, 360 e 720 segundos). Aplicar o método de árvore de decisão aleatória para obter três séries que indicam as tendências nos seus respectivos intervalos. O estudo de Velankar, Valecha e Maji (2018), porém, não foi concluído, por isso não apresenta resultados dos experimentos além da preparação dos dados, os quais indicariam se a estratégia utilizada obteve.

No que diz respeito a previsão de ativos da bolsa Oliveira et al. (2011) utilizaram um método de três fases: obtenção de amostras, pré-processamento e previsão. Para obtenção de amostras, foram analisadas as cotações diárias de uma série de ativos da bolsa e aplicadas técnicas de

análise técnica para obtenção de indicadores que auxiliassem na antecipação de tendências de preço. Na segunda fase, há uma correção de informações que possam levar a enganos, como dias sem compra/venda devido a feriados; finaliza-se esta fase de pré-processamento com normalização da informação. A fase da predição se resume a inserir a informação na rede neural a fim de treiná-la para a tarefa de predição utilizando a informação existente na amostragem normalizada. O objetivo da previsão de Oliveira et al. (2011) foi obter o preço da ação no dia seguinte, e conseguiu obtê-lo no melhor experimento com RMS (Root mean square, ou Valor quadrático médio) de 0.0129, no qual utilizou 5 dias para predizer um único dia. O valor da média quadrática abaixo de 1 se traduz em uma excelente previsibilidade. Estima-se que o objetivo do estudo de Oliveira et al. (2011) foi o de predizer o movimento do mercado.

O estudo de Jasemi, Kimiagari e Memariani (2011) apresenta uma abordagem moderna que utiliza uma rede neural e a análise técnica de candlesticks. Nesta abordagem, a rede neural não utiliza as velas para suas predições, mas apresenta um modelo de regressão baseando-se em variáveis independentes que representam padrões importantes da análise técnica. Na definição das variáveis independentes, duas abordagens são feitas: uma em dados brutos e outra em sinais com quinze e vinte e quatro variáveis. Jasemi, Kimiagari e Memariani (2011) ao comparar os resultados estimados com valores reais (obtidos do yahoo!finance), o modelo proposto apresentou performance excelente na primeira abordagem, e um pouco menor na segunda.

4. MATERIAIS E MÉTODOS

De modo entender a aplicação de indicadores de análise técnica e de redes neurais na predição do movimento de preço do Bitcoin, foi realizada a extração de dados a partir do portal yahoo!finance e foram organizados em uma estrutura utilizada nas diversas etapas experimentais, de modo a realizar simulações de investimento utilizando diferentes agentes, tendo como base um valor inicial de 10.000 dólares. Os dados extraídos compreendem os valores históricos do dia 1 de julho de 2015 até 30 de junho de 2020, ou seja, 5 anos.

Os agentes podem ser classificados em 2 categorias principais: agentes que utilizam apenas indicadores técnicos e agentes que utilizam um modelo híbrido utilizando indicadores técnicos e redes neurais. Para o cálculo dos indicadores técnicos a estrutura apresenta os seguintes dados: preço (máximo, mínimo, abertura e fechamento), volume e a data.

4.3 Protótipo de simulação de análise técnica

Para realização dos experimentos utilizando apenas indicadores técnicos foi construído um protótipo de simulação que utiliza agentes no decorrer de um período para observarem o mercado e tomarem decisões de compra e venda, tendo em vista os movimentos de tendência. Estes agentes utilizam dois tipos de estratégias: 1 - *buy and hold*, em que o agente compra o ativo no início e só vende no final do período, tal estratégia serve apenas como base de comparação para os outros agentes; 2 - baseada em médias móveis, em que o agente utiliza uma ou mais médias móveis e, observando os sinais emitidos, decide comprar ou vender. Após a simulação, os ganhos (com todos os agentes partindo de um montante de U\$10 000) foram comparados entre si.

A simulação da estratégia 2 (baseada em médias móveis) compreende 12 agentes utilizando estratégias de médias móveis:

1. Uma única média móvel simples de 14 dias
2. Uma única média móvel simples de 20 dias
3. Uma única média móvel de 14 dias com envelope de 3%
4. Uma única média móvel de 14 dias com envelope de 5%
5. Uma única média móvel de 20 dias com envelope de 3%
6. Uma única média móvel de 20 dias com envelope de 5%
7. Duas médias móveis de 7 e 14 dias
8. Duas médias móveis de 5 e 20 dias
9. Três médias móveis de 4, 9 e 18 dias
10. Uma única média móvel exponencial de 7 dias
11. Uma única média móvel exponencial de 14 dias
12. Uma única média móvel exponencial de 20 dias

Destes mesmos agentes também foi feita uma avaliação de como realizam ganhos em cada período, independente do montante financeiro que possuem no momento.

Foram realizadas três simulações com períodos diferentes de início, mas terminam no mesmo dia (30/06/2020): a primeira começa no dia 1 de julho de 2015; a segunda no dia 1 de julho de 2017; e a terceira no dia 1 de julho de 2019. Desta forma, os resultados estão individualizados para cada agente para cada uma destas simulações. O resultado de uma simulação não influencia na outra, pois o saldo de cada um deles (inicialmente 10.000 dólares) é reiniciado para cada nova simulação.

4.4 Protótipo de simulação com rede neural

Para a última série de experimentos, foi incluída uma estratégia 3 (híbrida) à estrutura do protótipo: possui uma série de médias móveis e um modelo de rede neural, e utiliza os sinais emitidos pelas médias móveis como entrada para a estratégia. Como o conjunto de dados total extraído é de 5 anos, 1 de julho de 2015 a 30 de junho de 2020, não há como replicar experimentos de agentes de redes neurais totalmente comparáveis com os agentes de análise técnica. Isto ocorre porque as redes neurais demandam de um conjunto de dados mínimo de alguns anos para treinamento, enquanto a análise técnica demanda de apenas algumas dezenas de dias para poder ser utilizado. Então dois conjuntos de experimentos foram realizados: experimentos com modelos treinados com dados históricos de 4 anos (jul/2015 a jun/2019); e experimentos com modelos treinados com dados históricos de 3 anos (jul/2015 a jun/2018). Cada modelo foi incluído em um agente participante de simulações de investimento, com capital inicial de 10.000 dólares, durante 1 ano.

Para a experimentação, todos os agentes utilizaram todas as doze diferentes combinações de médias móveis listadas na seção anterior, e um modelo de rede neural próprio (porém, todos possuindo a mesma entrada de 12 sinais). Para ajustar a rede neural e descobrir que parâmetros obtinham um melhor resultado, foi avaliado uma série de variações, incluindo funções de perda, existência ou não de *dropout* e velocidade de aprendizado do otimizador. Por meio de duas configurações diferentes de camadas (144/72/36/1 e 9/9/9/1), cerca de 1200 diferentes configurações foram executadas para cada um a fim de otimizar o resultado. Mais especificamente, utiliza-se um modelo sequencial com camadas densas disponibilizadas pela biblioteca. Um modelo sequencial é uma pilha de camadas em que cada uma possui um vetor de entradas e um vetor de saídas; uma camada densa é uma camada comum totalmente conectada.

5. ANÁLISE DE DADOS - SIMULADOR DE ANÁLISE TÉCNICA

Os experimentos foram realizados utilizando doze agentes, cada um utilizando uma combinação de médias móveis diferentes para definir suas ações, partindo de um saldo de dez mil dólares. Dois agentes utilizaram médias móveis simples (14 e 20 dias); dois agentes utilizaram médias móveis duplas (7/14 e 5/20 dias); um agente utilizou médias móveis triplas (4/9/18); três agentes utilizaram médias móveis exponenciais (7, 14 e 20 dias). Também, a partir de cada um dos dois agentes que utilizaram médias móveis simples, outros dois agentes foram criados de médias móveis simples com envelopes de 3% e 5%. Estes agentes foram nomeados baseando no tipo de média móvel que utilizavam.

A primeira simulação verificou a ação dos agentes pelo período de cinco anos, iniciando-se no dia 1 de julho de 2015 e finalizando no dia 30 de junho de 2020. Os resultados sintetizados de todos os agentes seguem a Tabela 1, ordenados do maior ganho percentual para o menor.

Agente	Saldo final	Ganho total	Ganho relativo
Enveloped (5.0% 14)	\$1.118.040,85	11080,41%	7626,62%
Exponential (7)	\$986.018,42	9760,18%	6306,39%
Triple (4/9/18)	\$919.151,45	9091,51%	5637,72%
Exponential (20)	\$866.339,54	8563,40%	5109,60%
Simple (20)	\$807.242,47	7972,42%	4518,63%
Simple (14)	\$798.937,01	7889,37%	4435,58%
Enveloped (3.0% 14)	\$772.992,11	7629,92%	4176,13%
Double (7/14)	\$734.500,75	7245,01%	3791,21%
Double (5/20)	\$733.282,50	7232,83%	3779,03%
Enveloped (3.0% 20)	\$542.649,81	5326,50%	1872,71%
Enveloped (5.0% 20)	\$540.177,30	5301,77%	1847,98%
Exponential (14)	\$478.874,31	4688,74%	1234,95%

Tabela 1 – Resultados simulação de cinco anos

Na Tabela 1, o ganho relativo foi calculado subtraindo o incremento no valor do ativo no período (vide Tabela 2, ou seja, foi o ganho obtido pelo agente que utilizou estratégia de *buy*

and holding) do ganho total de cada um dos agentes. Identificou-se, a partir dos ganhos sobre a estratégia hipotética de *buy and holding* — comprar determinado ativo e mantê-lo buscando sua valorização a longo prazo) —, quatro destaques: Envelopada de 5% por 14 dias; média móvel exponencial de 7 dias; triple moving average de 4/9/18 dias, conforme estabelecido na literatura (Murphy, 1999); exponencial de 20 dias. Próximos entre si, as médias móveis simples de 20 e 14 dias tiveram um desempenho muito satisfatório também.

Período	Valor inicial	Valor final	Varição percentual
julho/2015 a junho/2020	U\$ 258,62	U\$ 9.190,85	3453,79%
julho/2017 a junho/2020	U\$ 2.434,55	U\$ 9.190,85	277,52%
julho/2019 a junho/2020	U\$ 10.583,13	U\$ 9.190,85	-13,16%

Tabela 2 – Variação do preço do BTC

Uma segunda simulação utilizou os últimos três anos do período anterior para verificar o desempenho dos mesmos agentes, quando o início do período é em 1 de julho de 2017, com 10 mil dólares. Os resultados, conforme o padrão da simulação anterior, podem ser verificados na Tabela 3.

Agente	Saldo final	Ganho total	Ganho relativo
Simple (20)	\$155.934,06	1459,34%	1181,82%
Exponential (7)	\$139.871,37	1298,71%	1021,20%
Double (5/20)	\$139.699,70	1297,00%	1019,48%
Enveloped (5.0% 14)	\$132.929,86	1229,30%	951,78%
Exponential (20)	\$121.760,90	1117,61%	840,09%
Triple (4/9/18)	\$116.787,62	1067,88%	790,36%
Enveloped (3.0% 20)	\$107.911,43	979,11%	701,60%
Enveloped (3.0% 14)	\$101.349,84	913,50%	635,98%
Enveloped (5.0% 20)	\$98.577,12	885,77%	608,25%
Simple (14)	\$96.079,85	860,80%	583,28%
Double (7/14)	\$83.904,63	739,05%	461,53%
Exponential (14)	\$69.453,09	594,53%	317,01%

Tabela 3 – Resultados simulação de três anos

De acordo com as Tabelas 1 e 3, comparando o uso dos 6 indicadores técnicos com maior retorno, verifica-se que são os mesmos, contudo, em ordem diferente: simples de 20 dias; exponencial de 7 dias; dupla de 5 e 20 dias; envelopada de 5% para 14 dias; exponencial de 20 dias; tripla de 4, 9 e 18 dias. A diferença desta simulação para a anterior é decorrente da posição: enquanto na Tabela 1 a média móvel simples envelopada de 5% para 14 dias esteve entre os destaques, já na tabela 3 a média móvel simples de 20 dias se destacou. Nas tabelas, as simulações que não estiveram no topo, também não estavam próximas a ele; podem ser listadas entre estratégias menos ótimas.

No entanto, os valores percentuais destas simulações não podem ser comparados entre si devido ao fator de acúmulo de montante e reinvestimento. Ao passo que a segunda simulação teve ao

seu dispor 10 mil dólares ao início de seu período, neste mesmo dia a simulação maior (de cinco anos) já possuía um montante acumulado das transações efetuadas nos dois anos anteriores. Por isso o cálculo de ganho é feito baseando-se no ganho entre o início e o final da simulação, a comparação deve se limitar apenas a agentes de uma mesma simulação. Contudo, cabe destacar que diferentes períodos de investimento e diferentes indicadores técnicos tiveram resultados diferenciados.

5.1 Simulação de investimento com agentes de rede neural

Nesta seção são descritos os resultados obtidos nos experimentos utilizando modelos de rede neural. Considerando que estes modelos demandam um período de treinamento, e que o tamanho total da massa de dados compreende apenas 5 anos, não se faz possível realizar experimentos totalmente equivalentes e comparáveis aos experimentos realizados para a série de experimentos anterior. Portanto, dois conjuntos de experimentos compõem esta etapa: experimentos com modelos treinados com dados históricos de 4 anos (jul/2015 a jun/2019); e experimentos com modelos treinados com dados históricos de 3 anos (jul/2015 a jun/2018). Cada modelo foi incluído em um agente participante de simulações de investimento, com capital inicial de 10.000 dólares, durante 1 ano.

5.1.1 Simulação de especulação dos agentes de aprendizado de máquina - atuando durante jul/2019 a jun/2020

Dos vários agentes disparados executando especulações no período referente ao último ano, houve um total de 12 que se destacaram com um ganho maior do que o melhor agente do experimento anterior no mesmo período (Média móvel exponencial de 7 dias, que obteve o melhor resultado no experimento demonstrado na Tabela 4). Destes, metade estava utilizando a configuração I de camadas (144, 72, 36, 1) (Tabela 4); e a outra metade a configuração II (9, 9, 9, 1) (Tabela 5).

Nome	Número de Épocas	Taxa de Aprendizado	Função de Perda	Dropout	Saldo Final
IP-64-0.01	64	0,01	Poisson	-	\$16.222,99
IP-64-0.25	64	0,25	Poisson	AB	\$16.415,11
IK-1-0.25	1	0,25	KLD	AC	\$15.672,67
IK-1-0.001	1	0,001	KLD	A	\$15.912,94
IM-64-0.3	64	0,3	MSE	A	\$18.547,73
IM-4-0.2	4	0,2	MSE	AB	\$16.416,59

Tabela 4 – Melhores configurações de camadas I

Observa-se na Tabela 4 que os agentes que se destacaram tiveram combinações variadas das configurações disponíveis, referente às funções de perda e sua relação com o número de épocas, é que (i) O desempenho utilizando divergência Kullback e Leibler (1951) se destaca com um número baixo de épocas para treinamento — tal destaque pode indicar que estes modelos não sejam confiáveis — e (ii) A função de perda Poisson se destacou com 64 épocas. É importante notar que das demais configurações que estão nos melhores 50 resultados, apenas uma destas teve 128 épocas de treinamento, obtendo um saldo final de mais ou menos 12000 (Conquistaria

um 4o lugar comparando com a Tabela 16. As configurações de taxa de aprendizado e *dropout* foram variadas.

Na Tabela 5 em que cada camada tinha muitos neurônios a menos, os resultados se apresentaram de maneira diferente. Os três melhores resultados possuíam a menor taxa de aprendizado configurada dentro os testes, e o quarto melhor resultado possui a segunda menor taxa de aprendizado. Pode-se observar, também, que no agente IP-32-0.001, mesmo com uma taxa de aprendizado baixa, as configurações de *dropout* foram bastante utilizadas, permitindo uma boa generalização mesmo com poucos neurônios (quando comparado à quantidade da configuração I). Percebe-se que não há, entre os melhores resultados, a aparição da função de perda de Kullback e Leibler (1951) apareceu como o melhor resultado, com \$13.648,28 de saldo final.

Nome	Número de Épocas	Taxa de Aprendizado	Função de Perda	Dropout	Saldo Final
IIP-2-0.001	2	0.001	Poisson	A	\$17.394,42
IIM-1-0.001	1	0.001	MSE	AB	\$16.152,43
IIP-32-0.001	32	0.001	Poisson	ACDE	\$16.263,09
IIM-64-0.05	64	0.05	MSE	A	\$15.796,72
IIM-128-0.2	128	0.2	MSE	-	\$18.738,80
IIP-1-0.15	1	0.15	Poisson	ACDE	\$16.336,75

Tabela 5 – Melhores configurações de camadas II

Considerando o volume de testes, diversos agentes apresentaram resultados comparáveis aos resultados obtidos no experimento anterior, utilizando agentes de médias móveis comuns. Entretanto, para reduzir a quantidade de informação apresentada, utiliza-se os resultados que obtiveram um resultado acima do melhor resultado do experimento equivalente. A Tabela 6 compara os resultados das duas tabelas anteriores ao melhor resultado do período no experimento de indicadores técnicos, o de média móvel exponencial de 7 dias.

Agente	Saldo final	Ganho total	Ganho relativo
IIP-2-0.001	\$18.738,80	87,39%	100,55%
IP-64-0.01	\$18.547,73	85,48%	98,64%
IIM-1-0.001	\$17.394,42	73,94%	87,10%
IP-64-0.25	\$16.416,59	64,17%	77,33%
IK-1-0.25	\$16.415,11	64,15%	77,31%
IIP-32-0.001	\$16.336,75	63,37%	76,53%
IIM-64-0.05	\$16.263,09	62,63%	75,79%
IK-1-0.001	\$16.222,99	62,23%	75,39%
IIM-128-0.2	\$16.152,43	61,52%	74,68%
IM-64-0.3	\$15.912,94	59,13%	72,29%
IIP-1-0.15	\$15.796,72	57,97%	71,13%
IM-4-0.2	\$15.672,67	56,73%	69,89%
Exponential (7)	\$15.649,49	56,49%	69,65%

Tabela 6 – Ganhos percentuais dos melhores agentes durante último ano

A média móvel exponencial teve um desempenho considerado aceitável no período: em um período de queda do ativo, esta estratégia aumentou o montante financeiro. Nos anos 2018 e 2017, também apresentou um resultado satisfatório. Entretanto, teve um recuo no montante financeiro nos dois primeiros anos. As demais estratégias apresentadas na Tabela 8 foram treinadas utilizando informações dos quatro primeiros anos (e portanto, assume-se que estão bem adaptadas aos sinais emitidos durante tal período), e desempenharam um resultado acima da estratégia Exponencial (7), em período de queda.

5.1.2 Simulação de especulação dos agentes de redes neurais - atuando durante jul/2018 a jun/2019

O mesmo teste foi realizado para o penúltimo ano de dados (jul/2018 a jun/2019). Os agentes desta subseção tiveram três anos de treinamento (jul/2015 a jun/2018), e seus resultados (Tabela 7) são comparáveis se for utilizado média móvel de 14 dias com envelope de 5% e investisse no mesmo período (saldo final de \$32.350,37). Destes, um total de oito modelos de redes neurais com indicadores técnicos obtiveram um saldo final maior do que o saldo do melhor agente que usava apenas indicadores técnicos.

Na Tabela 7 pode-se observar que oito modelos utilizaram a configuração de camadas I. Nesta configuração, ao contrário do observado no tópico anterior, tem-se grande presença da função de perda KLD, e está se apresentou com variados números de épocas utilizada para o treino (128, 32, 16, 2). Este comportamento vem acompanhado da utilização de dropout e uma variação também na taxa de aprendizado. De forma geral, tem-se um desempenho melhor para a função de perda KLD, e é possível identificar que esta função de perda, para esta configuração de camadas, se beneficiou do comportamento diferente deste ano.

Nome	Número de Épocas	Taxa de Aprendizado	Função de Perda	Dropout	Saldo Final
IK-32-0.3	32	0,3	KLD	A	\$34.324,91
IK-128-0.25	128	0,25	KLD	ABF	\$33.696,86
IP-128-0.05	128	0,05	Poisson	-	\$33.362,92
IK-32-0.25	32	0,25	KLD	AB	\$33.265,20
IM-32-0.15	32	0,15	MSE	AB	\$32.929,98
IK-32-0.15	32	0,15	KLD	AB	\$32.793,53
IK-16-0.01	16	0,01	KLD	-	\$32.731,27
IK-2-0.001	2	0,001	KLD	AC	\$32.557,03

Tabela 7 – Melhores configurações de camadas I

Ao se utilizar a configuração II, apenas um modelo obteve saldo final acima do melhor saldo do agente Exponencial (5.0% 14). A configuração deste modelo, apresentada na Tabela 8, utilizou como função de perda o erro quadrático médio, com uma taxa de aprendizado mínima e dropout apenas na entrada.

Nome	Número de Épocas	Taxa de Aprendizado	Função de Perda	Dropout	Saldo Final
IIM-2-0.001	2	0,001	MSE	A	\$32.915,02

Tabela 8 – Melhores configurações de camadas II

Na Tabela 9 é possível observar comparativamente como cada um dos melhores modelos desempenharam comparado ao agente de melhor resultado utilizando apenas indicadores técnicos, ou seja, o de médias móveis simples de 14 dias envelopada com 5%. Entretanto, como ocorreu no experimento anterior, este modelo apenas treinou 2 épocas, e as versões deste que tiveram um período maior de treino, não tiveram desempenho como o esperado, o que pode indicar um modelo com resultados não muito confiáveis ou previsíveis, ou seja, que possivelmente não replicaria o resultado positivo com consistência.

Agente	Saldo final	Ganho total	Ganho relativo
IK-32-0.3	\$34.324,91	243,25%	155,97%
IK-128-0.25	\$33.696,86	236,97%	149,69%
IP-128-0.05	\$33.362,92	233,63%	146,35%
IK-32-0.25	\$33.265,20	232,65%	145,37%
IM-32-0.15	\$32.929,98	229,30%	142,02%
IK-32-0.15	\$32.793,53	227,94%	140,66%
IK-16-0.01	\$32.731,27	227,31%	140,03%
IK-2-0.001	\$32.557,03	225,57%	138,29%
IIM-2-0.001	\$32.915,02	229,15%	141,87%
Enveloped (5.0% 14)	\$32.350,37	223,50%	136,22%

Tabela 9 – Ganhos percentuais dos melhores agentes durante último ano

Como pode ser verificado pelos experimentos, as redes neurais apresentaram um desempenho financeiro melhor apenas em determinadas configurações de parâmetros, em detrimento do uso puro de indicadores técnicos. Cabe destacar ainda que, mesmo os indicadores técnicos não desempenham consistentemente no decorrer dos anos, assim como as redes neurais.

6. CONSIDERAÇÕES FINAIS

A especulação no mercado mais do que nunca é um nicho potencialmente lucrativo e um mercado que sempre aceita profissionais capacitados. Com boa utilização de técnicas de predição de movimentos de mercado, é possível buscar uma margem de lucro e com um risco controlado. Neste contexto, a análise técnica é utilizada para prever o movimento do mercado há séculos, porém depende do profissional realizar uma análise dos gráficos, dos indicadores, e fazer uma aposta calculada sobre a subida ou descida do valor de mercado.

Mais recentemente, em 2007, o Bitcoin (BTC) foi introduzido como a primeira moeda digital mundial descentralizada. Sua proposta era de construir um sistema econômico livre, rastreável, reversível e não dependente de instituições de confiança. É um sistema eletrônico bancário *peer-to-peer*, ou seja, sem um nó centralizador e detentor das informações válidas; a rede, como um todo, trabalha para validar ou invalidar as transações. Com uma abordagem ousada, o BTC vem

apresentando um crescimento em seu valor de mercado, e, portanto, é também utilizado para especulação de mercado, e pode ser considerado um ativo para este fim. Desta forma, seria possível aplicar as mesmas técnicas aplicadas há séculos no mercado da bolsa de valores.

Estas técnicas, como quaisquer outras, têm potencial de serem automatizadas, desde que estabelecidas as estratégias de sua utilização: cria-se regras para estabelecer objetivamente que decisão seria a melhor a ser tomada. Neste cenário, foi então estabelecido o seguinte objetivo: analisar a aplicação de uma rede neural para realizar predições do movimento de preço do Bitcoin em contraponto e uso da análise técnica.

Para atingir este objetivo, foram realizadas as seguintes atividades: (1) Levantamento de técnicas para predição de tendência de preço que possam ser utilizadas para BTC; (3) Levantamento teórico de técnicas que já tenham sido aplicadas na predição de BTC; (4) Analisar o desempenho de técnicas de predição do movimento de preço aplicadas ao BTC. Com a pesquisa, foi possível calcular vários indicadores técnicos com base no histórico de preços do BTC, para a criação de agentes de simulação de investimento utilizando indicadores técnicos e agentes de rede neural com base em indicadores técnicos, utilizando um valor inicial de 10.000,00 dólares. Foram identificadas nos experimentos algumas estratégias que obtiveram um resultado consideravelmente positivo, inclusive em momentos de queda como o último ano analisado (jul/2019 a jun/2020). Cada configuração de média móvel, entretanto, possui melhor desempenho em determinadas configurações, e, portanto, vê-se que para utilizar desta abordagem, o ideal seria construir um modelo que utiliza diversas médias móveis para conseguir se adaptar aos diferentes movimentos do mercado.

Considerando que o mercado possui uma natureza matemática, em que os movimentos seguem padrões, julga-se possível fazer a predição por meio de uma rede neural treinada para associar os padrões entre indicadores técnicos e a variação do preço no mercado.

Verificou-se que os resultados das redes neurais foram mais positivos do que os resultados de cada um dos agentes de médias móveis, obtendo até 80% de aumento do montante em um período em que o BTC perdeu 15% de seu valor de mercado: aproveitando das variações, mesmo em uma descendente de preço, para aumentar seu saldo.

Apesar de um desempenho positivo, a rede neural construída é um modelo simples: recebe como entrada 12 indicadores que indicam compra, venda, ou um sinal parcial. Utilizando esta entrada, treinada por um período, a rede neural se adaptou a considerar mais alguns indicadores que outros. Entretanto, o modelo proposto é alheio ao movimento do mercado propriamente dito: apenas conhece os dados que estes 12 indicadores lhe apresenta. Disto, pode-se considerar para os trabalhos futuros aplicar alguns avanços para o que foi desenvolvido neste trabalho: (1) Construir um modelo com possibilidade de identificar o movimento de mercado atual, com conhecimento dos preços dos últimos dias, ou da variação destes. Imagina-se que por meio da inserção dos preços (ou da variação destes) dos últimos dias, a rede neural se adaptaria melhor a escolher quais indicadores ignorar e quais mais considerar em determinados momentos. Um segundo aspecto que foi pouco explorado neste trabalho foi o tuning da rede neural: foi feito um tuning de força bruta, elencando-se possibilidades de variações e testando-as diretamente. Desta experiência, elenca-se um segundo desenvolvimento que se pode fazer a partir dos experimentos aqui demonstrados: (2) Construir uma estrutura com mais parâmetros para a

compilação da rede neural, e uma estrutura de algoritmo genético para averiguar qual o melhor conjunto de parâmetros (realizar o tuning da rede neural). Um terceiro ponto, e mais simplista, é utilizar diferentes tipos de camadas na rede neural, uma vez que neste trabalho apenas foram construídos modelos utilizando camadas densas, ou seja, não foi aplicado redes recorrentes, nem outros tipos de redes que não fossem uma rede neural simples..

REFERÊNCIAS

- Cárdenas-Monte, M. (2015). Sobreajuste - Overfitting. *Course notes*. Disponível em: <<http://www.wae.ciemat.es/~cardenas/docs/lessons/sobreajuste.pdf>>. Acesso em: 21 set. 2017.
- Chavarnakul, T., & Enke, D. (2008). Intelligent technical analysis based equivolume charting for stock trading using neural networks. *Expert Systems with Applications*, v. 34, n. 2, p. 1004 – 1017.
- Elder, A. (1993). Trading for a living: psychology, trading tactics, money management, *John Wiley & Sons*, v. 31.
- Gal, Y., & Ghahramani, Z. (2016) A theoretically grounded application of dropout in recurrent neural networks. *Advances in neural information processing systems*, v. 29, p. 1019–1027.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. *Springer Science & Business Media*, p. 699-745.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, v. 2, n. 5, p. 359–366.
- Jasemi, M., Kimiagari, A. M., & Memariani, A. (2011). A modern neural network model to do stock market timing on the basis of the ancient investment technique of japanese candlestick. *Expert Systems with Applications*, v. 38, n. 4, p. 3884 – 3890.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 22, n. 1, p. 79–86.
- Mackay, D. J. C. (2003). *Information theory, inference and learning algorithms*. England: Cambridge university press, 2003.
- Mitchell, T. M. (1997). *Machine Learning*. EUA: McGraw-Hill.
- Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. EUA: New York Institute of Finance.
- Oliveira, F. A., Zárate, L. E., Reis, M. A. & Nobre, C. N. (2011). The use of artificial neural networks in the analysis and prediction of stock prices. *IEEE International Conference on Systems, Man, and Cybernetics*, v. 1, p. 2151–2155
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, JMLR. org, v. 15, n. 1, p. 1929–1958.
- Velankar, S.; Valecha, S.; Maji, S. (2018). Bitcoin price prediction using machine learning. In: IEEE. *20th International Conference on Advanced Communication Technology (ICACT)*, p. 144–147.