

ESTRATÉGIAS DE MERCADO ACIONÁRIO UTILIZANDO PREVISÃO DE REDES NEURAIS EM COMPARAÇÃO COM MODELOS AUTORREGRESSIVOS

STRATEGIES OF STOCK MARKET USING PREDICTION OF NEURAL NETWORK IN COMPARISON WITH MODELS AUTOREGRESSIVE

Data de submissão: 17-08-2011

Aceite: 18-08-2012

André Pacheco Miranda¹

Paulo Sergio Ceretta²

Luis Felipe Dias Lopes³

RESUMO

Quando investidores decidem se aventurar pelo mercado de renda variável, como no mercado de ações, buscam um método que proporcione mais segurança durante a tomada de decisão. Na prática, não há como saber quais ativos se tornarão um investimento lucrativo nem se a predição de um método é melhor que a de outro. Diante disso, este artigo apresenta o desenvolvimento de um método heurístico, utilizando como variáveis de transição o volume negociado e o próprio retorno defasado. Após o desenvolvimento e a aplicação da rede neural, o resultado é confrontado com a predição de um modelo linear, utilizando alguns critérios de avaliação propostos posteriormente. O método heurístico, que conta com uma rede neural *multilayer perceptron*, treinada com o algoritmo de retropropagação de erro, foi comparado com um modelo autorregressivo de médias móveis (ARMA). Com base nos resultados, pode-se perceber que, embora os dois procedimentos tenham um desempenho satisfatório, a rede neural possui um poder de explicação maior do que aquele dos modelos ARMA.

Palavras-chave: Redes Neurais. Trade System. Modelos ARMA. Multilayer Perceptron.

¹Possui graduação em Ciência da Computação pela Universidade de Cruz Alta - UNICRUZ, Mestrado em Administração pela Universidade Federal de Santa Maria - UFSM. Santa Maria. Rio Grande do sul. Brasil. E-mail: andmirandapippi@gmail.com

²Possui graduação em Administração pela Universidade Federal de Santa Maria - UFSM, mestrado em Engenharia de Produção pela Universidade Federal de Santa Maria e doutorado em Engenharia de Produção pela Universidade Federal de Santa Catarina - UFSC. Atualmente é professor na Universidade Federal de Santa Maria – UFSM. Santa Maria. Rio Grande do Sul. Brasil. E-mail: ceretta10@gmail.com

³Possui graduação em Matemática Licenciatura Plena pela Faculdade Imaculada Conceição – FIC, mestrado em Engenharia de Produção pela Universidade Federal de Santa Maria – UFSM e doutorado em Engenharia de Produção pela Universidade Federal de Santa Catarina – UFSC. Atualmente é Chefe do Departamento de Administração e professor da Universidade Federal de Santa Maria. Santa Maria. Rio Grande do Sul. Brasil. E-mail: lflopes67@yahoo.com.br

ABSTRACT

When investors decide to venture into stock markets they search for a method that provides reliability to support their decision making process even though there is no means to know empirically with certainty which stocks will become profitable investments and which prediction method is the best to discover this. This article presents the development of a heuristic method that employs the trading volume and return lagged as transitions variables. Using assessment criteria proposed further, the result obtained by the development and application of the neural network is compared with the prediction extracted from a linear model. The heuristic method, that has a neural network multilayer perceptron trained with an algorithm for back propagation error, is then compared with an autoregressive moving average (ARMA) model. The results point out that the neural network offers a greater explanation power than ARMA models, even though neither approach presents a satisfactory performance.

Keywords: Neural Networks. Technical Analysis. Trade System. ARMA Models. Multilayer Perceptron.

1 INTRODUÇÃO

O uso de métodos e modelos matemáticos como forma de antever as possíveis movimentações do mercado financeiro não é uma prática nova. Nas últimas décadas, essa foi uma área de grande interesse para a comunidade de Inteligência Artificial (KIMOTO, 1990; STEINER, 1995; LAWRENCE, 1997; CERETTA et al., 2010).

Sistemas para a predição do mercado acionário vão desde as mais simples estatísticas, presentes nas análises técnicas (ROBERT, 2007), tais como a Teoria de Adam e a Teoria das Ondas de Elliot (NORONHA, 2009), até modelos econométricos AR, MA, ARMA (COELHO et al., 2008; OLIVEIRA et al., 2010), modelos de Markov (HASSAN, 2005) e outros modelos (LAWRENCE, 1997; CERETTA et al., 2010; OLIVEIRA et al., 2010). Entretanto, os modelos matemáticos mais utilizados para prever o retorno são as redes neurais e os modelos econométricos AR, MA, ARMA, como pode ser visto em Lawrence (1997), Kim (2004), Hassan (2005), Phua (2000), Smith e Gupta (2000), Ceretta et al., (2010), Oliveira et al. (2010) e Maia et al. (2008), apenas para citar alguns exemplos.

O comportamento não linear, aleatório e altamente complexo do mercado acionário configura as redes neurais como um dos modelos matemáticos mais adequados para a sua modelagem. Porém, as principais dificuldades em seu uso resultam da pouca familiarização entre os pesquisadores das áreas de ciências humanas, pois se trata de um método mais moderno do que a maioria dos métodos estatísticos e econométricos (CUNHA et al., 2010). Em contrapartida, os modelos lineares, que são mais simples, apresentam um poder computacional próximo ao das redes neurais (MAIA et al., 2008; RABONI, 2005).

Tais modelos de previsão do mercado financeiro podem ser aplicados de diversas formas. Alguns utilizam como informações de entrada resultados da análise fundamentalista; outros se baseiam apenas em informações básicas, como cotação da ação, retorno da ação e volume das negociações; e outros utilizam informações extraídas da análise técnica (LAWRENCE, 1997). Tendo em vista essa diversidade de modelos e variáveis de entradas, questiona-se: um método de previsão mais complexo, como as redes neurais, pode obter um melhor desempenho em comparação a um método econométrico ao utilizar o retorno do volume como variável de entrada?

Com o intuito de atender ao problema de pesquisa delineado neste estudo, objetivou-se desenvolver duas ferramentas matemáticas de previsão para auxiliar na tomada de decisão do investidor (pessoa física). Para isso, o presente artigo apresenta o desenvolvimento de dois modelos de previsão a fim de comparar qual método tem mais pré-disponibilidade de predição, utilizando o retorno do volume para alimentação das ferramentas matemáticas. Logo, as saídas do modelo que obtiver o melhor desempenho sugerirão ao usuário pontos de compra, venda ou manutenção de um ativo em carteira.

Um dos modelos desenvolvidos utiliza uma rede neural *perceptron* de múltiplas camadas (MLP), treinada com o algoritmo de retropropagação do erro (*backpropagation*) (HAYKIN, 2001; SCHLKOFF, 2001; NASCIMENTO J.R., YONEYANA, 2004; LUDWING JR.; MONTGOMERY, 2007; VALENÇA, 2009; SOBREIRO et al., 2009; GALÃO et al., 2011; HAYKIN, 2001) e tendo como saída desejada o dia seguinte à análise do índice Ibovespa. O segundo modelo de previsão desenvolvido consiste em um dos modelos autorregressivos de médias móveis (ARMA) (CAMPOS, 2008).

Para validação dos dois métodos, são realizadas baterias de testes utilizando um total de 3332 cotações situadas entre 02/01/1998 e 16/06/2011. Também são realizados testes, para comparar a previsão dos métodos já citados, com o critério do menor Erro Médio Quadrático e o Coeficiente de Explicação R^2 ajustado mais próximo de 1.

Tendo em vista os objetivos propostos, este artigo está estruturado em cinco itens, além desta introdução. Na seção 2, são descritas técnicas de análise financeira; na seção 3, realiza-se uma breve descrição de redes neurais; na seção 4, apresenta-se o sistema proposto; na seção 5, alguns resultados são discutidos; e, por fim, na seção 6, são expostas algumas conclusões.

2 MODELOS LINEARES

Nesta seção, são apresentados três modelos lineares de previsão do retorno no mercado acionário brasileiro e seus desenvolvimentos a partir de uma série real, para fins de comparação com a rede neural. O primeiro modelo apresentado consiste no autorregressivo (AR), um modelo que é desenvolvido, principalmente, para estimar o retorno de uma ação, utilizando determinadas séries temporais lineares (MAIA et al., 2008). Os modelos AR de Box e Jenkins (1976) são definidos de maneira que os valores das séries no tempo t dependem dos valores passados. Então, diz-se que Y_i (objeto a ser explicado) é explicado por X_i e seus consequentes, ou seja, é transformado em um processo autorregressivo de ordem p que satisfaz à seguinte equação de diferenças:

$$Y_i = \sum_{i=1}^p \alpha_i X_{i-1} + \sum_{i=1}^p \beta_i T_{i-1} + \sum_{i=1}^p \beta_i C_{i-1} + \varepsilon_i \quad (1)$$

Fonte: os autores.

Onde, α_i é a constante; $X_{i-1}, T_{i-1}, C_{i-1}$ são os parâmetros definidos para explicar Y_i ; p é número de termos autorregressivos; e ε_i é o ruído branco da equação. Em modelos AR(p), os operadores autorregressivos iguais a zero, de tal forma que ($j > p$), são estacionários de ordem p . A equação 1 que representa modelos da família AR também simula o efeito da magnitude do retorno de Y , o que indica que, quanto maior forem os parâmetros definidos, maior será o valor de Y .

O segundo modelo apresentado é o modelo de médias móveis de ordem q , representado por MA(q), onde os operadores de médias móveis são iguais a zero, de tal forma que ($j > q$) é considerado linearmente dependente de um número finito, q , de ruídos brancos. Isso significa que, em modelos de médias móveis de ordem q , Y_i representa o modelo linear, porém com o somatório em q termos (GALVÍNCIO et al., 2002). A equação 2 representado o operador de média móvel de ordem q :

$$Y_i = 1 - \sum_{i=1}^q \theta_q X_{i-1} + \varepsilon_i \quad (2)$$

Fonte: os autores.

Onde ϑ é o coeficiente de médias móveis, ou seja, determina o comportamento da variância condicional no sinal dos retornos positivos; ε_t é o ruído branco ou erro; e Y_t é a vazão diária no dia $t+1$. Assim as equações 1 e 2 são estacionárias se e somente se as suas raízes forem unitárias. O teste a ser analisado é o de Durbin-Watson (1950), que avalia a autocorrelação entre todos os valores das variáveis de todos os construtos, pois, em um modelo, pressupõe-se que o termo de erro de qualquer variável não é influenciado pelo termo de erro de suas conseqüentes nem o influencia. Para a rejeição de H_0 , os resíduos devem ser não autocorrelacionados, e o critério para solução do problema é a exclusão de variáveis problemáticas dos construtos iniciais.

O terceiro modelo a ser apresentado neste trabalho é o modelo ARMA (autorregressivos de médias móveis) de ordem (p, q) , considerado como uma generalização dos modelos $AR(p)$ e $MA(q)$. Os modelos ARMA são uma classe de modelos muito útil para descrever dados de séries temporais. Geralmente, são mais utilizados em algumas áreas das ciências humanas, onde é natural pensar no valor de alguma variável no instante t com função de valores defasados da mesma variável (MAIA et al., 2008). Portanto, um processo autorregressivo e de médias móveis, de ordem (p, q) , denotado por $ARMA(p, q)$ é definido por:

$$Y_t = \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (3)$$

Fonte: os autores.

Sendo ε_{t-q} o ruído branco; μ a média do processo, cabe salientar que as funções de autocorrelação ficam consideravelmente mais complicadas em processos de ARMA. De forma geral, para um processo $ARMA(p, q)$ estacionário, a função de autocorrelação tem um decaimento exponencial ou oscilatório após a defasagem q (IQUIAPAZA et al., 2010), o que indica que o modelo ARMA é considerado estacionário somente até certo grau na curva da função de autocorrelação.

3 REDES NEURAIS ARTIFICIAIS

Para entender o funcionamento básico de uma rede neural, é necessário compreender, antes, o funcionamento do cérebro humano e seus componentes, pois uma rede neural artificial é um modelo formal para representar o neurônio biológico (JUDD, 1990). Assim, percebendo esse funcionamento e baseado nos principais componentes de um neurônio biológico, Warren McCulloch, neurofisiologista, e Walter Pitts, matemático, de forma simples, desenvolveram o neurônio matemático, representando-o por meio de uma regra de propagação e uma função de ativação (VALENÇA, 2009). Eles combinaram os conceitos trazidos pela Neurofisiologia, como o processo de aprendizagem de um neurônio biológico, e pela Matemática, como pesos e *thresholds*, para desenvolver o neurônio matemático (MACCULLOCH; PITTS, 1943).

O ciclo de funcionamento de um neurônio biológico é composto, basicamente, de três etapas: receber os estímulos enviados pelas *sinapses* dos sensores (papilas gustativas, pele etc.), se o sinal for superior ao limiar, segue pelo axônio e é repassado para o corpo celular (LUDWIG JR.; MONTGOMERY, 2007); compilar o estímulo e fazer os cálculos necessários; e repassar, por meio dos dendritos, o resultado calculado pelo corpo celular para o próximo neurônio (VALENÇA, 2009). Para melhor entendimento de uma rede neural, a seguir, a Figura 1 apresenta a descrição, com base no neurônio biológico, de alguns dos principais componentes do neurônio matemático presentes em Ludwig Jr. e Montgomery (2007), Nascimento Jr. e Yoneyana (2004) e Ceretta et al. (2010).

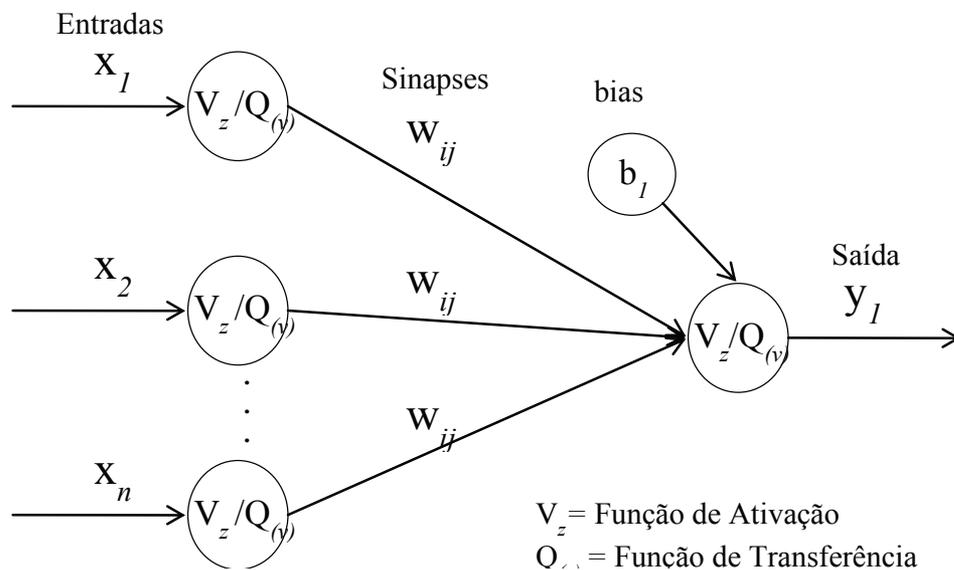


Figura 1: Neurônio Matemático

Fonte: Ludwing Jr. e Montgomery (2007); Nascimento Jr. e Yoneyana (2004); e Ceretta et al. (2010).

Onde: as entradas estão representadas em x_n , as *sinapses* em w_{ij} , o bias em b_1 e a saída em y_n . Na nomenclatura utilizada, a letra j representa o neurônio que recebe o sinal (saída), e a letra i o neurônio que emite o sinal (entrada). A letra n representa o número de entradas, s o número de neurônios na camada de saída e y_1 a saída $f(V_z)$ de cada neurônio da camada de saída.

No neurônio matemático, as *sinapses* desempenham a função dos dendritos no neurônio biológico. O papel dos axônios é desempenhado no modelo matemático pelo *bias* (LUDWING JR.; MONTGOMERY, 2007; FARIA, et al., 2008). O corpo do neurônio biológico, onde são processados os estímulos, é responsabilidade da função de transferência e de ativação em um neurônio (VALENÇA, 2009).

A função de ativação é responsável pela soma ponderada dos sinais de entrada para simular a primeira etapa dentro do neurônio. Entretanto, somente a função de ativação não fará o papel do corpo do neurônio. Para isso, é necessária a combinação com a função de transferência no caso sigmóide, descrita na equação 4, a seguir (LUDWING J.; MONTGOMERY, 2007). Onde: $Q(V_z)$ é o valor do resultado da função de transferência; V_z é o valor do resultado da função de ativação; e e é a base natural de logaritmos.

$$Q(V_z) = \frac{1}{(1 + e^{-[(V_z)])}} \tag{4}$$

Fonte: Valença (2009)

O grande papel da função de transferência, após a ativação, é determinar a saída do neurônio em função da soma ponderada (SCHALKOFF, 2001), evitando que os valores de saída se afastam do intervalo de valores máximos e mínimos determinados.

3.2 Redes neurais *perceptron*

Basicamente, todas as redes neurais são compostas de neurônio, camadas e sinapses. A grande diferença de uma rede neural para outra reside, contudo, nas suas formas de treinamento/aprendizado e de correção das sinapses pelo erro. Então, o número de neurônios e camadas de uma Rede Neural Artificial depende da sua funcionalidade ou do seu propósito.

Segundo Yonenaga e Figueiredo (1999), em 1958, Frank Rosenblatt construiu um dos primeiros tipos de redes neurais, a rede neural artificial *perceptron*. No seu trabalho, intitulado “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain”, Rosenblatt desenvolveu uma simulação computacional, confirmando que o reconhecimento de padrões é uma das características da retina.

Essa rede desenvolvida por Frank Rosenblatt atualiza as sinapses com um método de ajuste supervisionado quando o supervisor ajusta os parâmetros modificáveis da rede neural com base no valor do erro (YONENAGA; FIGUEIREDO, 1999). A menor representação da classe *perceptron* como ajuste supervisionado é um neurônio na camada de entrada ligado a um neurônio na camada de saída.

Entretanto, logo após a publicação do trabalho de Frank Rosenblatt, em 1958, Marvin Minsky e Seymour Papert descobriram e provaram, no trabalho intitulado “Perceptrons”, que as redes neurais de uma única camada não são tão eficientes assim (LUDWIG JR.; MONTGOMERY, 2007). Eles confirmaram que a *perceptron* não consegue resolver algumas situações em que ocorre o problema da resolução da operação da função lógica (OU) exclusivo (XOR) (VALENÇA, 2009). Com a publicação do trabalho de Minsky e Papert, muitos pesquisadores na área de redes neurais da época foram desestimulados (NASCIMENTO JR.; YONEYANA, 2004).

3.3 Perceptron múltiplas camadas (MLP)

Mais tarde, com os estudos de John Hopfield, em 1982, que estabeleceu uma descrição da rede neural de uma lesma, os pesquisadores voltaram a se interessar pelas pesquisas que compreendem as redes neurais (LUDWIG JR.; MONTGOMERY, 2007; NASCIMENTO Jr.; YONEYANA, 2004). Nesse contexto, Hopfield, em sua descoberta, propôs uma rede neural mais robusta, com algumas camadas adicionais entre a de saída e a de entrada, denominadas camadas ocultas. Com a criação dessa técnica, eliminou-se completamente o antigo problema do OU exclusivo (XOR).

Essa técnica, denominada *Multilayer Perceptron*, prevê uma camada de entrada com tantos neurônios quantos forem os pesos de entrada, uma ou de várias camadas ocultas e uma camada de saída, com o número de neurônios igual ao número de pesos de saída (VALENÇA, 2009; GONÇALVES, 2005; NASCIMENTO Jr.; YONEYANA, 2004).

As redes neurais artificiais de múltiplas camadas são extremamente precisas em suas respostas. A grande vantagem em utilizá-las consiste no fato de que uma rede neural imita o cérebro humano, isto é, transcreve os atos de aprendizagem com erros e acertos em suas tentativas (ANITHA et al., 2012). Porém, as redes neurais *perceptron*, na prática, não conseguem fornecer soluções para problemas muito complexos sozinhas. Então, o algoritmo de retropropagação do erro com aprendizagem supervisionada tem o papel de ajudá-las nessa tarefa (LUDWIG JR.; MONTGOMERY, 2007; SCHALKOFF, 2001).

3.4 Retropropagação do erro com aprendizagem supervisionada

Há vários algoritmos que produzem o aprendizado para as redes neurais artificiais, como a retropropagação do erro com aprendizagem não supervisionada (LUDWIG JR., MON-

TGOMERY, 2007) e o desenvolvimento da regra delta (VALENÇA, 2009). O algoritmo utilizado na rede neural artificial é o de retropropagação de erro com aprendizagem supervisionada, que funciona com um vetor de valores de entrada que inicia na primeira camada de neurônios. Dentro de cada neurônio, são realizados os cálculos de ativação e de transferência, repassando o resultado à entrada do neurônio na camada seguinte, até chegar à última camada de neurônios. Nesta camada, é calculada a diferença entre a saída desejada e a saída encontrada pela rede neural.

As técnicas supervisionadas, como o cálculo do erro do neurônio de saída e a taxa de aprendizado, dão aplicadas de trás para frente, do último neurônio para o primeiro, levando as redes neurais artificiais ao grande propósito de ajustá-las em sua precisão. Assim, se o erro não for aceitável, colocam-se novos pesos para as *sinapses* e o *bias* do neurônio.

Em uma rede neural de múltiplas camadas, utiliza-se um valor de erro para definir uma parada aceitável. A rede neural artificial procurará aproximar-se desse valor ou até mesmo igualar-se a ele em alguns casos. O erro não pode ser inferior ao escolhido, pois, se isso ocorrer, haverá uma parada forçada da rede, e novos valores terão de ser escolhidos.

Para medir o desempenho do treinamento, diminui-se o valor desejado do valor calculado do neurônio no treinamento, na validação e na verificação. A diferença indica o quanto a rede aprendeu. Esse procedimento chama-se treinamento por supervisão (GUIMARÃES, 2008).

3.2. Treinamento, validação e verificação

A confiabilidade dos resultados obtidos pela rede neural depende da ordenação correta dos cálculos utilizados para o treinamento da rede e do ajuste de cinco parâmetros modificáveis, que são: a taxa de aprendizado, o valor do erro desejado, o AIC, o R^2 ajustado, o valor do ciclo e a escolha da melhor topologia da rede neural. Quanto menor for o erro destes parâmetros, mais próximo da correção as *sinapses* e o *bias* estarão. Há três etapas para que a rede neural possa ser considerada validada: o treinamento/aprendizado, a validação cruzada e a verificação. O treinamento/aprendizado de uma rede neural leva em consideração uma regra formal, a utilização de informações do supervisor da rede, que atualiza os parâmetros modificáveis da rede, como o *bias* e as *sinapses*. A participação do supervisor pode ser classificada em métodos de treinamento. Quando essa participação é forte, o supervisor fornece um conjunto de entradas correlacionadas com o conjunto de saída desejada, ou seja, o supervisor apresenta os resultados almejados no vetor de saída desejada, para que se tenha a mesma “correlação” de entradas posteriores, aumentando o aprendizado a cada entrada do treinamento. Este método corresponde ao algoritmo de *backpropagation* (retropropagação). O erro é, então, retropropagado da saída para a entrada, fazendo com que a rede “imite” o supervisor, procurando sempre obter um erro menor (NASCIMENTO JR.; YONEYANA, 2004; SCHALKOFF, 2001). No treinamento, os valores recebidos no neurônio, para melhor refinamento da entrada, foram normalizados com a fórmula linear (VALENÇA, 2009).

A validação cruzada, por sua vez, é um ponto muito importante para o treinamento. Com ela, pode-se ter uma certeza de parada, ou seja, uma forma de descobrir se a rede neural não está super treinada somente para a entrada correspondente, localizando exatamente quando a rede neural está extremamente precisa e treinada. Para isso, comparam-se o erro médio quadrado (LUDWING JR.; MONTGOMERY, 2007; SCHALKOFF, 2001; GUIMARÃES, 2008), o erro médio global (GUIMARÃES, 2008; VALENÇA, 2009) e o erro padrão de predição (VALENÇA, 2009) do treinamento e da validação, procurando obter uma variação muito baixa entre as duas etapas do treinamento.

Para os erros, foi utilizado um critério de parada buscando sempre um valor abaixo do erro aceitável de resultado entre 0,08 a 0,05 por ciclo no erro médio global, um valor entre 0,008

a 0,005 por ciclo no erro médio quadrático e um valor entre 0,008 a 0,005 por ciclo no erro padrão de predição do treinamento (VALENÇA, 2009).

O último quesito a ser avaliado é a verificação, preceito em que se realiza uma comparação entre o erro médio quadrado da verificação e o erro médio quadrado da validação cruzada. A partir do momento em que o erro da validação cruzada, ocasionado pela entrada desejada menos a obtida pela rede neural, afastando-se cada vez mais do erro da verificação, ocasionado pelo valor desejado menos o obtido, considera-se a rede neural treinada até aquele exato momento (VALENÇA, 2009). Nesta etapa, também são feitas comparações entre outras análises ou outros tipos de redes diferentes, para uma confirmação do treinamento.

3.3. Topologias da rede neural

A topologia da rede neural consiste no número de neurônios nas camadas de entrada, nas camadas escondidas e na camada de saída. A topologia que se enquadra melhor ao problema somente é verificada no treinamento, com o método de tentativa e erro. A seguir, apresenta-se como exemplo, na Figura 2, uma rede com a topologia 2-3-3.

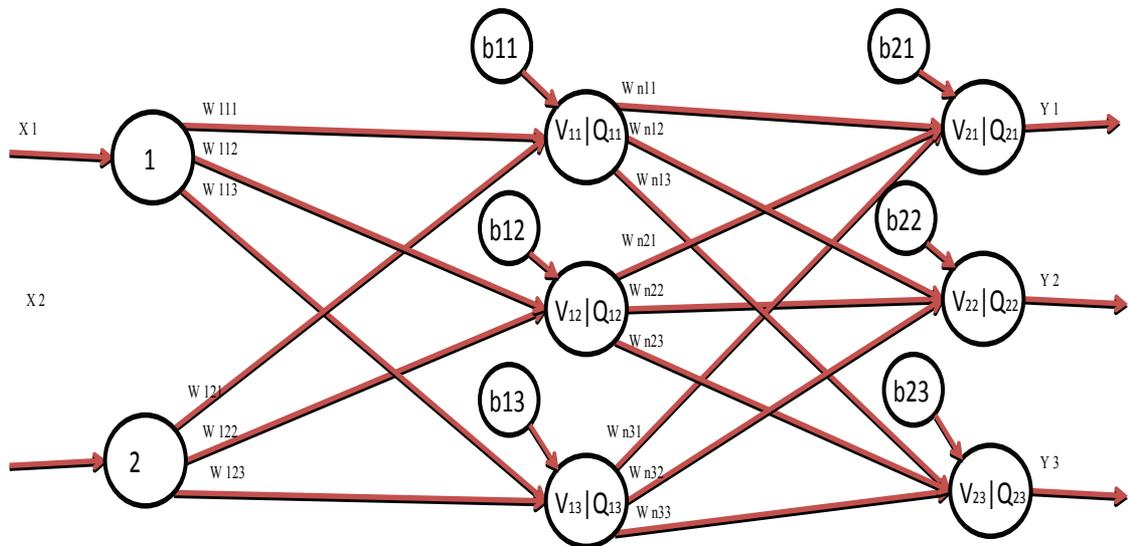


Figura 2: Topologia Rede Neural Artificial MLP

Fonte: os autores.

4 SISTEMA DE APOIO À DECISÃO

Neste trabalho, foram construídos dois modelos de apoio à decisão, voltados, principalmente, aos investidores com pouco ou nenhum conhecimento sobre técnicas de análise do mercado acionário. A rede neural presente no sistema foi desenvolvida na linguagem Java. O sistema

utiliza uma rede neural *perceptron* de múltiplas camadas para suas tomadas de decisão. Dessa forma, o sistema assegura maior garantia de sua saída, e a rede aprende com o comportamento passado do Ibovespa e faz um prognóstico de comportamento futuro deste ativo, fortalecendo a decisão do usuário e melhorando a confiabilidade da análise.

O segundo modelo desenvolvido é o modelo de ordem ARMA(p, q), que foi estabelecido com base no *software* estatístico R. Para chegar às definições das variáveis do modelo ARMA(p, q), foram feitas baterias de testes, ajustando os parâmetros modificáveis (p, q) do modelo.

Para atingir os fins propostos, é preciso entender que o principal objetivo da construção dos dois métodos de previsão descritos neste trabalho consiste em construir um rastreador de tendências que apoie a tomada de decisões de um investidor na bolsa de valores. Dessa forma, procura-se auxiliar a localização do momento certo de compra e venda da ação analisada, com o intuito de garantir que a solução encontrada seja tão generalizável quanto possível. Para isso, compara-se, a seguir, a previsibilidade dos dois métodos.

4.1. Manipulação dos dados

Para as entradas da rede neural, as cotações foram divididas em quatro categorias: saída desejada, treinamento, validação cruzada e verificação. Já o modelo ARMA não é composto pelas etapas de treinamento e validação cruzada, motivo pelo qual não foram divididas as entradas, utilizando-se 3332 cotações corridas da volatilidade para determinar a previsão com base no dia anterior.

O treinamento da rede neural satisfaz 50% das cotações obtidas de um total de 3332 cotações de 02/01/1998 até 16/06/2011. A validação cruzada correspondeu a 25% do total, e a verificação correspondeu aos 25% restantes. A saída desejada foi desenvolvida a partir da previsão do dia anterior, como mostra a equação 1, citada anteriormente. Geralmente, as séries temporais não são de raiz unitárias, por isso os dados foram estacionados com base na diferenciação entre a série do *log* retorno, aplicando a fórmula para estimar Y com os valores defasados de X menos X , multiplicados por *log*, conforme indica a equação 5, a seguir:

$$Y_i = \log \sum_{i=1}^p \alpha_i X_{i-1} - X_i + \epsilon_i \quad (5)$$

Fonte: Greene (2000).

Para o treinamento da rede neural, recebeu-se, no primeiro neurônio (que fica na camada de entrada, como três neurônios), um vetor de cotações diárias da volatilidade do Ibovespa analisado. O segundo neurônio recebeu o tempo correspondente das cotações. E o terceiro neurônio recebeu um vetor de número um para melhor obtenção de um erro mais baixo.

A análise da validação cruzada utiliza a fórmula do erro médio quadrático e do R^2 ajustado como um critério de parada aceitável, com entradas distintas do treinamento. Com isso, pode-se concluir que as sinapses e o *bias* correspondem a uma correção entre a média de erro e os valores aceitáveis obtidos na rede neural, assim como à ligação entre um nó e outro dos demais métodos utilizados com a saída desejada.

O último estágio a ser concluído consiste na verificação. Nesta etapa, é possível constatar que os métodos não perdem sua precisão na troca de entradas, ou seja, que seu erro médio quadrado permanece dentro os valores aceitáveis.

Neste estágio, também há a primeira comparação da previsão dos métodos, utilizando

o critério do R^2 ajustado, de modo que, quanto menor for o erro médio quadrático do método, melhor será seu poder computacional e melhor treinado estará o método. Além disso, quanto mais próximo do valor 1 estiver o R^2 ajustado (equação 6), maior a garantia de previsibilidade correta, o que indica que, quando maior o R^2 ajustado, maior o poder de explicação de X em Y .

$$\bar{R}^2 = 1 - \frac{n-1}{n-k} (\bar{e}^2 - R^2) \quad (6)$$

Fonte: Alves (2007).

Onde: n é o número de observações; K é o número de parâmetros; e R^2 ajustado é o coeficiente de determinação ajustado, ou seja, o coeficiente mais preciso.

O segundo critério a ser avaliado nos modelos e a comparação do AIC de cada método, onde: n é o número de observações; K é o número de parâmetros; R^2 ajustado é o coeficiente de determinação; e $e'e$ é a soma dos quadrados dos resíduos, como mostra a equação 7, a seguir. Conforme Greene (2000), uma medida alternativa de ajuste sugerida é o critério de informação de Akaike, que é reportado no *software* R, segundo a equação 5.

$$AIC = \log\left(\frac{e'e}{n}\right) + \frac{2K}{n} \quad (7)$$

Fonte: Alves (2007).

5 ANÁLISE DOS RESULTADOS

Nesta seção, foram analisados, por meio de diversos testes, os resultados obtidos quanto à modificação das entradas, aos valores, à saída desejada e aos métodos utilizados no treinamento e na validação da rede neural e do método ARMA(p, q). Com a finalidade de obter uma baixa variação entre a saída desejada e a saída calculada pela rede neural e pelo ARMA(p, q), procuraram-se um menor erro médio quadrático, um menor erro médio global, um menor erro padrão na rede neural e um valor do R^2 ajustado mais próximo de 1 no modelo ARMA(p, q).

5.1. Análise dos resultados do modelo ARMA

Como já mencionado, para que a série seja estacionária, procedeu-se à diferenciação dos valores com o *log* retorno, como mostra a equação 5. Esta técnica, referida anteriormente, por meio do teste de Durbin e Watson, não detectou autocorrelação dos resíduos. Porém, a Figura 3, a seguir, que apresenta os resíduos da regressão deste primeiro modelo, sugere que ele apresenta heterocedasticidade condicional autorregressiva (ALVES, 2007). Ressalta-se, ainda, que os índices de retornos se desenvolvem ao longo do tempo, ao redor de um valor imutável zero, e apresentam uma forma de balanceamento constante, podendo-se concluir que os retornos dos índices são estacionários, como mostra a Figura 3 (FARIAS, 2008).

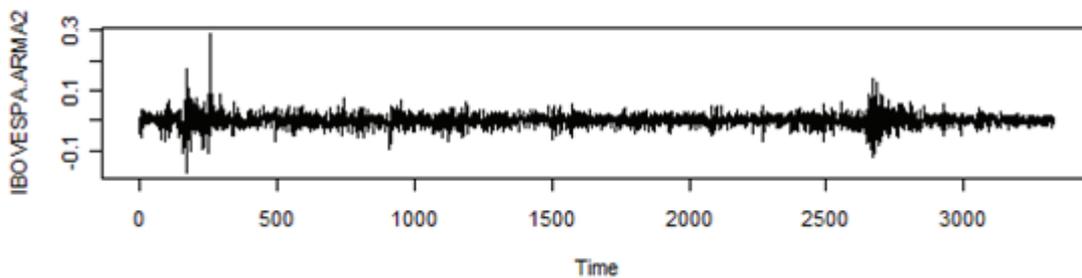


Figura 3: Série log retorno diária (Ibovepa) com Raiz Unitária

Fonte: os autores.

Para o critério de seleção dos modelos AR, MA e ARMA, foi utilizado o teste do AIC dos resultados da regressão, como mostra a Tabela 1, exposta a seguir. Nesta seleção, foi considerado o menor valor para a melhor significância dos parâmetros, comparando-se todos os modelos autorregressivos.

Tabela 1: Critério de seleção da distribuição dos modelos

Modelos	Valor de AIC
AR(p,0)	-15968.13
MA(0,q)	-15959.98
ARMA(p, q)	-15969.47

Fonte: os autores

A Figura 4, a seguir, demonstra o teste gráfico ACF e o teste gráfico ACF Parcial no *software* R. Considerando os *log* retorno sem *p* como as variáveis de entrada em ARMA(*p*, -), os dois testes foram realizados até que os valores da série não se encontrassem com autocorrelação parcial. Isso indica que o teste ACF dos modelos corrobora a correlação de *X* em *Y* até certo valor testado (ALVES, 2007; FARIAS, 2008). Então, na análise gráfica dos testes, pode-se observar que, levando-se em consideração os valores entre o *lag* zero e o *lag* dez, somente o quarto *lag* transpõe a linha pontilhada, demonstrando que apenas até o quarto *lag* a série temporal não dispõe de autocorrelação parcial e que o valor de *p* como as variáveis de entrada em ARMA(4,-) é quatro (Figura 4).

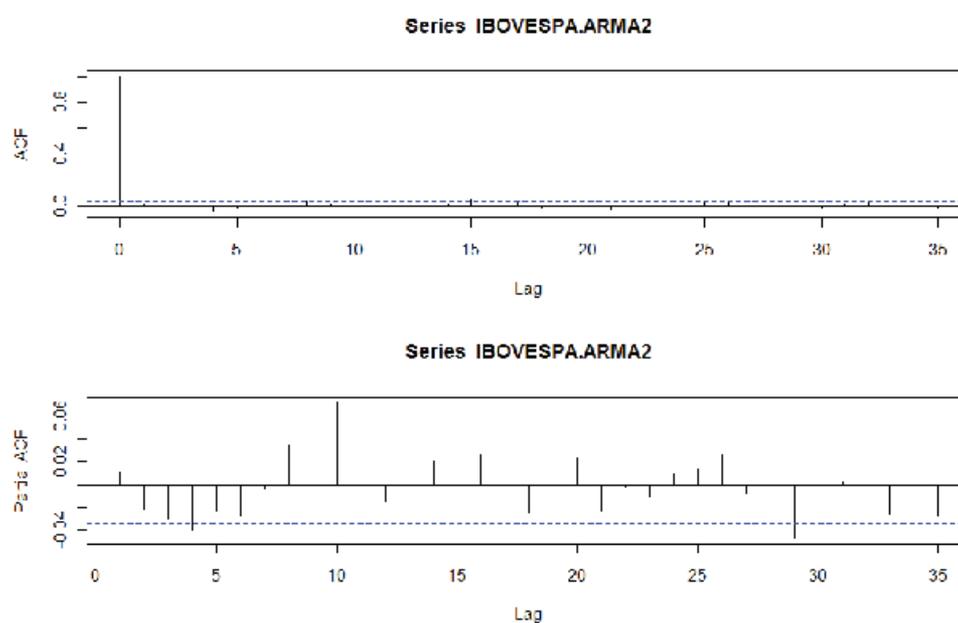


Figura 4: Teste ACF da função de autocorrelação dos log retornos.

Fonte: os autores.

Após o teste de autocorrelação parcial dos *log* retornos, foi feito o primeiro critério de seleção da defasagem da série, considerando-se a letra *q* como as variáveis de entrada em ARMA(4, *q*). Para este teste, utilizou-se a significância dos parâmetros com os erros padrões robustos, conforme exposto na Tabela 2, a seguir. Na Tabela 2, pode-se observar que todos os valores de $Pr(>|t|)$ importantes foram significativos, ou seja, menores do que 0,05, o que indica que são todos válidos ao valor de cinco defasagens da série representada pela letra *q* como as variáveis de entrada em ARMA(4, 5).

Tabela 2: Parâmetros e significância dos coeficientes

	Estimate	Std. Error	t value	Pr(> t)
ar1	0,2453	0,0549	4,468	7,91e-06
ar2	0,9429	0,0468	20,140	< 2e-16
ar3	0,4702	0,0385	12,204	< 2e-16
ar4	0,7913	0,0578	-13,687	< 2e-16
ma1	0,2375	0,0567	-4,186	2.84e-05
ma2	0,9780	0,0452	-21,621	< 2e-16
ma3	0,5100	0,0458	-11,119	< 2e-16
ma4	0,8001	0,0553	14,444	< 2e-16
ma5	0,0648	0,0184	3,506	0,0004
intercept	0,0005	0,0003	1,340	0,1802

Mas, somente o teste de significância dos parâmetros não demonstra totalmente a dependência dos *log* retornos. Por isso, utilizou-se o teste Box-Ljung para o segundo critério de seleção da defasagem da série, representada pela letra *q* como as variáveis de entrada em ARMA(4, *q*), para a confirmação da indicação da não dependência do *log* retorno (KUMAR et al., 2010). Neste teste de Box-Ljung, considerou-se a presença de um valor maior que 0,5 de *p-value* na rejeição de H_0 como a não correlação nas estatísticas sobre os resíduos padronizados. Como demonstração à rejeição de H_0 , a saída do teste para *p-value* foi 0,8446, ou seja, maior que 0,5, validando e confirmando o valor de cinco defasagens da série representada pela letra *q* como as variáveis de entrada em ARMA(4, 5). Com estes critérios, pode-se perceber que o modelo ARMA(4,5) obteve um melhor desempenho dentre os vários testados, já que todos os parâmetros foram significativos, com menor AIC e *p-value* maior do que 0,5 nos resíduos padronizados normais.

5.2. Análise dos resultados da rede neural

Para que uma rede neural com o aprendizado supervisionado possa ser considerada treinada, como mencionado anteriormente, o supervisor deve modificar os parâmetros modificáveis da rede neural seguindo seis grandes etapas: a modificação da taxa de aprendizado; a escolha da topologia da rede neural; a escolha do valor de ciclo; a escolha do erro desejado; o desenvolvimento do AIC; e o R^2 ajustado da rede que mais se destacou segundo critérios já mencionados.

Na etapa da taxa de aprendizagem por interação, o valor foi escolhido a partir dos testes feitos na fase do treinamento da rede neural. Neste processo, ocorreu uma comparação entre os valores escolhidos com os resultados dos erros obtidos. Para a escolha do valor da taxa de aprendizagem, os critérios utilizados foram o menor erro por interação, o menor erro global, o menor erro médio quadrático e o menor tempo de execução.

Na análise dos resultados, a partir dos testes realizados, concluiu-se que a taxa de aprendizado por interação 20 obteve o melhor custo-benefício dentre as testadas, com o menor erro médio global no ciclo 10000, o menor erro por interação na posição 30 do vetor de saída do erro e o segundo menor tempo de execução total. Para melhor esclarecimento, a posição 30 do vetor de saída do erro foi escolhida aleatoriamente. Os resultados foram apresentados na Figura 4, que mostra a taxa em que ocorreu o menor erro médio global, o menor erro médio quadrático e o menor erro na interação escolhida. Ao comparar a saída calculada pela rede neural (vermelho fraco) e a saída desejada (vermelho forte), pode-se perceber que há pouca variação entre as saídas, cumprindo um dos objetivos deste artigo, que consiste em obter um erro baixo e um treinamento satisfatório.

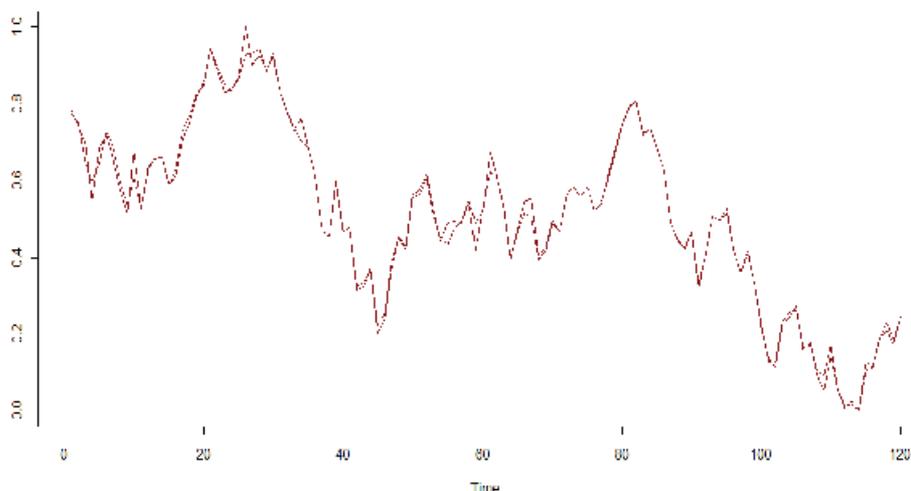


Figura 4: Gráfico da melhor taxa de aprendizado

Fonte: os autores.

Na segunda grande etapa, a escolha da topologia, a rede neural conta com uma topologia 3-8-1, pois foi a melhor resposta encontrada no treinamento dentre as 13 diferentes topologias testadas, variando de dois a 14 neurônios na camada oculta. A melhor situação correspondeu a três neurônios na camada de entrada, oito neurônios na camada oculta e um neurônio na camada de saída. Nesta etapa, também se observou quantos ciclos foram necessários para que a predição ocorresse em um erro médio quadrático aceitável, como mostra a Tabela 3, exposta a seguir, com os resultados mais relevantes do erro médio quadrático na saída da rede neural.

Tabela 3: Tabela da Topologia da MLP

	3-4-1	3-6-1	3-8-1	3-10-1
Ciclo 100	0,00365	0,00301	0,00253	0,00408
Ciclo 500	0,00351	0,00309	0,00257	0,00416
Ciclo 1000	0,00342	0,00338	0,00296	0,00428
Ciclo 8000	0,00373	0,00348	0,00352	0,00463
Ciclo 10000	0,00352	0,00365	0,00351	0,00456

Fonte: os autores.

Na Tabela 3, pode-se observar que o melhor resultado do erro médio quadrático foi obtido no ciclo 500, com oito neurônios na camada oculta. Nesta etapa, os menores erros entre os ciclos e as topologias testadas comprovam um dos objetivos do treinamento, mapear as oscilações da Bovespa.

Na quarta grande etapa do treinamento e da validação cruzada, o valor do erro desejado por interação foi escolhido a partir da aproximação do erro obtido com o erro desejado. O critério utilizado foi o menor valor do erro calculado sem danificar os resultados. A seguir, na Tabela 4, apresentam-se os resultados desta comparação.

Tabela 4: Erro Desejado

	EMQ	EMG	EP
Erro Desejado 0,70	0,00431	0,05195	0,07285
Erro Desejado 0,50	0,00257	0,04891	0,0682

	EMQ	EMG	EP
Erro Desejado 0,30	0,00675	0,06347	0,07931
Erro Desejado 0,09	0,01979	0,06899	0,18659

EMQ = Erro Médio Quadrático

EMG = Erro Médio Global

EP = Erro Padrão de Previsão

Fonte: os autores.

Então, concluiu-se que o valor do erro desejado 0,50 teve o melhor custo-benefício dentre os valores comparados, com a melhor aproximação do erro desejado e o menor erro propiciado pelo esforço da rede neural para alcançá-lo. Levando-se em consideração o treinamento inteiro, a rede neural com 500 ciclos, topologia 3-8-1, taxa de aprendizagem 45 e erro desejado de 0,5 foi considerada a mais satisfatória no conjunto do treinamento e da validação cruzada.

Para validar os resultados do treinamento da rede neural, na Figura 5, a seguir, apresenta-se um gráfico para melhor visualização dos resultados da variação dos parâmetros modificáveis escolhidos. Nesta etapa de validação da rede neural, foram utilizadas entradas diferentes daquelas do treinamento. Entretanto, os resultados da etapa de validação reafirmaram os resultados obtidos na etapa do treinamento, com menor variação em relação à saída desejada e o menor erro médio global com uma rede neural de 500 ciclos, topologia 3-8-1, taxa de aprendizagem 45 e erro desejado de 0,5, conforme mostra a Figura 5, exposta a seguir.

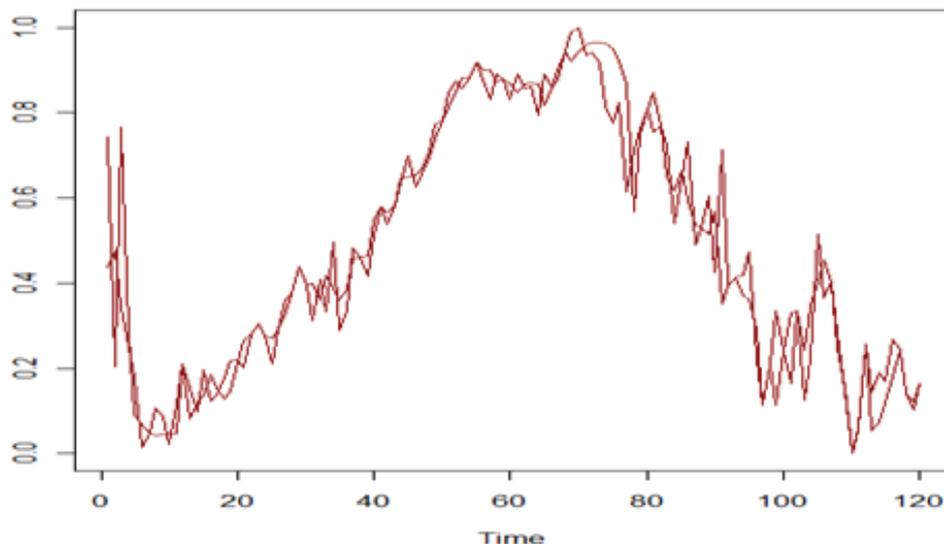


Figura 5: Gráfico da validação da rede neural

6 ANÁLISE DE PREDIÇÃO

Nesta seção, apresenta-se a comparação dos dois modelos que se destacaram dentre todos os testados na análise dos resultados, levando-se em consideração três critérios: o menor erro médio quadrático; o menor erro médio global; e o maior R^2 ajustado obtido. Portanto, será comparado o modelo ARMA(4, 5), que apresentou um melhor desempenho do que os modelos lineares testados, como demonstrado no item 5.1, e a rede neural de 500 ciclos, com a topologia

3-8-1, a taxa de aprendizagem 45 e o erro desejado de 0,5, que apresentou uma melhor previsão que suas modificações testadas no item 5.2.

6.1 Análise dos modelos ARMA e rede neural

Como mencionado, foram comparadas duas ferramentas, uma não heurística e uma que utiliza o treinamento e a aprendizagem supervisionada, com o intuito de observar qual método tem a melhor disponibilidade de previsão, sendo, portanto, mais confiável. Contudo, o objetivo principal deste item também é validar os dois métodos que não poderão se afastar das implicações obtidas na análise dos resultados. Neste caso, foi analisada a previsão do retorno do volume do Ibovespa de uma série temporal distinta da empregada no treinamento e na validação da rede neural, com os critérios de menor EMG e EMQ e maior R^2 ajustado, conforme demonstra a Tabela 4, a seguir.

Tabela 4: Comparação do EMQ e EMG dos métodos

	ARMA	Rede Neural
EMG	0,07621	0,03792
EMQ	0,00894	0,00257
R^2 ajustado	0,03687	0,06285

Na análise da Tabela 4, pode-se perceber que a rede neural obteve, tendo em vista os três critérios utilizados na comparação, melhores resultados na previsão do retorno do volume do Ibovespa do que o modelo ARMA(4, 5). No entanto, os seis resultados das três equações foram bastante satisfatórios se levarmos em consideração o desempenho das duas ferramentas. Assim, conclui-se que os dois modelos obtiveram resultados aceitáveis e, portanto, foram validados.

7 CONSIDERAÇÕES FINAIS

Este trabalho obteve resultados positivos na previsão de oscilações do retorno do volume do índice BOVESPA. Na fase de treinamento e nas demais fases da rede neural, pode-se perceber que o desempenho do modelo ARMA foi também muito satisfatório. Levando-se em consideração a complexidade dos dois modelos analisados e os resultados positivos obtidos, pode-se concluir que os dois modelos mostram um poder computacional satisfatório.

Três critérios de avaliação foram considerados na comparação dos dois modelos. Um destes critérios levou em consideração o menor valor do erro médio quadrático. Nesta comparação, pode-se perceber que o EMQ da rede neural obteve um valor relativamente mais baixo do que o do modelo ARMA, fazendo com que o primeiro critério de validação seja favorável à rede neural. O segundo critério de avaliação foi o menor erro médio global dos modelos, em que é possível perceber, conforme Tabela 4, que a rede neural também obteve um melhor desempenho que o modelo ARMA. No terceiro critério de avaliação, foi utilizado o R^2 ajustado, com resultado também favorável à rede neural, o que a indica como a melhor ferramenta utilizada para a previsão do retorno do volume do Ibovespa no período analisado. No entanto, como já mencionado, o modelo ARMA não obteve resultados insatisfatórios. Em vista disso e levando-se em consideração a complexidade dos dois algoritmos, é lícito afirmar compreender que o modelo menos complexo pode apresentar um desempenho mais satisfatório em alguns casos.

Quanto às limitações deste estudo, pode-se apontar a não possibilidade de descrição do teste de White (para detectar heterocedasticidade irrestrita dos modelos) e do teste ARCH-LM (para detectar heterocedasticidade condicional), realizados para execução desta pesquisa,

devido à necessidade de respeitar a extensão máxima aceitável para um artigo. Como trabalhos futuros, espera-se desenvolver um método com mais critérios de avaliação para os modelos, utilizando, também, mais variáveis de entrada, a fim de proporcionar resultados mais concretos. Sugere-se, também, comparar o retorno do volume do Ibovespa com outros modelos da família GARCH, como o I-GARCH e o T-GARCH e E-GARCH.

REFERÊNCIAS

- ALVES, J. S. Análise comparativa e teste empírico da validade dos modelos CAPM tradicional e condicional: o caso das ações da Petrobrás. **Rev. Ciênc. Admin., Fortaleza, v. 13, n. 1, p. 147-157, 2007.**
- BANIFINVEST. "BI Desktop Streamer5". São Paulo. 1999. **Sistema para Análise Técnica**, Disponível em (<https://www.banifinvest.com.br/tr/bi/produtos/bidsbeta/index.jsp>). Acesso em: 10 Dez. 2009.
- BOX, G. E. P.; JEANKINS, G. M. **Time Séries Analysis: Forecasting and Control**. San Francisco: Holden-day. 1976.
- BOX, G. E. P.; JEANKINS, G. M. **Time Séries Analysis: Forecasting and Control**. Englewood Cliffs NJ: Prentice-Hall, New York. 1994.
- CAMPOS, R. J. **Previsão de séries temporais com aplicações a séries de consumo de energia elétrica**. Dissertação de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais. Belo Horizonte - MG. 2008.
- CERETTA, Paulo S; RIGHI, Marcelo B; SCHLENDER, Sergio G. PREVISÃO DO PREÇO DA SOJA: uma comparação entre os modelos ARIMA e redes neurais artificiais. **Revista Informações Econômicas, Instituto de Economia Agrícola - Centro de Comunicação e Transferência do Conhecimento**. São Paulo. v.40, n.9. 2010.
- COELHO, L. dos S.; SANTOS, A. A.P.; COSTA J.R. N. C. A; Podemos prever a taxa de cambio brasileira? Evidencias empírica utilizando inteligência computacional e modelos econométricos. **Revista Gestão de Produção, São Carlos, v.15, n. 3. 2008.**
- CUNHA, G. B.; MOURA, J. F. L.; NAVES, E. L. M.; ANDRADE, A. O.; PEREIRA, A. A.; MILAGRE, Selma T.. A utilização de uma Rede Neural Artificial para Previsão de Incidência da Malária no Município de Cantá, Estado de Roraima. **Revista da Sociedade Brasileira de Medicina Tropical**. Uberaba-MG. set./out. . 2010.
- FARIAS, H. P.; SÁFADI, T.. Causalidade Entre as Principais Bolsas de Valores do Mundo. **RAM - Revista de Administração Mackenzie**, V. 11, N. 2, São Paulo – SP. 2010.
- GALÃO, O. F.; BORSATO, D.; PINTO, J. P.; VISENTAIRNE, J. V.; GREENE, W. **Econometric analysis**. New Jersey: Prentice Hall, 2000.
- GALVÍNCIO, J.D.; SOUSA, F. de A. S; SOUSA I. F.. **Uso de Modelos Auto-Regressivos e Medias Moveis para Geração de Vazões Médias Mensais Na Bacia Hidrográfica do Alto São Francisco**. XII Congresso Brasileiro de Meteorologia, Foz de Iguaçu-PR, 2002
- GRAHAM, B.; DODD, D.. **Security Analysis**. McGraw-Hill. 2004.
- HAYKIN, S. **Redes Neurais: Princípios e Prática**. 3ª Edição, São Paulo. Bookman. 2001.
- HASSAN, M.D. R.; BAIKUNTH, Nath. **Stock Market Forecasting Using Hidden Markov Model: A New Approach**. *Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA'05)*. IEEE Computer Society. 2005.
- IQUIAPAZA, R. A.; BRESSAN, A. A.; AMARAL, H. F.. Previsão Não-linear de Retornos na BOVESPA: Volume Negociado em um Modelo Auto-Regressivo de Transição Suave. **ANPAD - RAC, Curitiba, v. 14, n. 1, art. 8, pp. 149-171, Jan./Fev. 2010**

- JUDD, J. S. **Neural Network design and Complexity of Learning**. MIT Press, USA, 1990.
- KIM, K.; LEE, W.B. *Stock market prediction using artificial neural networks with optimal feature transformation*. **Journal Neural Computing & Applications**. Springer-Verlag, 13: 255-260, 2004.
- KIMOTO, T, K; ASAKAWA, M; YODA, M. Takeoka. **Stock market prediction system with modular neural networks**. In Proceedings of the International Joint Conference on Neural Networks, volume 1, pages 1–6, 1990.
- KUMAR, U.; RIDDER Koen. *GARCH modelling in association with FFT e ARIMA to forecast ozone episodes*. **Journal Institute for Technological Research, Elsevier**. Belgica. 2010.
- LAWRENCE, R. **Forecasting stock prices using neural networks**. Department of Computer Science, University of Manitoba, December 1997.
- LUDWING J.R., OSWALDO; MONTGOMERY, Eduard. **Redes Neurais: Fundamentos e Aplicações com Programa em C**. Rio de Janeiro: Ciência Moderna, 2007.
- MAIA, M.; ALVES, T. S; **Métodos Bayesianos de estimação e comparação de modelos na família GARCH**; Trabalho de conclusão de curso - Estatística – UFPR; Curitiba, 2008.
- NASCIMENTO J.R., L Cairo; YONEYANA, Takashy. **Inteligência Artificial em Controle de Automação**. São Paulo: FAPESP, 2004.
- NELSON, D. B. *Conditional Heteroskedasticity in Asset Returns: A New Approach*. **Journal Econometrica**, v. 59, p.247-370, 1991.
- NORONHA, M.. **Análise Técnica: Teorias Ferramentas Estratégicas**. 7ª Edição. Rio de Janeiro, EDITEC, 2009.
- OLIVEIRA, A. C. S.; SOUZA, A. A.; LACERDA, W. S.; *Aplicação de Redes Neurais Artificiais na Previsão da Produção de Alcool*. **Revista Ciência e Agrotecnologia**, Lavras-MG. V. 34, n. 2, p. 279-284, mar./abr. 2010.
- PANIZZI, M. C. C. *Artificial Neural Network in the Classification and Identification of Soybean Cultivars by Planting Region*. **Journal of the Brazilian Chemical Society**. Vol. 22.No. 1, 142-147, 2011.
- PHUA, P.K.H.; MING, D.; LIN, W. **Neural Network With Genetic Algorithms For Stocks Prediction**. *Fifth Conference of the Association of Asian-Pacific Operations Research Societies*, 5th - 7th July, Singapore, 2000.
- RABONI, P. L. **Anomalias no Mercado de Capitais Brasileiro: Ensaios com testes empíricos na BOVESPA**. Tese de Doutorado em Administração de Empresas da PUC-Rio. Rio de Janeiro. 2005.
- ROBERT E. D.; MAGEE, J.W.H.C. *Technical Analysis of Stock Trends*. **American Management Association**, 2007.
- SANTOS, Douglas Gomes. **Estimação de Volatilidade em Séries Financeiras: Modelos Aditivos Semi Paramétricos e GARCH**. Dissertação de Mestrado em Economia – UFRGS, Porto Alegre. 2008.
- SCHALKOFF, R. J. **Artificial neural networks**. SAGE Publication - Progress in Physical Geography, New York: McGraw-Hill, 2001.
- SMITH, K. A.; GUPTA, J. N. D. *Neural networks in business: techniques and applications for the operations researcher*. Volume 27, **Oxford: Computers & Operations Research**, Elsevier, 2000.
- SOBREIRO, V. A.; ARAÚJO, P. H. S. L.; NEGANO, M. S. *Precificação do Etanol Utilizando Técnicas de Redes Neurais Artificiais*. **Revista de Administração**. São Paulo. v.44, n.1. Nicolau Reinhard. 2009.
- STEINER, M.; WITTKEMPER, Hans-Georg. **Neural networks as an alternative stock market model**. In *Neural Networks in the Capital Markets*, chapter 9, pages 137–148. John Wiley and Sons, 1995.
- VALENÇA, M. **Fundamento das Redes Neurais: Exemplos em Java**. 2ª Edição. Olinda: Livro Rápido – Elógica, 2009.