

UNIVERSIDADE FEDERAL DO PARANÁ

GABRIEL MOREDO ANELLI

**REDES NEURAS ARTIFICIAIS E A TEORIA DO PASSEIO ALEATÓRIO APLICADOS
À VOLATIDADE DA CRIPTOMOEDA RIPPLE**

CURITIBA

2018

GABRIEL MOREDO ANELLI

**REDES NEURAS ARTIFICIAIS E A TEORIA DO PASSEIO ALEATÓRIO
APLICADOS À VOLATIDADE DA CRIPTOMOEDA RIPPLE**

Trabalho de Conclusão de Curso
apresentado à disciplina de TCC II,
do Curso de Engenharia de
Produção, do Departamento de
Engenharia de Produção, do Setor de
Tecnologia, da Universidade Federal
do Paraná.

Profa. Dra. Mariana Kleina

CURITIBA

2018

RESUMO

A movimentação de dinheiro quando digital é intermediada por entidades financeiras, como os bancos de varejo. Com a crise financeira americana de 2008, surgiu uma nova discussão e consequente proposta de sistema descentralizado para eliminar estas entidades. Nesta proposta as transações seriam realizadas entre pessoas por meio de um sistema criptografado e com uma moeda digital: o *Bitcoin* via *Blockchain*, sistema descentralizado de armazenamento de informações. A partir deste marco, diversos projetos foram criados com o intuito de resolverem problemas utilizando a tecnologia *Blockchain* e criando novas criptomoedas. É o caso do *Ripple*, utilizada para a realização de transações interbancárias e internacionais, acelerando a velocidade de liquidação entre as duas pontas da transação. Processo que tradicionalmente demora entre 3 a 5 dias, por meio do *Ripple* pode ser feito em 4 segundos. No entanto, há uma grande volatilidade nas criptomoedas, apresentando em suas cotações o comportamento de Bolhas Financeiras e consequentemente levando investidores destes ativos a obterem grandes ganhos ou perdas. Dentro da ótica da previsibilidade dos ativos financeiros, existe divisão entre vertentes que defendem os mercados como previsíveis ou aleatórios. O objetivo deste trabalho é realizar uma análise da criptomoeda *Ripple* por meio de dois métodos: Redes Neurais Artificiais e Passeio Aleatório, que carrega a tese de que o mercado é imprevisível. Estes métodos são aplicados em dois cenários: (A) fora de Bolha Financeira e no cenário de Bolha Financeira (B). A análise comparativa utilizada é a do cálculo do Erro Quadrático Médio para definir qual método possuiu maior assertividade, sendo a Rede Neural o método que obteve resultados 16,10 vezes mais assertivos que o Passeio Aleatório no Cenário A e o Passeio Aleatório obteve resultados 2,04 vezes mais assertivos no Cenário B.

Palavras-chave: *blockchain*, *Ripple*, redes neurais artificiais, passeio aleatório, previsão, inteligência artificial.

ABSTRACT

Peer-to-peer transactions rely exclusively on financial institutions serving as trusted third parties to process electronic payments, such as retail banks. Since the subprime crisis in 2008, a new decentralized network was proposed with the goal of eliminating the middleman and instead of it, using cryptography and a digital currency for peer-to-peer transactions: Bitcoin via Blockchain, a decentralized ledger. Since then, many projects were created using the Blockchain technology for many kinds of purposes and launching new cryptocurrencies. One of these new cryptocurrencies is called Ripple and the main goal of the Project is to provide a network capable of doing interbank transactions on an average of 4 seconds instead of 3 to 5 days. There is a lot of volatility on the crypto market and in some cases, it is frequent to watch Financial Bubbles patterns, which can bring huge gains or losses to investors. Though the forecasting side, analysts are divided between the Random Walk hypothesis which says that the Markets are unpredictable and the non-Random Walk hypothesis. The main goal of this paper was to analyze the cryptocurrency Ripple through two quantitative methods: Artificial Neural Networks and Random Walk. These methods were applied on different scenarios: a non-Financial Bubble scenario (A) and a Financial Bubble scenario (B). The comparison between methods was made by the mean square error to define which one was more accurate, resulting on the Artificial Neural Network being the best method for scenario (A), 16,10 times more accurate than the Random Walk and in the scenario B, Random Walk was 2,04 times more accurate than the Neural Network.

Key words: blockchain, Ripple, artificial neural networks, random walk, forecasting, artificial intelligence.

1. INTRODUÇÃO	7
1.1 OBJETIVO GERAL	8
1.2 OBJETIVOS ESPECÍFICOS	8
2 REVISÃO DE LITERATURA	9
2.1 CRIPTOMOEDAS	9
2.1.1 <i>BLOCKCHAIN</i> CONTEXTUALIZADO	9
2.1.2 RIPPLE E AS TRANSFERÊNCIAS GLOBAIS	9
2.2 BOLHAS FINANCEIRAS	10
2.2.1 O CONCEITO DE BOLHA FINANCEIRA	10
2.3 REDES NEURAS ARTIFICIAIS (RNA)	12
2.3.1 DEFINIÇÃO DE REDES NEURAS ARTIFICIAIS	12
2.3.2 FUNÇÕES DE ATIVAÇÃO	14
2.3.3 TREINAMENTO DA RNA	14
2.4 O ALGORITMO DO PASSEIO ALEATÓRIO	15
3 METODOLOGIA	16
3.1 FLUXOGRAMA DA METODOLOGIA	16
3.2 O SOFTWARE R	17
3.3 BIBLIOTECAS PARA TREINAMENTO DA RNA E PREVISÃO	17
3.4 DEFINIÇÃO DOS CENÁRIOS	18
3.5 SCRIPTS EM R	19
3.6 TOPOLOGIA DA RNA	22
3.7 INTERPRETAÇÃO DAS ANÁLISES E CENÁRIOS GERADOS	25
4 RESULTADOS	26
5 CONCLUSÃO	28
5.1 SUGESTÕES E PRÓXIMOS PASSOS	28

6	REFERÊNCIAS	30
----------	--------------------------	-----------

1. INTRODUÇÃO

Próximo à Crise Financeira de 2008, um documento publicado pelo pseudônimo de “Satoshi Nakamoto”, intitulado “*Bitcoin: A Peer-to-Peer Electronic Cash System*” (NAKAMOTO, 2008) iniciou discussões e a corrida pelo desenvolvimento de uma tecnologia que pode transformar a maneira com que são feitas transações financeiras e registradas diversas informações. Sendo assim, o *Blockchain* pode ser definido como um banco de dados descentralizado, onde movimentações são assinadas por meio de criptografia.

Juntamente ao progresso tecnológico do *Blockchain*, surgem diversas criptomoedas que podem ser comercializadas em bolsas de valores ao redor do mundo. A alta volatilidade destes ativos atrai investidores de risco que buscam ganhos financeiros expressivos em meio as oscilações.

No mercado acionário, o comportamento de uma ação pode ser expresso matematicamente, no entanto, a ausência de linearidade pode resultar em pouca acurácia para previsões (REFENES e ALIPPI, 1993). Por conta disto, Redes Neurais Artificiais são uma alternativa para a previsão da cotação de ativos.

Mueller (1996) e Albuquerque (2008) fizeram uso de redes neurais artificiais para a previsão do mercado de ações. Propondo um modelo preditivo interpretando o comportamento de séries temporais. O modelo desenvolvido por meio de redes neurais artificiais, identifica regularidades no comportamento dos preços das ações. Analogamente a isto, serão desenvolvidas análises quantitativas que envolvem redes neurais artificiais e o método do Passeio Aleatório para a previsão de outra classe de ativo financeiro, as criptomoedas.

O presente trabalho tem como objetivo aplicar métodos de previsão quantitativos na cotação da criptomoeda *Ripple*, analisando o erro quadrático médio destes métodos em um cenário de Bolha Financeira e em um cenário fora de uma Bolha Financeira. Sendo uma Bolha Financeira um período de análise de uma série temporal em que há um crescimento exacerbado de um valor e, em seguida, uma queda abrupta deste mesmo número analisado.

OBJETIVOS

1.1 OBJETIVO GERAL

Aplicar métodos de previsão quantitativos no problema de previsão de cotação de uma criptomoeda utilizando implementação computacional em linguagem R.

1.2 OBJETIVOS ESPECÍFICOS

O desenvolvimento do estudo envolve:

- Tratamento de dados para análise preditiva;
- Implementação de uma Rede Neural Artificial utilizando a linguagem de programação R;
- Implementação computacional em R do método de Passeio Aleatório;
- Comparar, mediante análise de erro, os resultados dos métodos quantitativos implementados.

2 REVISÃO DE LITERATURA

2.1 CRIPTOMOEDAS

2.1.1 BLOCKCHAIN CONTEXTUALIZADO

Segundo a definição de Nakamoto (2008), uma moeda digital é uma corrente formada por assinaturas digitais. Um agente A, possuidor de certo montante X, transfere o montante X para um agente B, ao assinar digitalmente o *hash*, ou seja, uma sequência numérica da transação anterior e a chave pública do agente B, adicionando isto ao final do código da moeda digital.

Além disto, todas as transações possuem um *timestamp*, que é o momento quando ocorreu a transação e fica gravado permanentemente no *Blockchain*, sem a possibilidade de ser alterado.

Para a viabilização das transações do *Blockchain*, Nakamoto (2008) introduz o conceito de prova de trabalho, de modo a eliminar a terceira parte intermediadora da transação entre duas partes. O conceito de prova de trabalho está ligado diretamente à mineração de *Bitcoins*, tratando-se do fluxo contínuo de scanear os *hashs* das transações.

Para fins de simplificação, o *Blockchain* funciona como um livro contábil em que qualquer indivíduo consegue visualizar as transações feitas, sem que possa alterá-las de qualquer forma. Os agentes das transações permanecem no anonimato de suas pessoas físicas, sendo apenas simbolizados por uma série de letras e números que representa o endereço de suas carteiras virtuais.

Após o surgimento do *Bitcoin*, outras redes foram criadas utilizando a tecnologia do *Blockchain* por exemplo, a criptomoeda Ripple, proveniente de um instituto privado e que tem como proposta de valor a resolução do problema de transferências globais a um baixo custo e em altas velocidades.

2.1.2 RIPPLE E AS TRANSFERÊNCIAS GLOBAIS

Dentre os diversos institutos e empresas que lançaram criptomoedas e serviços baseados no *Blockchain*, o Ripple encontra-se como a solução mais

valorizada quanto ao seu valor e conquistas no mercado de pagamentos internacionais, comumente chamado por *payments cross-border*.

Segundo Ripple (2018), a criptomoeda proporciona o benefício de eliminar agentes intermediários (e.g bancos e instituições financeiras) que atuam entre 2 pessoas durante uma transação, substituindo-os pela rede Ripple. Este processo é economicamente vantajoso, por eliminar custos e proporcionar a vantagem de reduzir o tempo de transferência de uma média de 3 à 5 dias para apenas 4 segundos.

Desta forma, o Ripple apresenta-se como uma possível solução para o problema de pagamentos internacionais. O desenvolvimento de análises quantitativas para a previsão de sua cotação é de grande valia para instituições financeiras e pessoas que pretendem utilizar a tecnologia para remessas internacionais.

2.2 BOLHAS FINANCEIRAS

2.2.1 O CONCEITO DE BOLHA FINANCEIRA

De acordo com Phillips e Yu (2011), acadêmicos que estudaram a Bolha Financeira da Crise de 2008 (crise do *subprime*), uma Bolha Financeira é determinada pelo crescimento acelerado de um ativo financeiro de modo que isto esteja nitidamente desconectado de uma lógica que acompanhe a valorização intrínseca do ativo. Em outros termos, é uma valorização irracional e desproporcional de um ativo.

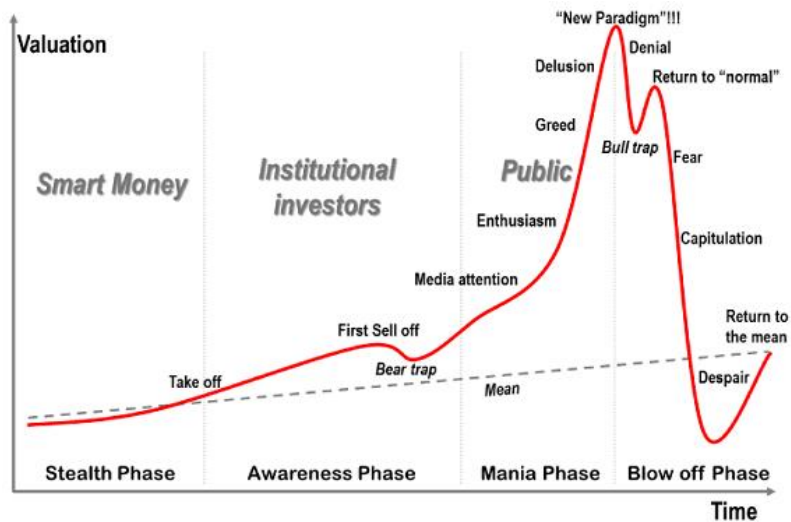
Uma clássica Bolha Financeira, foi a Bolha da Internet nos anos 2000. Graficamente é possível ver a semelhança entre o gráfico dos pontos do pregão da Nasdaq (FIGURA 1) e do gráfico clássico de uma Bolha Financeira (FIGURA 2).

FIGURA 1: PONTOS DA NASDAQ



FONTE: Forbes (2013)

FIGURA 2: GRÁFICO CLÁSSICO DE UMA BOLHA FINANCEIRA



FONTE: Forbes (2013)

Há estudiosos do campo econômico que acreditam que as criptomoedas sejam uma Bolha Financeira. De qualquer forma, discutindo ou não os pontos qualitativos deste fenômeno, há de fato uma grande semelhança entre os gráficos das Figuras 1, 2 e 3.

O gráfico da Figura 3 representa os valores de fechamento do Bitcoin em dólares americanos. Pode-se notar semelhança com uma Bolha Financeira nos

intervalos de tempo entre setembro de 2017 e janeiro de 2018 quando a cotação do Bitcoin subiu cerca de 300,00% e em seguida perdeu cerca de 75% deste valor.

FIGURA 3: COTAÇÃO DO BITCOIN EM DÓLAR



FONTE: CoinMarketCap (2017)

2.3 REDES NEURAIS ARTIFICIAIS (RNA)

2.3.1 DEFINIÇÃO DE REDES NEURAIS ARTIFICIAIS

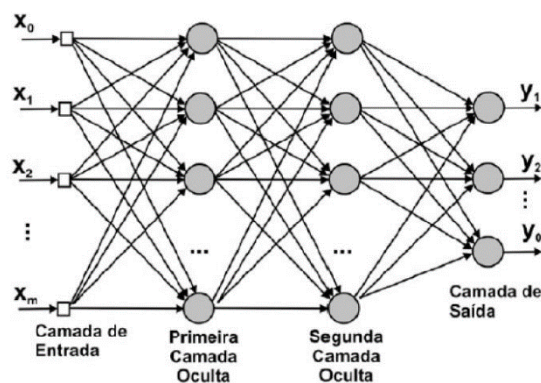
As RNAs (Redes Neurais Artificiais) têm seus primórdios em 1943, quando McCulloch e Pitts estabeleceram as bases da neurocomputação por meio de modelos matemáticos. O desenvolvimento deste campo de estudo, nas décadas seguintes passou a ter grande aplicação em diversos setores, como no Mercado Financeiro, Medicina e Marketing. As RNA's podem ser utilizadas para problemas de classificação ou regressão. Sendo que possuem grande eficiência na resolução de problemas não-lineares (TATIBANA e KAETSU, 2015).

Segundo Tatibana e Kaetsu (2015), a RNA é organizada em três camadas:

1. **Camada de Entrada:** fase de estímulo ou recepção, ou seja, a entrada dos dados. Cada neurônio corresponde à uma variável X_i ;
2. **Camada Oculta:** fase de processamento da informação;
3. **Camada de Saída:** resultado final do processamento.

Com base na Figura 4, é possível visualizar a arquitetura de uma RNA, com suas camadas plotadas.

FIGURA 4: ESTRUTURA DE UMA RNA



FONTE: Tatibana e Kaetsu (2015)

O neurônio de uma Rede Neural Artificial é definido como uma unidade de processamento de informação, sendo imprescindível para o funcionamento da RNA (HAYKIN, 2001).

Matematicamente, o neurônio k é formado por n terminais de entradas (ou variáveis explicativas), sendo estes terminais simbolizados por x_1, x_2, \dots, x_n . Analogamente a um neurônio biológico, estes terminais representariam os dendritos. O axônio, por sua vez, é representado pela variável de saída y .

O comportamento das sinapses de um neurônio biológico é emulado por meio da atribuição de pesos associados, sendo estes simbolizados por w_1, w_2, \dots, w_n . Os índices de w_i , assumindo $i = 1, \dots, n$ representam um comportamento excitatório quando possuem valores positivos e inibitórios quando possuem valores negativos (ROQUE, 2009).

Na RNA, sequencialmente, há a ponderação dos sinais de entrada pelos pesos sinápticos, produzindo uma combinação linear que pode ser definida matematicamente pela Equação (1).

$$\sum x_i w_i \geq \theta \quad (1)$$

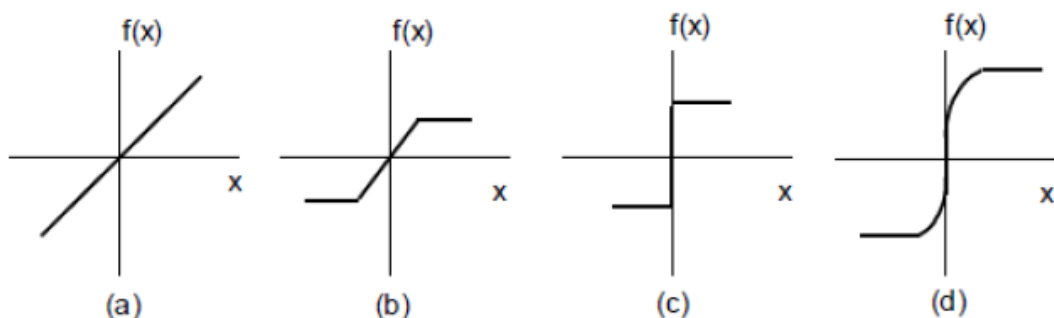
Simbolizado por θ , o limiar quando superado, transfere o resultado do somatório dos produtos dos terminais de entrada com os pesos das sinapses para uma função de ativação. Esta, por sua vez, tem como objetivo restringir a amplitude de saída do neurônio. Assumindo valores geralmente normalizados, o intervalo de restrição pertence ao domínio $[0, 1]$ ou $[-1, 1]$. E as saídas podem ser

interpretadas como estímulo positivo para quando o valor é igual a 1 ou estímulo negativo (inibidor), caso o valor de saída seja igual a -1.

2.3.2 FUNÇÕES DE ATIVAÇÃO

Podendo também serem chamadas por funções restritivas, as funções de ativação podem ser descritas matematicamente por distintas formas. No entanto, a literatura, explora com maior vigor quatro funções de ativação mais promissoras para os estudos das RNA's: (a) função linear, (b) função rampa, (c) função degrau ou step e (d) função sigmoidal (FIGURA 5).

FIGURA 5 – FUNÇÕES DE ATIVAÇÃO DE UMA RNA



FONTE: Roque (2009)

As funções de ativação não são escolhidas especificamente para um tipo de problema. O emprego delas, altera o resultado do estudo e devem ser testadas em diferentes cenários até que seja gerado um resultado com o menor erro. Na academia, há um maior emprego para a função sigmoidal (d), já que ela possui uma transição suave entre os valores de inibição e excitação.

2.3.3 TREINAMENTO DA RNA

Por meio de um processo iterativo de ajustes nos pesos sinápticos, possibilitado por um algoritmo de aprendizagem, a RNA encontra padrões relevantes de informações na base de dados. Desta forma, ocorre o aprendizado por meio da minimização entre a saída fornecida e a saída verdadeira. Segundo

Roque (2009), é no peso sináptico, representado por w_n que a RNA guarda o aprendizado adquirido.

Para o caso em que o estudo se refere à previsão de valores, ou seja, problemas não-lineares de regressão, como é o caso da previsão da cotação de criptomoedas, o aprendizado da RNA pode ser classificado como aprendizagem supervisionada. A base de dados fornecida à Rede Neural é então dividida entre base de treino e base de teste, sendo que a proporção entre estas pode variar.

Segundo Priddy (2005), a base de treino é utilizada para estimar os pesos da RNA e após esta estimação, a base de teste é utilizada exibindo valores em que a rede neural não está familiarizada, para colocá-la em prova.

2.4 O ALGORITMO DO PASSEIO ALEATÓRIO

Há no mercado financeiro uma dualidade entre teorias que dividem acadêmicos entre defensores da preditividade do mercado contra a crença da aleatoriedade nas oscilações dos preços de ativos financeiros. A vertente que acredita que o Mercado é aleatório possui a tese do *Random Walk* (em português, Passeio Aleatório) e tendo como precursor Spitzer (1976). Embora esta teoria tenha sido introduzida há mais de 40 anos, foi precursora de estudos matemáticos que a esmiuçaram em outros estudos atuais.

O raciocínio básico desta teoria pressupõe que a previsão em questão seja composta pelo valor do dia anterior acrescida com um fator de aleatoriedade. Este algoritmo pode ser utilizado para confrontar previsões feitas por meio de métodos quantitativos e, desta forma, validá-las além de medidas de erro já aplicadas, como é demonstrado quantitativamente na Equação (2), desenvolvida por Andrew e Mackinlay (2002):

$$X_t = X_{t-1} + \mu + \epsilon_t \quad (2)$$

Onde:

X_t é o dado previsto;

X_{t-1} é o valor do dia anterior;

μ é uma constante de desvio;

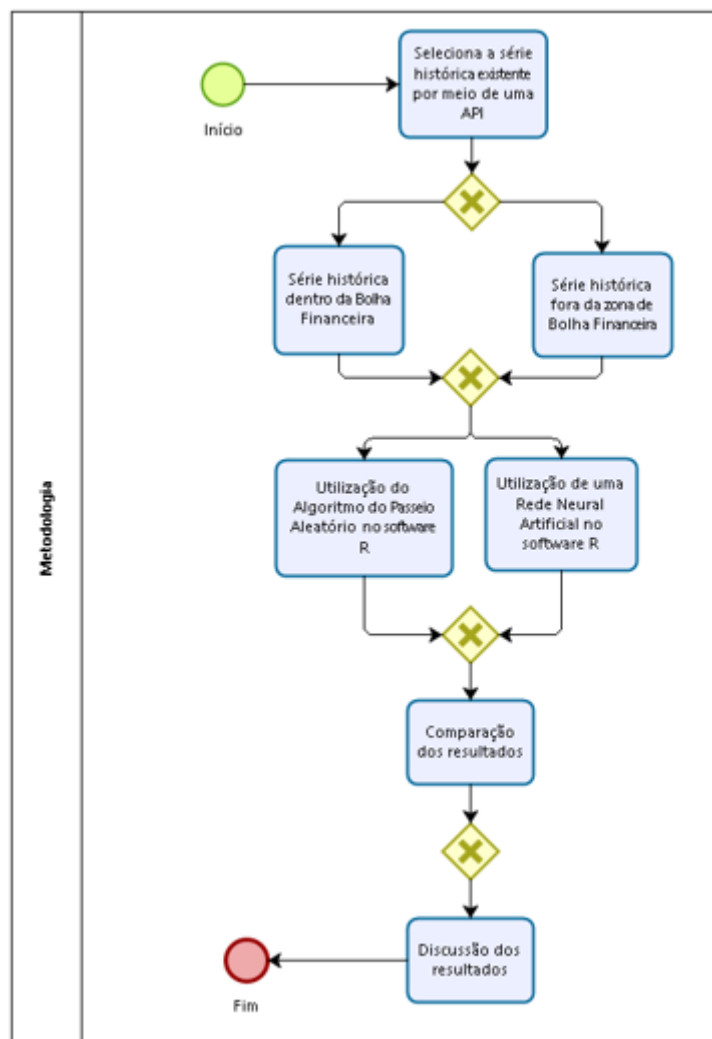
ϵ_t é um número aleatório que segue uma média e desvio padrão.

3 METODOLOGIA

3.1 FLUXOGRAMA DA METODOLOGIA

Por meio do fluxograma da Figura 6, é possível visualizar como foi aplicada a metodologia do trabalho. Desde a extração e segmentação dos dados, até a modelagem matemática, coleta de resultados e discussão.

FIGURA 6: METODOLOGIA



FONTE: O autor (2018)

3.2 O SOFTWARE R

A implementação computacional foi realizada por meio do software R®, e interpretada pela interface R *studio*®. Este software suporta uma série de bibliotecas munidas de funções específicas para cada tipo de modelagem. No atual problema, os algoritmos foram construídos com a utilização das bibliotecas *nnetar*, *forecast*, *xts* e *crypto* (R CORE TEAM, 2018)

A biblioteca *crypto* foi utilizada exclusivamente para a extração dos dados. A biblioteca *nnetar* concentra a inteligência da modelagem e as bibliotecas *forecast* e *xts* são auxiliares a ela.

3.3 BIBLIOTECAS PARA TREINAMENTO DA RNA E PREVISÃO

Após a extração dos dados, estes são transformados em uma série temporal por meio da biblioteca *xts*, então modelados por meio da função *nnetar*, que gera um número pré-determinado de Redes Neurais Artificiais e realiza a média dos pesos sinápticos para a determinação da topologia da rede, para então ativar a função *forecast* e realizar o número de previsões desejadas.

A função *nnetar* gera Redes Neurais com apenas uma camada oculta e com um número de neurônios nesta camada com a metade do número de neurônios de entrada com o acréscimo de um neurônio, como é especificado no manual da biblioteca.

Como na modelagem de séries temporais utilizando Redes Neurais Artificiais não há um roteiro para a determinação da arquitetura ideal para cada tipo de problema de previsão, as arquiteturas são variadas manualmente ou com algumas otimizações para a determinação de um modelo mais ajustado possível. Com a utilização do *nnetar* no lugar de outras bibliotecas também utilizadas para problemas com RNAs, como é de outras funções do R que fazem uso de RNAs, reduz-se a chance de criar um modelo pouco acurado.

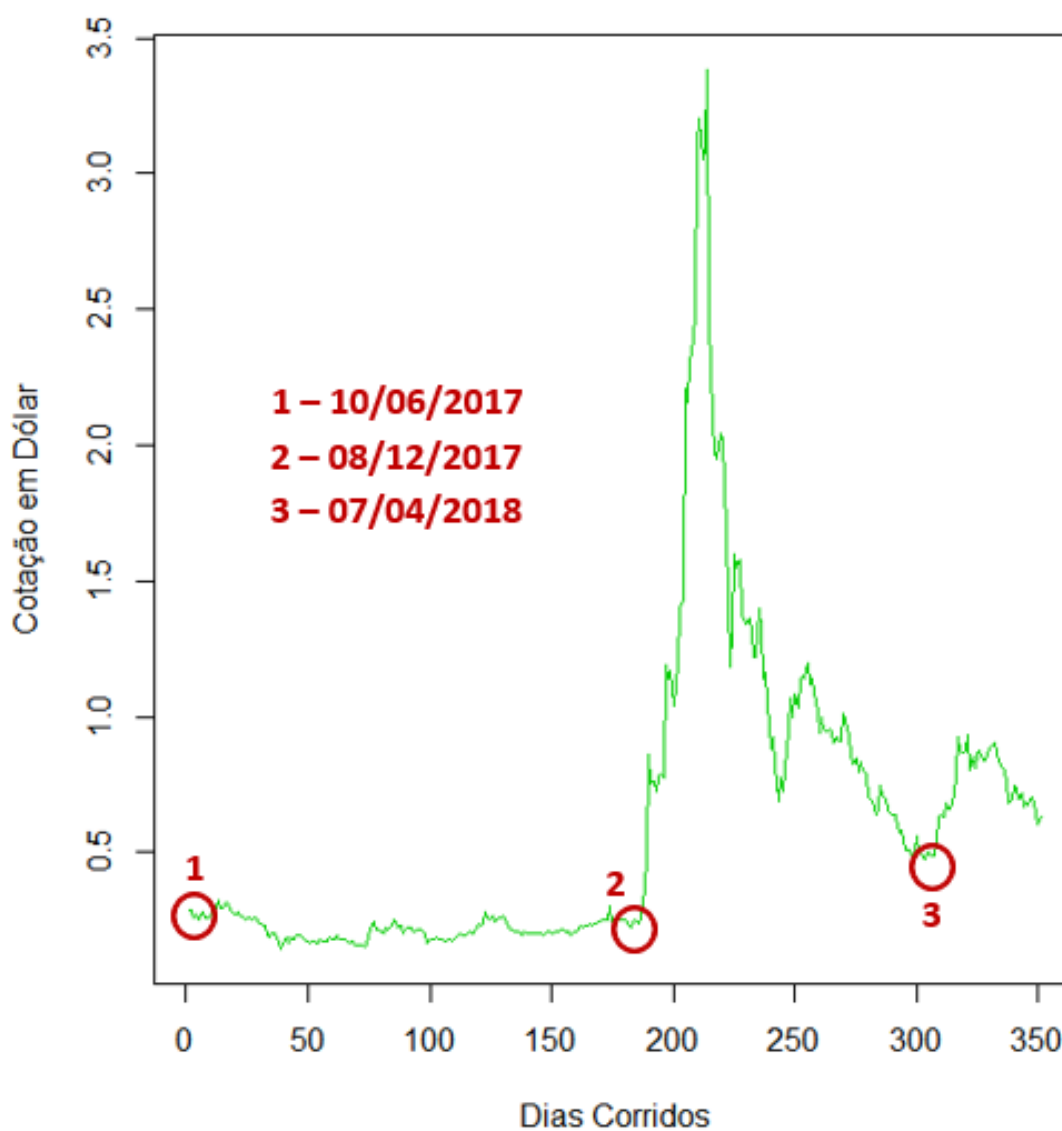
A arquitetura da RNA utilizada no trabalho foi obtida por meio de uma Rede Neural do tipo *Multiple Layer Perceptron*, cuja função de ativação é sigmoïdal, por *default* da função *nnetar()* do R.

3.4 DEFINIÇÃO DOS CENÁRIOS

Os 350 dias de cotação foram segmentados em dois períodos: (1) até (2) e de (2) até (3), como pode ser representado graficamente por meio da Figura 7.

O período de (1) à (2) é o fora da zona da Bolha Financeira, doravante nomeado por Cenário A. Já o período de (2) à (3), doravante nomeado de Cenário B pertence ao intervalo da formação da Bolha Financeira. Desta forma, constroem-se dois cenários em que há um comportamento distinto dos dados para que os métodos sejam testados.

FIGURA 7: VISUALIZAÇÃO DOS DADOS UTILIZADOS



FONTE: O autor (2018)

3.5 SCRIPTS EM R

Os dados que serão utilizados na implementação computacional e análises preditivas, tanto para o método do Passeio Aleatório quanto da Rede Neural Artificial foram extraídos do CoinMarketCap (2017), sendo esta referência uma fonte de cotações em tempo real de todas as criptomoedas negociadas em bolsas de valores ao redor do mundo. A extração dos dados foi realizada por meio de uma API (*Application Programming Interface*), ou seja, um conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseado na Web. Neste caso, um canal de comunicação com uma base de dados do software R utilizando a biblioteca *crypto* (FIGURA 8).

FIGURA 8: SCRIPT EM R

```

2 #Rede Neural Artificial - Cenário A
3
4 install.packages("crypto")
5 install.packages("forecast")
6 install.packages("xts")
7 install.packages("lubridate")
8 library('crypto')
9 library('forecast')
10 library('xts')
11 library('lubridate')
12 set.seed(123)
13 Ripple <- crypto::getCoins(coin = "Ripple")
14 Ripple <- Ripple[,c(4,9)]
15 ripple_xts <- xts(Ripple$close, order.by = Ripple$date)
16 primeiro_dia <- "2017-06-10"
17 ultimo_dia <- "2017-12-08"
18 cenario_a_xts <- ripple_xts[paste0(primeiro_dia, "/", ultimo_dia)]
19 cenario_a <- data.frame(data = as.Date(character()),
20                       previsto = double(),
21                       real = double())
22
23 Acf(cenario_a_xts)
24 dia_predito <- .index(tail(head(cenario_a_xts, 0.8 * length(cenario_a_xts)), 1))

```

FONTE: o Autor (2018)

Primeiramente, foram instaladas as bibliotecas que serão utilizadas, como se observa nas linhas 4 à 7 da Figura 8. Em seguida, estas bibliotecas foram ativadas nas linhas 8 à 11. Para que os resultados do trabalho não fossem afetados pela aleatoriedade, fixou-se uma semente, como observado na linha 12. Desta forma o *script* deve obter os mesmos resultados em qualquer computador.

Após a extração dos dados por meio da função *getCoins* e filtrando apenas a criptomoeda Ripple, é selecionado apenas o Preço de Fechamento da Ripple e sua data, como é possível observar no *script* da Figura 8 nas linhas 13 e 14. A variável *ripple_xts* recebe um tratamento de dados proveniente da bibliotecas *lubridate* e *xts* para converter os dados a uma série temporal (linha 15).

Esta série temporal é então segmentada entre a datas de início e fim de cada um dos cenários de análise. Na Figura 8, as variáveis das linhas 16 e 17 recebem os valores das datas de início e fim e na linha 18 estas datas são acopladas. A estrutura de um *data frame* é criada na linha 19 do *script*, sendo que este *data frame* está vazio em um primeiro momento e posteriormente será preenchido com os resultados produzidos.

Por último, na Figura 8, observa-se na linha 23 a execução da função *Acf* que gera a análise de auto-correlação entre os dias de fechamento da criptomoeda, servindo desta forma para embasar a arquitetura da RNA criada. Na linha 24, há a segmentação da base de dados entre base de treino e base de teste, sendo que esta proporção é de 80% para treino e 20% para teste.

FIGURA 9: SCRIPT EM R

```

26 while (dia_predito < ultimo.dia) {
27   print(paste0("Predizendo o dia ", format(as_datetime(dia_predito), "%d/%m/%Y"), "."))
28   treino_xts <- cenario_a_xts[paste0(primeiro_dia, "/", as_datetime(dia_predito))]
29   teste_xts <- cenario_a_xts[as_datetime(dia_predito + 86400)]
30   rede_neural <- nnetar(treino_xts, repeats = 100, p = 15)
31   predito <- forecast(rede_neural, h = 1)$mean
32   resultados <- data.frame(data = as_datetime(dia_predito + 86400), previsto = as.numeric(predito), real = as.vector(teste_xts[1]))
33   cenario_a <- rbind(cenario_a, resultados)
34   dia_predito <- dia_predito + 86400
35 }
36
37 #Cálculo do erro
38 Real_RNA_A <- cenario_a$real[13:37]
39 Previsto_RNA_A <- cenario_a$previsto[13:37]
40 EQM_RNA_A <- sum(((Real_RNA_A - Previsto_RNA_A) ^ 2) / length(Previsto_RNA_A))

```

FONTE: O autor (2018)

Na continuação do *script* criado para as análises, observa-se na Figura 9 que o código utiliza da estrutura de um *while* para fazer as previsões da base de teste. A linha 27 escreve no console do R qual dia está sendo predito dentro da iteração atual do *while*, na linha 28 a variável *treino_xts* recebe o intervalo de tempo da base de treino e na linha 29 a variável *teste_xts* recebe o dia que será

previsto acrescido de 86400 segundos, o que resulta sempre em um dia acrescido do dia anterior, ou seja, D+1.

Na linha 30 do *script* é o momento em que a RNA é criada com os dados de treinamento. O parâmetro *repeats* determina a quantidade de RNAs criadas e o parâmetro *p* determina o número de neurônios da camada de entrada da Rede. O parâmetro *repeats* neste caso, foi responsável para que 100 Redes Neurais fossem criadas à partir de diferentes pesos aleatórios iniciais, após estas 100 iterações, a média dos pesos sinápticos resultantes provenientes da minimização dos erros por meio do método *backpropagation* é utilizada como os pesos de cada sinapse da RNA que fará a previsão.

Na linha 31 do *script*, a previsão é então realizada. A variável *predito* recebe o valor proveniente da função *forecast* que utiliza a RNA criada na linha 30 e o parâmetro $h = 1$, ou seja, a previsão do dia posterior ao atual (D+1). Na sequência, as linhas 32 à 34 são utilizadas para computarem os dados preditos em uma base de análise e então repetir o *loop* do *script* enquanto o dia predito for menor que o último dia de análise.

O cálculo do Erro Quadrático Médio é realizado nas linhas 38 à 40 do *script*. Como no Cenário B, a quantidade de dias preditos foi de 25 e no Cenário A foi de 37, nas linhas 38 e 39, são selecionados os últimos 25 dias da série de 37 dias preditos. Isto é feito para que haja uma mesma base de comparação entre os dois cenários além da análise de erro dos dois métodos. Por fim, na linha 40 é realizado o cálculo do EQM.

FIGURA 10: SCRIPT EM R

```

44 #Passeio Aleatório - Cenário A
45
46 Ripple <- crypto::getCoins(coin = "Ripple")
47 DailyClose <- Ripple$close
48 new_daily <- DailyClose[1406:1587]
49 n<-length(new_daily)
50 returns.daily.simple <- new_daily[-n]/new_daily[-1]-1
51 mu <- mean(returns.daily.simple)
52 sd <- var(returns.daily.simple)
53 RW <- new_daily[-1] + rnorm(new_daily[-1],mu,sqrt(sd))
54
55 #Calculo do Erro
56 RealRW<-new_daily[157:181]
57 PrevistoRW<-RW[157:181]
58 EQM_RW_A <- sum((PrevistoRW - RealRW) ^ 2) / length(RealRW)
59
60
61 #Faz o gráfico
62 plot(Real_RNA_A, type = 'l', col="red", lwd=2, xlab='Dias corridos', ylab='Fechamentos',ylim=range(Real_RNA_A,Previsto_RNA_A,PrevistoRW))
63 lines(Previsto_RNA_A,col="green",lwd=2)
64 lines(PrevistoRW,col="purple",lwd=2)
65 legend("bottomleft",legend = c("Real","Previsão RNA","Previsão RW"),col=c("red","green","purple"),lty=1,lwd=2)

```

FONTE: O autor (2018)

Na continuação do *script* criado para as análises, observa-se na Figura 10 a aplicação do método do Passeio Aleatório. Nas linhas 46 à 48 são extraídos novamente os dados de fechamento do Ripple e então a série temporal é seccionada com os dados do Cenário A. Na sequência, a variável n na linha 49 recebe o tamanho do vetor da série temporal do Cenário A que é utilizado na linha 50 para o cálculo do retorno diário do ativo. A variável μ recebe a média dos retornos diários na linha 51, a variável sd recebe a variância dos retornos diários na linha 52 e, por último, na linha 53 ocorre a execução do método do Passeio Aleatório.

Do mesmo modo como ocorrido na análise utilizando a RNA, no caso do Passeio Aleatório são selecionados os últimos 25 dados da série para o cálculo do Erro Quadrático Médio, como pode-se observar nas linhas 56, 57 e 58 do *script*.

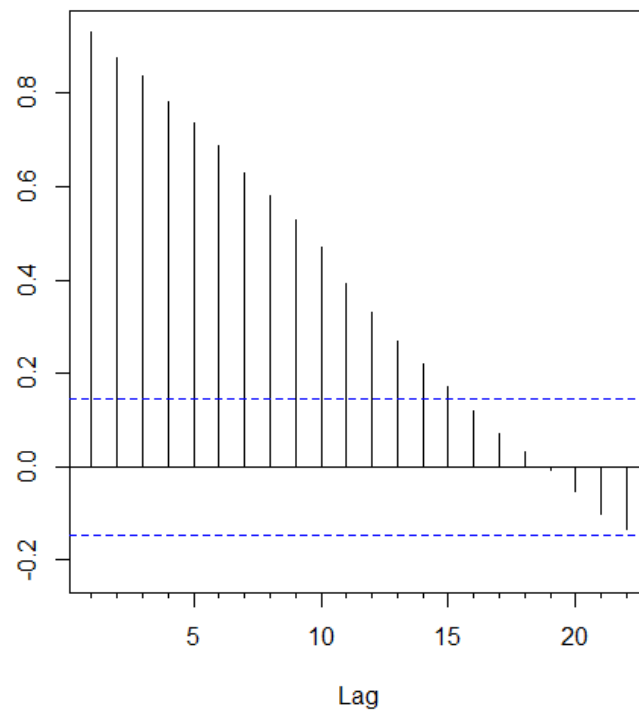
Por fim, as linhas 62 à 65 criam os gráficos para a visualização entre os dados reais da série temporal e os dados preditos pelos métodos da RNA e Passeio Aleatório.

3.6 TOPOLOGIA DA RNA

Não há uma fórmula para a determinação da arquitetura correta de uma RNA, ou seja, de uma seleção precisa do número de neurônios de entrada, número de camadas ocultas ou neurônios na camada oculta. No entanto, é possível observar a correção dos dados de uma série temporal para então tomar a decisão de quantos neurônios devem ser utilizados na camada de entrada da RNA (HYNDMAN, 2015).

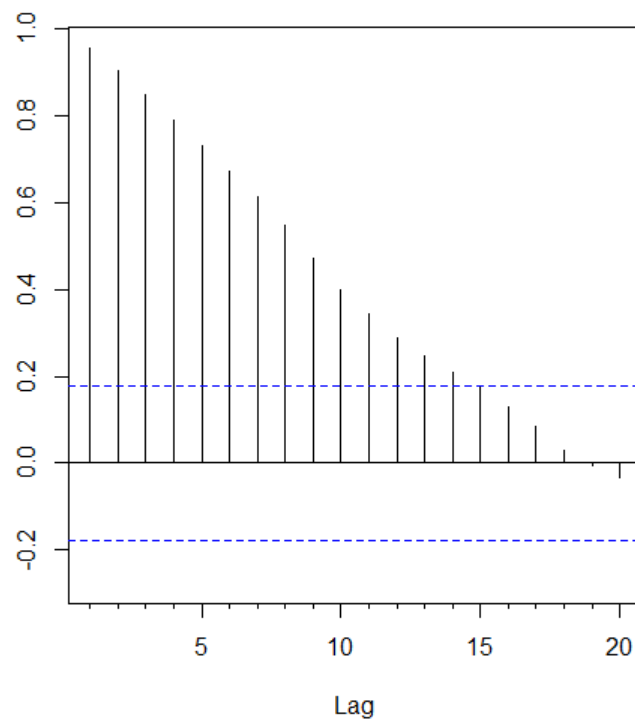
Esta análise foi realizada separadamente nos cenários A e B utilizando a função *Acf* do software R. Deste modo, pode-se observar graficamente a auto-correlação entre os dias de fechamento da criptomoeda Ripple (FIGURAS 11 e 12).

FIGURA 11: AUTOCORRELAÇÃO DO CENÁRIO A



FONTE: O Autor (2018)

FIGURA 12: AUTOCORRELAÇÃO DO CENÁRIO B



FONTE: O Autor (2018)

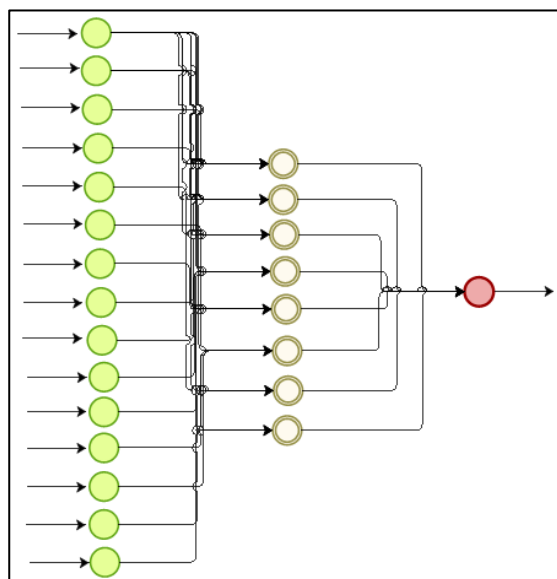
A análise de auto-correlação, ou seja, a influência da cotação futura entre os dias anteriores, por meio da função *Acf* proporciona direcionar a modelagem da arquitetura da RNA para um número de neurônios menor ou igual ao limiar superior representado pela linha azul hachurada, sendo esta plotada automaticamente pela função. Ou seja, o número de neurônios de entrada deve ser menor ou igual à 15. Cada neurônio de entrada é um dia anterior à previsão, sendo a camada de entrada composta por dados de D-1 até D-15 e a previsão é em D+1.

Desta forma, em ambos os cenários, a RNA foi estruturada para possuir 15 neurônios de entrada, sendo implementada por meio do código:

```
rede_neural <- nnetar(treino_xts, repeats = 100, p = 15)
```

em que p é o parâmetro que determina o número de neurônios de entrada. O número de neurônios da camada oculta é 8, como determinado automaticamente pela função *nnetar* (neurônios da camada oculta = ((neurônios da camada de entrada/2) + 1). Além disso, há apenas uma camada oculta, como determinado pelo pacote *nnetar* e o número de neurônios de saída é 1, representando o dia predito.

FIGURA 13: REPRESENTAÇÃO DA ARQUITETURA DA RNA



FONTE: O Autor (2018)

3.7 INTERPRETAÇÃO DAS ANÁLISES E CENÁRIOS GERADOS

Como resultado da análise, espera-se comparar qual método de previsão possui o menor erro quadrático médio em cada um dos cenários: Cenário A e Cenário B.

Como métrica quantitativa, foi implementado em R a Equação (3) do Erro Quadrático Médio:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (3)$$

Onde:

n : tamanho da série temporal;

i : linha de referência da série temporal;

y_i : valor i verdadeiro;

$\hat{f}(x_i)$: valor predito.

Espera-se, desta forma, chegar à conclusão de qual método possui o menor Erro Quadrático Médio no Cenário A e no Cenário B.

4 RESULTADOS

Após os *scripts* em R serem executados, os dados foram analisados para que houvesse a comparação tendo como base o Erro Quadrático Médio. Como a base de treinamento da RNA foi de 80% para ambos os cenários, a base de validação que foi considerada para fins de análise como os dados previstos que tem tamanho de 37, sendo cada dia predito um de cada vez para o Cenário A e de 25, sendo cada dia predito um de cada vez para o Cenário B. Por conta disto, tanto na análise do Passeio Aleatório, quanto da RNA, foram apenas utilizados para o cálculo do Erro Quadrático Médio os últimos 25 resultados de cada série. Desta forma, garantindo que não houvesse distorção na comparação dos resultados.

Conforme Quadro 1, A RNA mostrou-se superior ao método do Passeio Aleatório no Cenário A. Já no Cenário B, o Passeio Aleatório mostrou-se como método mais assertivo. A RNA obteve resultados 16,10 vezes mais assertivos que o Passeio Aleatório no Cenário A. No Cenário B, o Passeio Aleatório obteve resultados 2,04 vezes mais assertivos que a RNA.

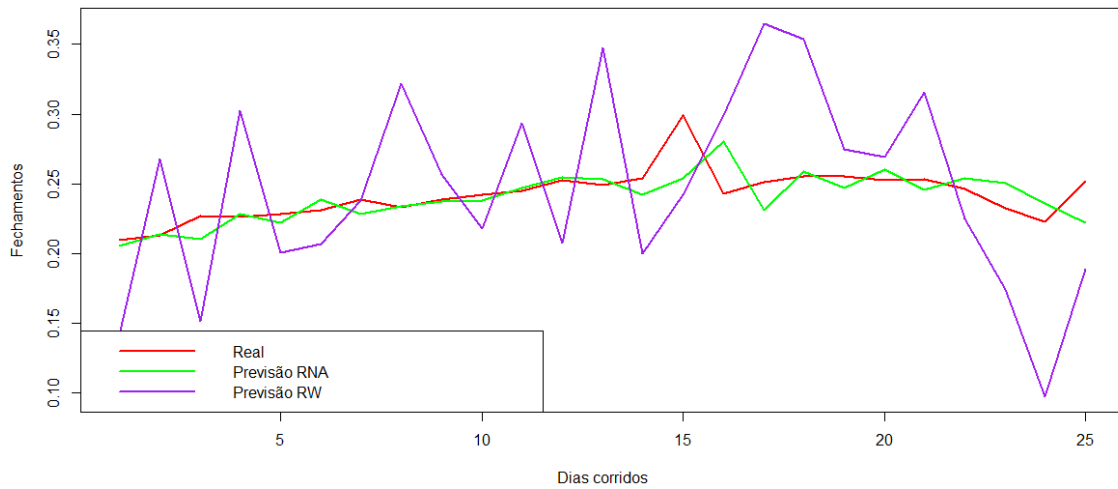
QUADRO 1: RESULTADOS COMPARATIVOS

	Passeio Aleatório	RNA	Delta
Cenário A	0,004009550	0,00024908	16,10x
Cenário B	0,007414807	0,01515618	2,04x
Delta Cenários	1,85x	60,85x	

FONTE: O Autor (2018)

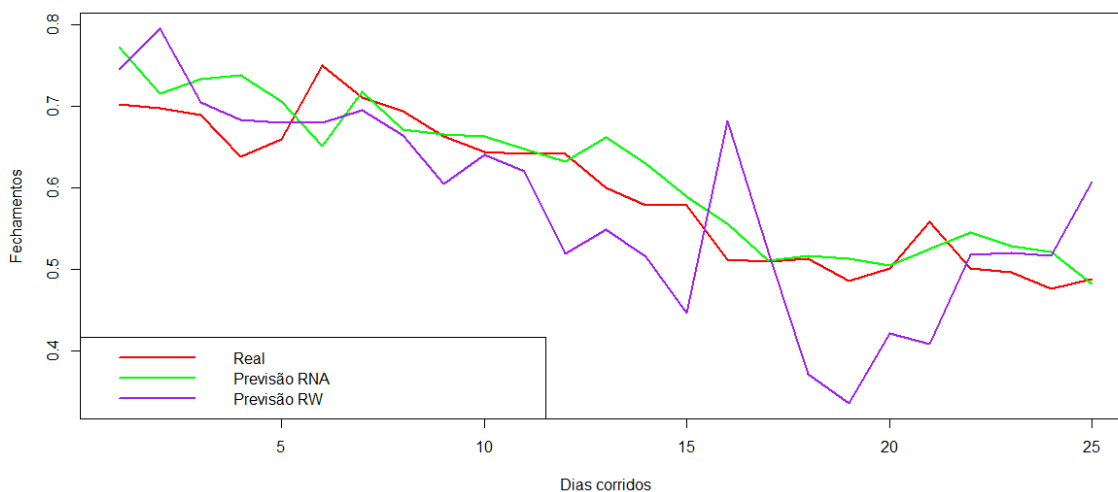
Ainda de acordo com o Quadro 1, com relação à comparação entre cenários e tendo como base o mesmo método, observou-se maior erro para os dois métodos no Cenário B. O método do Passeio Aleatório obteve uma menor variação entre os dois cenários (1,85 vezes), já a RNA obteve uma variação de 60,85 vezes no comparativo dos cenários.

FIGURA 14: RESULTADOS CENÁRIO A



FONTE: O Autor (2018)

FIGURA 15: RESULTADOS CENÁRIO B



FONTE: O Autor (2018)

É possível notar graficamente por meio das Figuras 14 e 15 o comportamento dos métodos durante suas previsões. Na Figura 14 nota-se a maior aderência da RNA no Cenário A, o que deixa de ocorrer na Figura 15, representada pelo Cenário B.

5 CONCLUSÃO

O objetivo deste trabalho foi compreender o funcionamento de uma classe de Rede Neural Artificial, a MLP, assim como do algoritmo do Passeio Aleatório, aplicando ambos à previsão da cotação da criptomoeda Ripple. Por tratar-se de um método capaz de encontrar padrões, esperava-se que a RNA superasse o Passeio Aleatório nos dois cenários gerados: Cenário A e Cenário B. Além disto, o trabalho teve como intuito investigar o impacto da volatilidade ocasionada por uma Bolha Financeira com o cálculo do erro dos dois métodos.

Na análise do Cenário A, que envolvia uma série temporal fora de uma Bolha Financeira, a RNA obteve um EQM de 0,00024908 contra os 0,004009550 do Passeio Aleatório, o que resulta em 16,10 vezes mais assertividade. Já no Cenário B, de Bolha Financeira, a RNA obteve resultados menos assertivos que o algoritmo do Passeio Aleatório. O EQM do Passeio Aleatório neste cenário foi de 0,007414807 contra os 0,01515618 da RNA, representando 2,04 vezes mais assertividade para o Passeio Aleatório.

Quanto à comparação dos erros entre o mesmo método em cenários diferentes, observa-se uma maior variação do EQM para a RNA do que para o Passeio Aleatório. A volatilidade de Bolha Financeira ocasionou um crescimento de 60,85 vezes no EQM da RNA ante o crescimento de apenas 1,85 vezes do EQM do Passeio Aleatório. Possivelmente, o Passeio Aleatório obteve uma menor variação em seu EQM devido à variável de volatilidade que é incorporada ao método.

Assim, a RNA não obteve o melhor resultado nos dois cenários e teve o seu EQM aumentado consideravelmente em um cenário de Bolha Financeira. Contudo, a RNA mostrou-se adequada à um cenário de previsão que possua menor volatilidade.

5.1 SUGESTÕES E PRÓXIMOS PASSOS

Este projeto teve como um intuito prever a cotação da criptomoeda Ripple utilizando Redes Neurais Artificiais e o método do Passeio Aleatório. Com base no *script* desenvolvimento em R para a implementação dos algoritmos, há a possibilidade da continuação do estudo analisando outros períodos de cotação com apenas poucas modificações do código em R. Além disto, o código é

facilmente adaptado para analisar outras criptomoedas, basta alterar o filtro de “Ripple” para qualquer outra criptomoeda (COINMARKETCAP, 2017).

Em trabalhos futuros há a possibilidade da análise de outros cenários de Bolhas Financeiras, de outras criptomoedas, da maior variação dos parâmetros que formam a arquitetura da RNA. Há também a possibilidade da variação de outras classes de dados para análise, como o preço de abertura do ativo, volume negociado nas últimas 24 horas, dentre outros.

Por fim, a aplicação da RNA poderia ter sido direcionada para a identificação de padrões de Bolha Financeira ao invés apenas da previsão da criptomoeda. Há, portanto, uma rica gama de aplicações das Redes Neurais Artificiais no comportamento das criptomoedas.

6 REFERÊNCIAS

ALBUQUERQUE, M. Pontes. **Previsão do Mercado de Ações Brasileiro utilizando Redes Neurais Artificiais**. Rio de Janeiro: CBPF, 2008.

ANDREW, L. W., MACKINLAY, A. C. **A Non-Random Walk Down Wall Street**, 5ª ed, 2002.

COINMARKETCAP. **Cryptocurrency Market Capitalizations**. 2017. Disponível em: <coinmarketcap.com>. Acesso: 10/12/2017.

FORBES. **Bitcoin May Be Following This Classic Bubble Stages Chart**. 2013. Disponível em: <<https://www.forbes.com/sites/jessecolombo/2013/12/19/bitcoin-may-be-following-this-classic-bubble-stages-chart/#749ce2c836b8>>. Acesso em 04/06/2018.

HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre: Bookman, 2001.

HYNDMAN, R.J., **High-dimensional autocovariance matrices and optimal linear prediction**. Electronic Journal of Statistics, 9, 792-796, 2015.

MCCULLOCH W., PITTS W. **A Logical Calculus of the Ideas Immanent in Nervous Activity**, Bulletin of Mathematical Biophysics, 5, p. 115-133, 1943.

MUELLER, A. **Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário**. Dissertação (Mestrado em Engenharia de Produção) – Departamento de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1996.

NAKAMOTO, S. **Bitcoin: A Peer-to-Peer Electronic Cash System**, 2008. Disponível em: <https://bitcoin.org>. Acesso em 17/11/2017.

PHILLIPS, P. C. B., YU, J. **Dating the timeline of financial bubbles during the subprime crisis. Quantitative Economics**, 2, 455-491, 2011.

PRIDDY, L. K. **Artificial Neural Network**. Washington: The International Society for Optical Engineering, 2005.

R Core Team (2017). R: **A language and environment for statistical computing. R Foundation for Statistical Computing**, Vienna, Austria. URL <https://www.R-project.org/>.

REFENES, A. N., ALIPPI, C. **Histological Image Understanding by Error Backpropagation**. North-Holland: Microprocessing and Microprogramming, 1993.

RIPPLE. **Ripple**. 2018. Disponível em: <ripple.com>. Acesso: 10/12/2017.

ROQUE, R. C. **Estudo sobre a empregabilidade da previsão do índice BOVESPA usando Redes Neurais Artificiais**. 2009. Dissertação (Mestrado em Engenharia Eletrônica) – Departamento de Eletrônica e de Computação, Universidade Federal do Rio de Janeiro.

SPITZER, F. **Principles of Random Walk**, Springer-Verlag, 1976.

TATIBANA, C. Y.; KAETSU, D. Y. **Uma Introdução às Redes Neurais**. 2015. Disponível em: < <http://www.din.uem.br/ia/neurais/> >. Acesso em: 18/11/2017.