

REDES NEURAI APLICADAS À PREVISÃO DE SÉRIES TEMPORAIS

BATTAGELLO, Vinícius Antonio*, CORDEIRO, Marciso Xavier, DAL PINO Jr, Arnaldo
Divisão de Ens. Fundamental - Instituto Tecnológico de Aeronáutica – ITA
Pça. Mal. Eduardo Gomes, 50 – Vila das Acácias – CEP: 12228-900
São José dos Campos – SP – Brasil – battagello@gmail.com, marcisox@terra.com.br, dalpino@ita.br

Resumo

As Redes Neurais Artificiais (RNA) vêm sendo utilizadas com frequência na construção de modelos para solucionar problemas relacionados à previsão de séries temporais em diferentes áreas, para que seja possível minimizar gastos, aumentar a lucratividade e solucionar diversos problemas das atividades humanas. Este artigo objetiva o uso da capacidade das RNA de aprendizagem de padrões não-lineares para a previsão do Índice da Bolsa de Valores de São Paulo e do Índice Pluviométrico Anual da cidade de Fortaleza, envolvendo diferentes critérios de treinamento e de arquitetura das redes utilizadas.

Abstract

Artificial Neural Networks (ANN) have been utilized frequently in the construction of models to solve problems related to the time series forecasting in different areas, so that it is possible to reduce expenses, raise the profits and solve many problems of the human activities. The current paper objectives the use of the RNA ability to learn nonlinear patterns to forecast the São Paulo Stock Exchange Ibovespa Index and the Fortaleza Annual Raining Index, involving different training and configuration criteria of the used networks.

1. INTRODUÇÃO

As Redes Neurais Artificiais (RNA) constituem uma ferramenta de grande importância na atualidade, por sua capacidade de “aprender” padrões através de treinamento, o que torna seu uso importante no desenvolvimento da Inteligência Artificial. A principal vantagem de uma rede neural é a sua capacidade de aproximar relações funcionais, particularmente quando as relações não são bem definidas e/ou são não-lineares, o que torna difícil a utilização de métodos convencionais para a tentativa de prever as variações futuras dessas relações [1].

Por sua vez, séries temporais, como o próprio nome diz, nada mais são do que uma seqüência de medições relativas a um determinado evento organizadas cronologicamente, que aparentemente não seguem nenhuma lei ou tendência. No entanto, em séries temporais relacionadas a fenômenos naturais, econômicos ou físicos, podemos notar certas características que se repetem após certo período de tempo (sazonalidade) e outras que se mantêm durante o intervalo considerado (tendências), mesmo que não obedeçam a padrões lineares.

Esse trabalho baseou-se na tentativa de fazer com que a rede neural “aprendesse” características temporais para o Índice Pluviométrico Mensal de Chuvas de Fortaleza e o Índice BOVESPA. Através da comparação entre os resultados previstos e os dados reais, foram estudados os diferentes fatores influentes nas dadas previsões, de forma a ser ajustada uma configuração ideal na previsão desses fatores.

2. DESCRIÇÃO DO TRABALHO

O início dos trabalhos consistiu basicamente em realizar o recolhimento e agrupamento dos dados históricos das séries a serem previstas (variação do índice de chuvas mensal da capital cearense e variação do índice de ações da Embraer, da Companhia Vale do Rio Doce e do próprio Índice BOVESPA da Bolsa de Valores de São Paulo) através da Internet [2]. Escolheu-se o Índice Mensal de

* Bolsista do CNPq - Brasil

Chuvas de Fortaleza porque esta série é provavelmente a maior disponível do índice pluviométrico de uma região que sofre frequentemente com secas severas, e por isso tem sido objeto de grande número de estudos na área de prognósticos temporais [3], [4]. Assim, além de ser de grande importância, essa série reúne dados de um período grande, que vai de Janeiro de 1849 a Janeiro de 2003.

A escolha do Índice de Ações da BOVESPA como série principal baseou-se também na sua importância econômica, seja ela para os investidores ou para qualquer empresário, que visa a saber as tendências do Mercado Econômico em um Futuro Próximo, e, por isso, essa série também é alvo de diversos estudos atualmente [5], [6]. A seguir, prosseguiu-se com a confecção e ajuste do algoritmo da Rede Neural, adaptando-o para as diferentes configurações desejadas.

Na primeira parte do trabalho, foi utilizada uma Rede Neural do Tipo *Feedforward*, cuja regra de aprendizado (supervisionado) era o Gradiente Descendente (*Backpropagation Descent*). Composto a rede havia, como regra geral, 1 camada de entrada, 1 de saída e 2 camadas escondidas, sendo que a função de ativação era linear para a camada de saída e tangente-sigmóide para as demais camadas. Ainda nessa primeira parte, a taxa de aprendizado utilizada foi de 0,05, com 300 épocas de treinamento (iterações) para a Rede, em que o Erro Quadrático Médio procurado foi de 10^{-4} .

Para o armazenamento e manipulação dos dados relativos às séries temporais foi utilizado o Microsoft Excel 2003[®], por sua facilidade de operação com tabelas. Na programação das diferentes Redes Neurais, foi utilizado o Matlab[®] 6.5, uma vez que essa linguagem, além de permitir a simplificação do tratamento de operações matemáticas necessárias (manipular matrizes e traçar gráficos), apresenta já implementado um suporte para aplicações de Redes Neurais [7] e permite um intercâmbio fácil de dados com o Microsoft Excel[®], através do plug-in Excel Link[®]. Dessa forma, o foco do trabalho poderia se concentrar na obtenção de resultados através das diversas configurações, sem excessiva preocupação com questões funcionais.

Nessa parte inicial, o trabalho concentrou-se na previsão do Índice Pluviométrico de Fortaleza, do Índice de Fechamento das Ações da Embraer e do Índice BOVESPA, testando as diversas configurações dos dados de entrada. Também se analisou a influência das arquiteturas das redes no resultado final, variando os números de camadas e o número de neurônios por camada, bem como o tamanho da janela de dados apresentados à rede (o tamanho da janela representa os 'n - 1' valores anteriores da série fornecidos à rede para a previsão do enésimo valor).

Na segunda parte do trabalho, o foco foi tentar generalizar os resultados obtidos para a primeira parte, tentando descobrir outros fatores que influíam nos resultados obtidos e analisando os resultados que outras configurações de RNAs forneciam. Assim, além da Rede *FeedForward*, foi feita a tentativa de implantar um modelo de Rede Recorrente de Elman (RRE), na busca de melhores resultados.

Buscou-se essa rede porque a sua particularidade reside no fato de que há conexões que partem da saída de uma unidade em direção a uma outra unidade da mesma camada ou de uma camada anterior a esta. Este tipo de conexão permite a criação de modelos que levam em consideração aspectos temporais e comportamentos dinâmicos, onde a saída de uma unidade depende de seu estado em um tempo anterior. Dessa forma, em busca de melhores resultados, esse foi um outro modelo escolhido para tentativa de implementação.

Além disso, foi feito um estudo sobre a influência do número de épocas de treinamento da rede e da taxa de aprendizado no Erro Quadrático Médio da previsão, bem como a utilização de diferentes configurações de rede, ajustando a que melhor se comportava em cada caso.

Essa parte final concentrou seus testes no índice de ações, observando se os resultados obtidos para o Índice de Ações da Embraer se estendiam para o Índice de Ações da Vale do Rio Doce. Outro foco foi a tentativa de realizar previsões para o Índice de Abertura e para o Índice de Fechamento das Ações da Embraer, seja através de previsões separadas, de previsão conjunta (rede com 2 saídas) ou da previsão da diferença entre o índice de fechamento e de abertura. Essa tentativa se baseou na intenção de permitir, ao investidor, saber qual será a tendência de variação do índice no dia, fornecendo informações para um controle diário sobre os investimentos.

Em todo o trabalho, antes que os dados fossem apresentados à RNA, foi utilizado um processo de normalização [8], no qual o vetor com os dados de entrada foi normalizado ao longo de intervalo [0, 1] através da divisão de cada elemento pelo maior elemento do vetor.

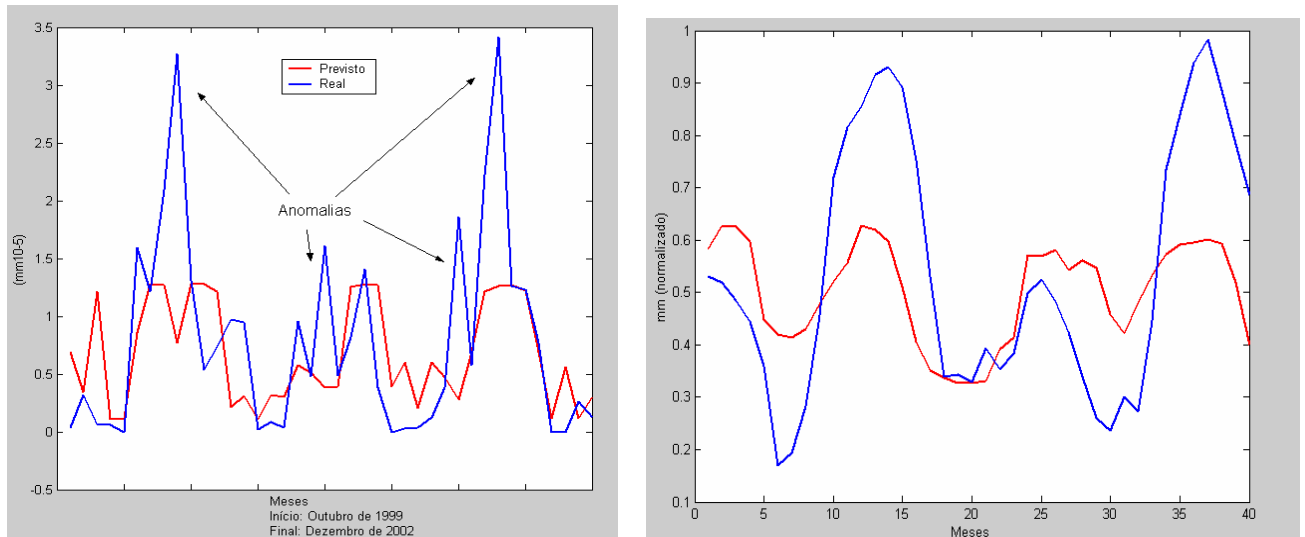
De forma geral, o algoritmo utilizou a seguinte construção: o histórico de dados coletados foi dividido em 2 partes: uma que seria apresentada à rede para treinamento, e a outra que corresponderia à rede prevista. Logo, apresentava-se um vetor de validação à rede, e, uma vez treinada, a rede atuava então sobre cada janela anterior a cada “dia” a ser previsto, realizando a previsão desse “dia”. Assim, supondo que a rede utilize uma janela de t dias, para prever o dia d_n , é apresentado à rede treinada o vetor $(d_{n-t}, \dots, d_{n-3}, d_{n-2}, d_{n-1})$, em que d_i representa o valor real do dia i .

Finalmente, basta desfazer a normalização dos dados no período considerado (multiplicando o vetor previsto pelo módulo do maior valor real no período que foi utilizado para previsão). Portanto, de posse dos dados reais e dos dados previstos para o período em questão, é possível traçar os gráficos comparativos e calcular o erro quadrático médio.

3. RESULTADOS

3.1 Previsão do Índice Pluviométrico de Fortaleza

Para a previsão do Índice Pluviométrico de Fortaleza, foi inicialmente realizado um teste sobre a influência do número de camadas e do número de neurônios por camada no Erro Quadrático Médio da previsão, para a escolha da melhor configuração da rede. Os resultados desses 25 testes podem ser observados na Tabela 1, e a melhor previsão realizada pode ser observada na Figura 1(a).



(a) Previsão Normal

(b) Previsão Dessazonalizada

Figura 1: Previsões para o Índice Pluviométrico da Cidade de Fortaleza

O próximo passo foi dessazonalizar a série, para que fosse apresentada à rede uma série menos complexa, com menor estocasticidade. Tomando como período de 12 meses, a fórmula usada foi:

$$mes(t) = \frac{mes(t) + \sum mes(t-1:t-5) + mes(t-6)/2}{12}$$

obtendo-se assim uma série mais suave, mas sem perda de dados relevantes, cuja previsão tornou-se mais fácil. O resultado pode ser observado na Figura 1(b).

Janela	1 camada 5 neurônios	1 camada 10 neurônios	2 camadas 5,5 neurônios	2 camadas 10,5 neurônios	2 camadas 10,10 neurônios
4	0,815585	0,811477	0,708886	0,692017	0,694357
5	0,81956	0,815585	0,703341	0,697156	0,692624
6	0,813379	0,810967	0,703341	0,692017	0,691509
7	0,708886	0,69964	0,69475	0,692017	0,690629
8	0,814444	0,81956	0,813468	0,703341	0,690772

TABELA 1: Erros Quadráticos Médios em Função do Tamanho da Janela e das Configurações das Camadas Escondidas da Rede

É possível notar em [4] um método diferente, que consiste na divisão dos dados de índice pluviométrico em componentes cíclico e irregular antes de apresentar esses valores à RNA, para a previsão do índice anual de chuvas. As principais dificuldades encontradas com esse método, segundo os autores, se relacionaram à obtenção de erros pequenos quanto à previsão, obtendo erros percentuais que ficaram ao redor de 5% (chegando até a 33,64%). Assim, nota-se que a precisão atingida pelo método descrito em [4] é inferior à previsão atingida nesse trabalho, cujos resultados foram plenamente satisfatórios.

3.2 Previsão do Índice de Fechamento das Ações da Embraer e da Vale do Rio Doce

A previsão de ambas as séries foi feita fornecendo, inicialmente, apenas os índices de abertura e fechamento das ações de ambas as empresas (Figura 2, para o caso da Vale do Rio Doce), e, a seguir, foi dado à rede uma maior gama de dados de entrada, como o Índice de Abertura, Fechamento, Máximo e Mínimo Diários (formato *candlestick*), além do Volume Diário de Negociações, como notado nas Figuras 3(a) e 3(b).

Com a previsão da Figura 2, foi feito um estudo sobre a influência do número de épocas de Treinamento e da Taxa de Aprendizado sobre o EQM da previsão da rede, sendo testadas 75 testes para 25 configurações diferentes. Os resultados estão na Tabela 2, que contém a média entre 3 testes para cada configuração, sendo sombreados o melhor e o pior EQM obtidos.

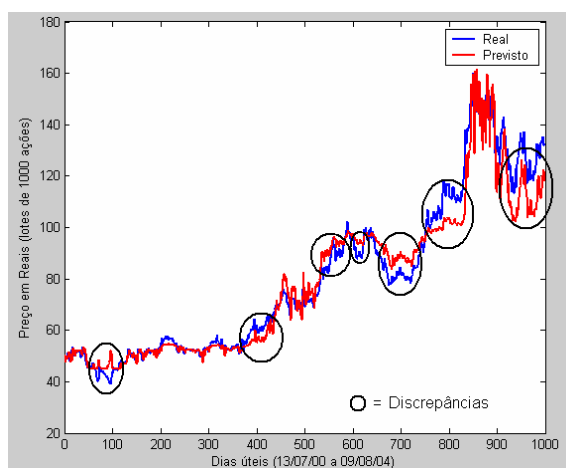


Figura 2: Previsão das Ações da Vale do Rio Doce (dados de entrada: Abertura e Fechamento)

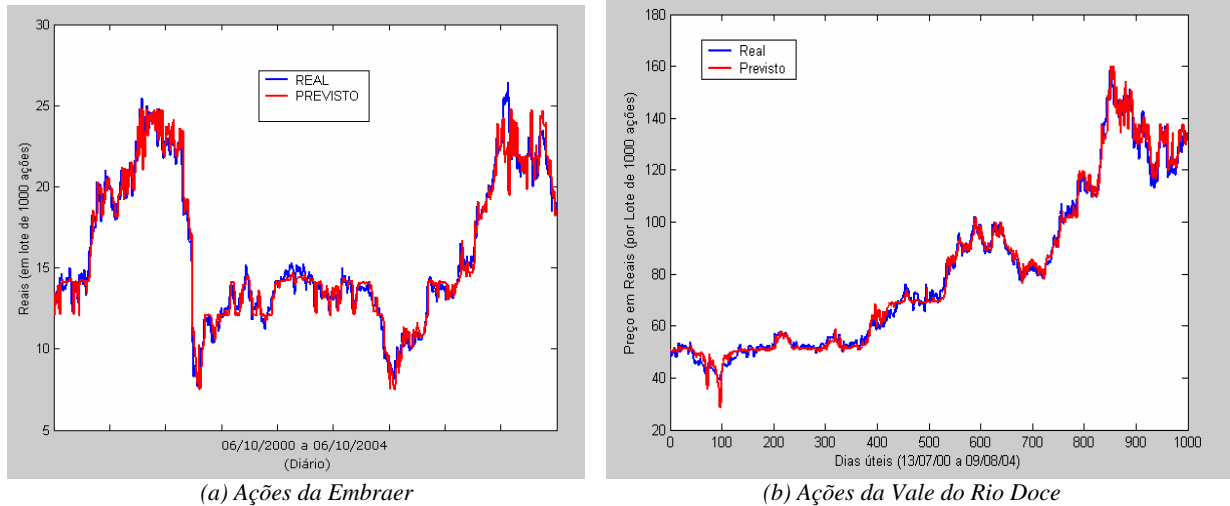
Épocas	Taxa de Aprendizado					
	0,025	0,05	0,075	0,1	0,125	0,15
300	0.0112201	0.0046878	0.00848698	0.0024039	0.0027917	0.00190462
450	0.00774659	0.00170792	0.00201479	0.00250286	0.00114917	0.00111825
600	0.00688915	0.0033622	0.00153792	0.00117834	0.00116572	0.000652299
750	0.00198636	0.00166458	0.00166498	0.000844233	0.00132017	0.00064661
900	0.00140559	0.00317863	0.00109325	0.00097225	0.000492896	0.000458966

TABELA 2: Média Aritmética dos Erros Quadráticos Médios em Função da Taxa de Aprendizado e do Número de Épocas de Treinamento

Um outro método para a previsão de índice de ações pode ser encontrado em [5], em que foi realizada a previsão do IBOVESPA utilizando-se de um critério diferente para o treinamento da RNA. Ao invés do critério da minimização do erro quadrático médio, utilizou-se o critério da lucratividade, em que é mensurada a “acumulação, a cada negociação, da diferença entre o valor previsto e o valor real, ponderado por um sinal de compra (positivo) ou venda (negativo), ao longo do período”. Mesmo concentrando-se em um critério diferente do utilizado neste trabalho, os gráficos obtidos traçando-se os valores previstos e os reais apresentam resultados semelhantes, e por isso pode-se considerar o critério utilizado nesse trabalho bem sucedido.

3.3 Previsão da Diferença entre Fechamento e Abertura das Ações da Embraer

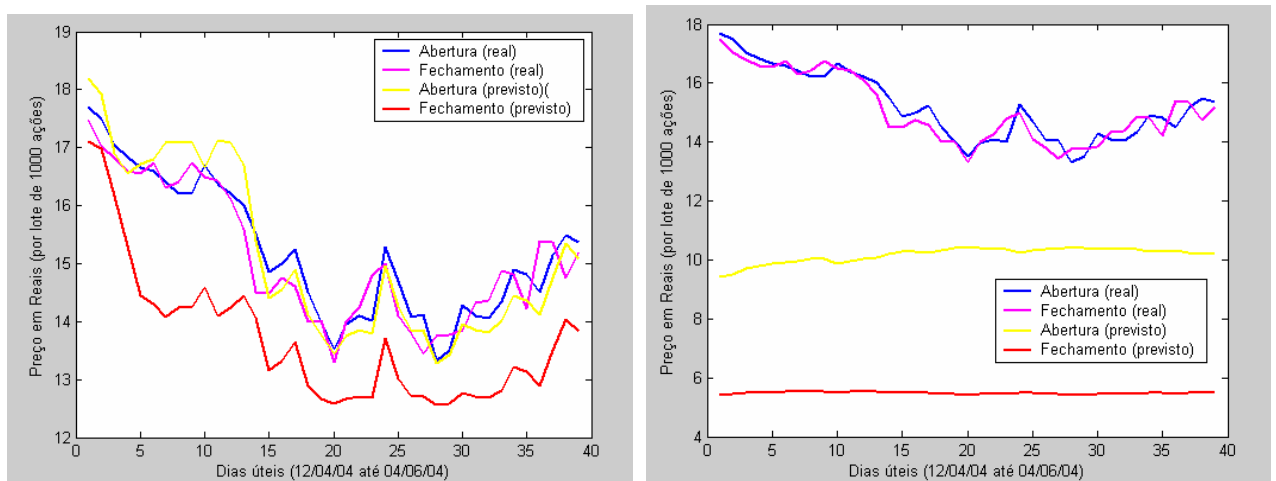
Inicialmente, foi realizada a previsão do Índice de Abertura e Fechamento das Ações da Embraer, fornecendo os dados para que a rede fizesse uma previsão em separado (como é possível ver na Figura 4a) e em conjunto, com uma rede neural de duas saídas (Figura 4b). Para a previsão da Diferença entre Abertura e Fechamento, os dados no formato *candlestick* das Ações da Embraer foram fornecidos como dados de entrada a uma rede *feedforward* e a uma Rede de Elman, sendo obtidas as previsões da Figura 5a e 5b, respectivamente.



(a) Ações da Embraer
(b) Ações da Vale do Rio Doce
Figura 3: Previsão do Índice de Fechamento (dados de entrada: Abertura, Fechamento, Máximo, Mínimo e Volume de Negociações)

4. ANÁLISE E CONCLUSÕES

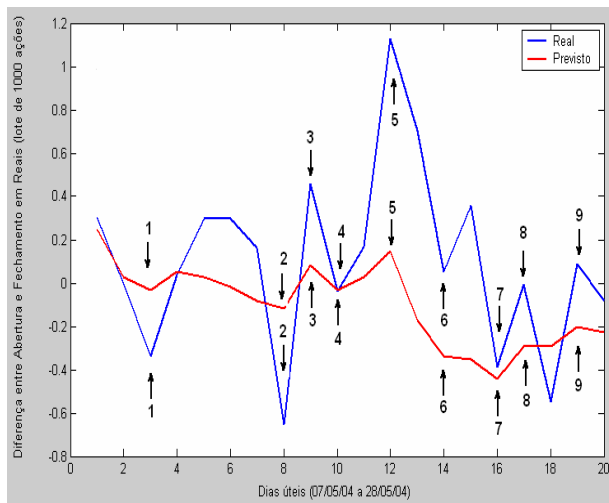
Em relação à previsão do Índice Pluviométrico de Fortaleza, pode-se considerá-la relativamente satisfatória para a série não dessazonalizada, uma vez que a rede aprendeu a sazonalidade da série, mas a previsão deixou a desejar na previsão das tendências. Já a previsão da série dessazonalizada foi melhor, pois a rede já conseguiu prever se em um mês ocorrerá mais chuva que o mês anterior, embora com um considerável erro em seu valor. Contudo, este resultado pode ser de grande valia para o agricultor, que poderá saber se deve plantar em maior ou em menor quantidade do que no mês anterior.



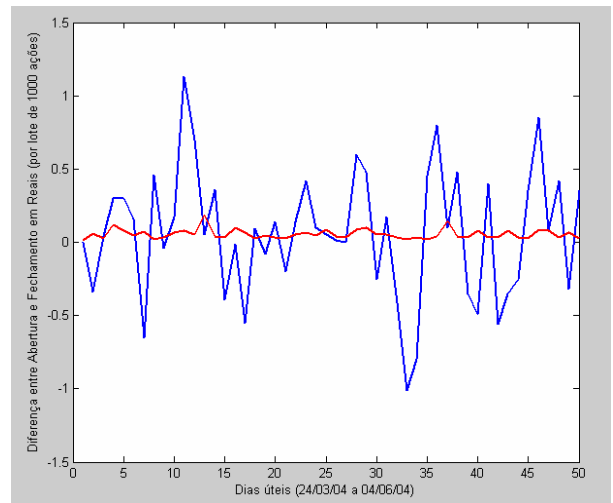
(a) Duas previsões (Rede com 1 saída)
(b) Uma previsão (Rede com 2 saídas)
Figura 4: Previsão do Índice de Abertura e de Fechamento das Ações da Embraer

Através da Tabela 1, é possível notar que o tamanho da Janela de dados de entrada pouco influi no EQM final, sendo mais importante a configuração da rede neural. O que se notou é que, quanto maior a rede, menor o EQM, pois, “se a camada for muito pequena, a rede levará muito mais iterações para atingir a precisão desejada” [9]. No entanto, deve-se procurar uma configuração ideal, pois, “se a camada (escondida) for muito grande, a rede estará memorizando os padrões e com isso perdendo a capacidade de generalização” [9]. Assim, “a rede que melhor generaliza os dados é a menor rede apta a realizar o treinamento de dados” [10].

Através da análise da Figura 2 e das Figuras 3a e 3b, é possível afirmar que a forma como a série temporal é apresentada à rede é muito importante para a obtenção de um resultado preciso, pois notamos que a precisão da Figura 3a é muito maior do que a precisão da Figura 2. Isso porque, com uma maior variedade de dados relacionados entre si, a rede possui maior diversidade de informações para “aprender” a série temporal, o que facilita a sua previsão.



(a) Rede 'Feedforward'



(b) Rede Recorrente de Elman

Figura 5: Diferença entre o Índice de Abertura e Fechamento das Ações da Embraer

Também podemos concluir, a partir da Tabela 2, que o número de épocas de treinamento e a taxa de aprendizado também são dois fatores importantes para diminuir o EQM da previsão, mesmo que estejam sujeitos a certo limite. Isso porque o EQM se estabiliza após um determinado tempo (e aumentar o número de épocas a partir daí influi pouco no EQM) e a previsão da série diverge quando a Taxa de Aprendizado se torna maior que determinado valor crítico (que foi de 0,15 para a previsão em questão). Caso a Taxa de Aprendizado se torne maior do que esse valor crítico, o EQM cresce exponencialmente, chegando à ordem de 10^{120} em 200 épocas de treinamento, pois a rede erra a previsão e se afasta do valor procurado.

Em relação à Previsão da Diferença entre Abertura e Fechamento do Índice das Ações da Embraer, podemos concluir que a melhor previsão foi a feita em separado através da rede *Feedforward* com 1 saída apenas (Figura 4a). Realizando a previsão para uma rede com 2 saídas, notamos que o resultado foi não-satisfatório, vez que a rede não conseguiu nem ao menos identificar a sazonalidade ou as tendências da série.

Já a previsão para a diferença do Índice de Fechamento e de Abertura, para a rede *Feedforward*, pode ser considerada bem-sucedida, já que conseguiu prever a sazonalidade e a tendência da variação, mesmo que apresentando erro em seus valores previstos. A implementação da Rede de Elman, como visto na Figura 5b, não apresentou resultados satisfatórios, pois todas as configurações utilizadas não permitiram que a rede identificasse a tendência da rede. Ao que parece, a rede realiza uma série de aproximações, o que impede a previsão correta da sazonalidade e da tendência da diferença entre os Índices de Abertura e Fechamento.

Dessa maneira, pudemos concluir que o fato mais importante na previsão das séries apresentadas é a maneira como os dados de entrada são apresentados à rede neural, influenciando no EQM também o número de épocas de treinamento, a configuração da rede e a Taxa de Aprendizado utilizada

pela rede. Em todos os casos, foi possível notar que, quanto maior o período a ser previsto pela RNA, maior o Erro Quadrático Médio apresentado, problema comum no estudo das previsões de séries temporais.

A cargo de trabalhos posteriores fica a sugestão de implementar outros recursos em redes neurais que permitam uma previsão mais acertada dos dados, como a utilização de um filtro na rede, a utilização de uma função custo (para determinação exata da arquitetura das camadas escondidas da rede [11]), a aplicação de uma Transformada de Fourier na série ou então a implementação correta da Rede Recorrente de Elman, que não apresentou resultados satisfatórios no presente trabalho, mesmo com os dados de entrada no formato *candlestick*.

AGRADECIMENTOS

Gostaria de primeiramente, e em principal, agradecer ao estimado orientador Professor Dr. Arnaldo Dalpino, por sua compreensão, paciência e apoio seguro durante a fase final do projeto. Devo também agradecimentos aos professores do Curso Fundamental, cujo incentivo foi decisivo para a busca de algo que fosse além das matérias obrigatórias. Finalmente, agradeço ao Professor Massi pela atenção e paciência e ao CNPq, pelo apoio financeiro.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Kovács, Z. L.; “Redes Neurais Artificiais: Fundamentos e Aplicações”, 2. ed., São Paulo: Collegium Cognitio, **1997**.
2. Bovespa – <http://www.bovespa.com.br>.
3. Morettin, P. A.; Mesquita, A. R.; Rocha, J. G. C.; “Rainfall at Fortaleza in Brazil Revisited”, O. D. Anderson, E. A. Robinson and K. Ord (eds.), *Time Series Analysis: Theory and Practice 6, North Holland, Amsterdam, 1985*.
4. Fernandes, L. G. L.; Portugal, M. S.; Navaux, P. O. A.; “Previsão de Séries de Tempo: Redes Neurais Artificiais e Modelos Estruturais”, **1995**.
5. Vieira, R. S.; Thomé, A. C. G.; “Avaliação de Redes Neurais Aplicadas à Previsão de Índices de Mercados de Ações”, **2000**.
6. Oliveira, M. A.; Siqueira, J. O.; “Aplicação do Algoritmo de Rede Neural de Aprendizagem Recorrente em Tempo Real (RTRL) para Previsão da Série do IBOVESPA”, **2003**.
7. Demuth, H.; Beale, M.; “Neural Network Toolbox for use with MATLAB®”, **2003**.
8. Azoff, E. M.; “Neural Network time series forecasting of financial markets”, *John Wiley & Sons, 1994*.
9. Caudill, M.; “Neural Network Primer”, *Miller Freeman Publications, 1990*.
10. Rumelhart, D. E.; Weigend, S. A.; “Predicting the Future: A Connectionist Approach”, *Stanford PDP-90-01, PARC-SSL-90-20, 1990*.
11. Levenbach, H.; Cleary, J. P.; “The Beginning Forecaster: The Forecasting Process Through Data Analysis”, *Lifetime Learning Publications, 1981*.