

Predição do Valor Futuro de Ações Através de Redes Neurais

Fábio Daros de Freitas

Alberto Ferreira de Souza

Francisco Negreiros Gomes

Depto. Informática
U.F.E.S.
Vitória, E.S

Computer Science Dept.
University College London
London, U.K.

Depto. Informática
U.F.E.S.
Vitória, E.S.

Resumo

A utilização de redes neurais na predição de séries históricas tem se mostrado uma alternativa bastante viável aos tradicionais modelos estatísticos, apresentando resultados comparativamente superiores. Este trabalho tem como objetivo relatar os resultados obtidos com a utilização de modelos baseados em redes neurais na predição das cotações futuras de ações negociadas em bolsa de valores. Erros médios absolutos variando entre 4% e 2% foram alcançados.

Abstract

The adoption of neural networks in time series prediction is becoming a strong alternative to the traditional statistics models and is achieving better results. This paper reports the results achieved using networks models to predict future stock market prices. Mean errors range between 4% to 2% has been achieved.

1 Introdução

A abordagem clássica para o problema de predição de séries históricas é baseada em modelos estatísticos. No entanto, a utilização de redes neurais tem se mostrado uma alternativa bastante viável a estes modelos, apresentando resultados significativamente superiores, comparados aos da regressão linear [1], bem como a modelos mais sofisticados como o de *Box-Jenkins* [2].

A predição de séries econômico-financeiras tem atingido bons resultados através dos modelos neurais, em áreas tão distintas quanto indicadores macroeconômicos [3] e mercado de capitais [4]. Muitos dos problemas inerentes à modelagem das séries e a seleção de arquiteturas e algoritmos que garantam determinados níveis de erro de predição ainda são objeto de intensa pesquisa [5, 6]. Este trabalho tem como objetivo relatar os resultados obtidos com a utilização de modelos baseados em redes neurais na predição da cotação futura das séries de ações negociadas em bolsa de valores. Após esta introdução, o problema de predição e a abordagem utilizada são apresentados. Em seguida, o modelo neural proposto é discutido e são apresentados os experimentos e resultados. Finalizando, os resultados obtidos são comentados e são indicados alguns tópicos para futuras investigações.

2 Descrição do Problema

A teoria econômica moderna formula duas hipóteses básicas para o comportamento das séries históricas dos títulos de mercado [7]. A Hipótese do Mercado Eficiente (*Efficient Market Hypothesis -EMH*) estabelece, na sua forma fraca, que os preços dos títulos seguem um movimento aleatório (*random walk*), o que torna impossível prever este movimento baseado nas informações públicas disponíveis sobre o mercado. Diferente da EMH, a formulação das séries como um sistema não linear dinâmico, infere que o preço dos títulos num determinado instante tem componentes que são funções dos seus valores passados e de outros ativos econômicos. Existem

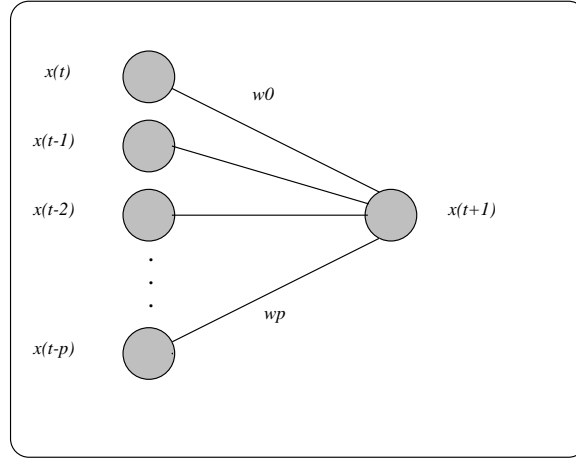


Figura 1: Rede Neural Monocamada $Ar(p)$

argumentações contra e a favor das duas hipóteses, mas nenhum experimento conclusivo, que permitisse eliminar completamente uma delas, foi ainda apresentado.

Neste trabalho, o comportamento das séries históricas dos títulos é tratado como se fosse gerado por um sistema não linear dinâmico, cujas entradas são representadas por valores históricos e a saída é um valor futuro. Assim, problema da predição de títulos foi aqui definido da seguinte forma: Dada a série histórica

$$V = (v_t, v_{t-1}, v_{t-2}, \dots). \quad (1)$$

achar um sistema S , tal que:

$$v_{t+1} = S(V) + \varepsilon \quad (2)$$

onde ε é o erro da predição.

3 Modelo Neural Utilizado

O sistema não linear dinâmico S , descrito na seção anterior, foi implementado através de uma rede neural. A utilização de um modelo neural para a implementação deste sistema foi motivada pelas características de aprendizado e generalização apresentadas pelas redes neurais [8]. As redes neurais conseguem “aprender” a representar o sistema através de um algoritmo de aprendizado apropriado. A rede extrai esta representação de um conjunto de amostras dos sinais de entrada e saída deste sistema. Durante, e também após este processo, a rede demonstra a capacidade de filtrar amostras ruidosas, que é uma característica bastante desejável quando se implementa sistemas que modelam o mundo real.

Um modelo básico para esta rede neural seria baseado no do modelo *Linear Autorregressivo de ordem p* - $Ar(p)$, descrito da forma:

$$x(t) = w_0 x(t-1) + w_1 x(t-2) + \dots + w_p x(t-p) + \varepsilon. \quad (3)$$

onde w_i são constantes, p é a ordem da regressão e ε é o erro da estimativa. Esta rede, que poderia ser utilizada para implementar S , teria uma topologia como a da figura 1. Esta topologia, contudo, apresenta capacidade

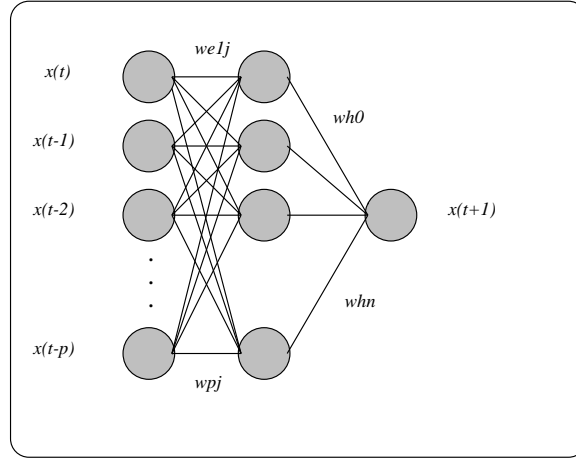


Figura 2: Rede Neural Multicamada $Ar(p)$

representativa bastante limitada [9]. Uma forma de aumentar esta capacidade é introduzir uma camada interna de neurônios e aplicar uma função de transferência não linear ao somatório das entradas de cada neurônio [10]. Desta forma, chegamos à topologia da figura 2. Analiticamente, temos o seguinte modelo matemático para esta rede neural:

$$x(t+1) = \varphi(w_{h_0} NET_1 + w_{h_1} NET_2 + \dots + w_{h_n} NET_n), \quad (4)$$

com,

$$NET_j = \varphi(w_{e_{1j}} x(t-1) + w_{e_{2j}} x(t-2) + \dots + w_{e_{pj}} x(t-p)), \quad (5)$$

e,

$$\varphi(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

onde $\varphi(x)$ é a função de transferência, w_{h_i} é o peso da conexão do neurônio i da camada interna com o neurônio de saída, $w_{e_{ij}}$ é o peso da conexão da entrada $x(t-i)$ com o neurônio j da camada interna e NET_j é saída do neurônio j da camada interna.

Uma rede neural deste tipo, com 3 camadas, foi utilizada nos experimentos realizados. Os números de neurônios da primeira e segunda camadas foram parâmetros topológicos estudados neste trabalho; os resultados são apresentados na seção 4. O número de neurônios da terceira camada é igual a 1.

As redes neurais utilizadas nos experimentos tiveram camadas adjacentes completamente conectadas e foram treinadas com o algoritmo *backpropagation*.

4 Experimentos e Resultados

Os experimentos foram conduzidos com o objetivo de se verificar a capacidade da rede neural de prever o comportamento das séries, observando a sua evolução durante um período de aproximadamente um ano. Utilizou-se um conjunto de treinamento (*training set*) contíguo de 400 amostras das cotações diárias dos títulos, portanto excedendo o período anual, e um conjunto de teste (*test set*) de 97 amostras.

Título	Série
Acesita PN	ACE4
Belgo Mineira PN	BEL4
Metal Leve PN	LEV4

Tabela 1: Séries Históricas dos Experimentos

4.1 Títulos Utilizados

Os experimentos foram baseados nas séries históricas de ações preferenciais nominativas (PN) do setor metal-indústrias, cotadas na Bolsa de Valores de São Paulo (BOVESPA) entre maio de 1994 e maio de 1996. Foram escolhidos os títulos da tabela 1.

4.2 Índice Deflacionador

Face aos níveis dos indicadores inflacionários no período das amostras, foi necessário deflacionar as séries através de um índice não tendencioso a nenhuma delas. De fato, a utilização de ações de empresas do mesmo setor industrial objetivou, em última análise, a facilidade na escolha deste índice. A série histórica do Dólar Comercial foi utilizada como índice deflacionador.

4.3 Pré-processamento

A maioria das aplicações de redes neurais em problemas do mundo real necessita de algum tipo de pré e pós-processamento. Esta situação é causada tanto pela precisão numérica e limitações impostas à representação pelos computadores convencionais, quanto pelos algoritmos de aprendizado existentes. A necessidade de processamento adicional nos experimentos se deu, principalmente, por duas razões:

1. Permitir a correta convergência do algoritmo *backpropagation* durante a fase de treinamento - Quando as entradas contêm amplitudes muito pequenas, observa-se que a baixa relação sinal/ruído estaciona a rede em algum nível de energia mínimo não ótimo. Por outro lado, a presença de amplitudes muito grandes na entrada causa uma saturação dos neurônios da rede, levando a uma saída máxima constante. Estas duas situações impedem a convergência do treinamento para um estado útil. Como processamento adicional, utilizou-se a adição de ruído normal às entradas (o que permite a convergência para entradas pequenas) e a normalização das entradas dos experimentos para uma amplitude numérica máxima igual a 10,0.
2. O modelo de neurônio utilizado tem função de transferência limitada no intervalo $[-1.0, 1.0]$ - Como processamento adicional, os valores desejados (*targets*) foram normalizados para uma amplitude máxima absoluta de 0,8. O objetivo desta escolha é afastar os valores máximos aprendidos pela rede dos níveis de saturação do neurônio da camada de saída.

Os valores utilizados nos processamentos adicionais acima citados foram obtidos experimentalmente.

4.4 Ambiente de Simulação e Parâmetros de Treinamento

Os experimentos foram realizados com o simulador de redes neurais Aspirin/MIGRAINES [11] versão 6.0, em uma plataforma PENTIUM 75MHz com sistema operacional LINUX e interface gráfica X-Windows. Foi utilizado uma taxa de aprendizado de 0,0009 e inércia de 0,95.

4.5 Determinação da Topologia

Foram avaliadas algumas variações da topologia proposta. A primeira avaliação foi feita com o objetivo de determinar o número de neurônios mais apropriado para a segunda camada *hidden*. Durante esta avaliação, foram utilizados 4 neurônios na camada de entrada.

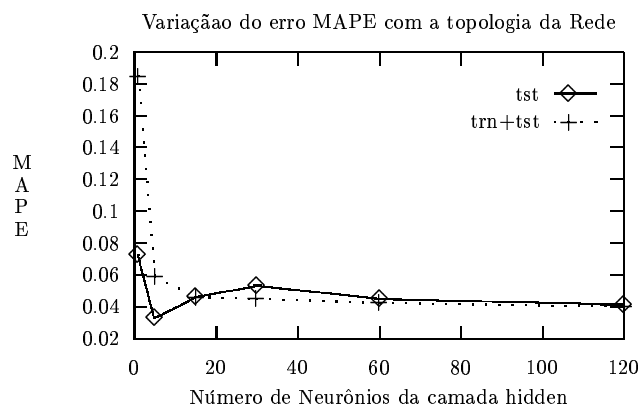


Figura 3: Variação do número de neurônios da camada hidden

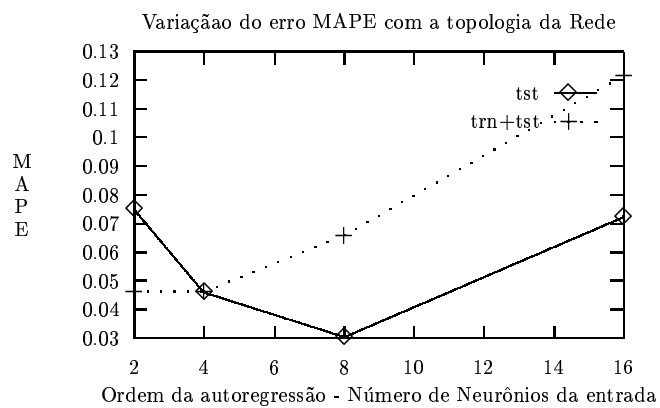


Figura 4: Variação do número de neurônios da entrada

Título	RMS	Pontual Max	MAPE
ACE4	0,028811	0,061787	0,038661
BEL4	0,021528	0,046355	0,026971
LEV4	0,042094	0,092283	0,135753

Tabela 2: Erros de predição obtidos nos experimentos

Verificou-se que, com um número grande de neurônios nesta camada (mais de 20), melhorou a capacidade de representação dos valores de pico da série de treinamento, porém, piorou a representação dos valores pequenos observados na série usada para teste. De fato, com o aumento do número de parcelas somadas pelo neurônio de saída, a representação de pequenos sinais necessitaria (intuitivamente) de pesos bastante pequenos nestas sinapses, o que normalmente não é verificado após a convergência do treinamento. A figura 3 mostra as variações de erro MAPE para a série completa (treinamento + teste) e para a série de teste apenas, em função do número de neurônios da camada interna. O erro *MAPE* - *Mean Absolute Percent Error*, é definido como:

$$MAPE = \frac{1}{n} \sum_{i=1}^{i=n} \left| \frac{target_i - out_i}{target_i} \right|, \quad (7)$$

onde, $target_i$ é o valor real da série, out_i é o valor estimado pela rede e $|\cdot|$ é valor absoluto da diferença. Tendo em vista os resultados obtidos, optou-se por escolher $n = 15$ para a camada interna.

A segunda avaliação topológica realizada foi o efeito da variação do número de neurônios da entrada (*i.e.* a ordem da regressão p) no erro de predição. Foi verificado que com poucos neurônios nesta camada, a capacidade de rastreamento (*tracking*) da rede é prejudicada, e com muitos neurônios, a capacidade de representação de transientes é afetada pela inércia dos p valores passados. Os resultados para as séries completa e de teste são mostrados na figura 4. Foi escolhido $p = 4$.

4.6 Resultados Obtidos

A tabela 2 sumariza os resultados obtidos na predição do valor futuro de cada uma das ações, utilizando a rede neural cuja topologia foi determinada na seção anterior. Além do erro *MAPE*, são apresentados os erros *RMS* e pontual máximo. Os gráficos das figuras 5, 6 e 7 mostram as predições das cotações futuras para cada conjunto de teste (*i.e.* amostras que não foram apresentadas à rede durante o treinamento) após 40.000 interações (*epochs*) de treinamento.

5 Conclusão

O uso de redes neurais para solução do problema de predição de séries econômico-financeiras mostrou-se bastante promissor, mesmo um modelo de variável única (uma única ação é usada como entrada da rede neural). Erros médios absolutos variando entre 4% e 2% foram alcançados.

Futuras investigações serão conduzidas visando a melhoria destes resultados, particularmente através da utilização de um modelo de variáveis múltiplas. Acredita-se que a inclusão de mais variáveis de entrada deve permitir a diminuição dos erros de predição, visto que existem fatores que afetam a economia como um todo [5, 12].

Uma das principais motivações para a realização e continuidade deste trabalho, através dos novos experimentos que ele sugere, é a construção de um sistema otimizador de carteiras baseado na predição de retornos reais com preditor auto-adaptativo [13].

Resultado de ACE4 após 40.000 interações.

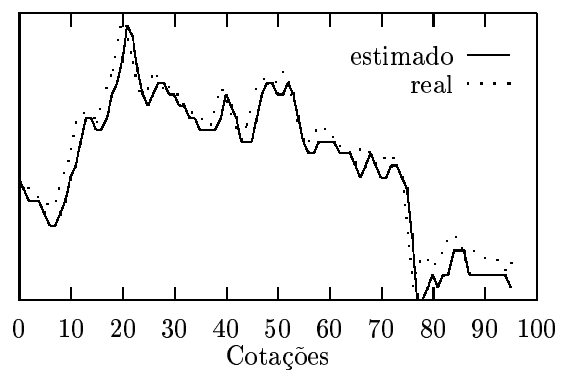


Figura 5: Predição de 01 cotação futura de ACE4

Resultado de BEL4 após 40.000 interações.

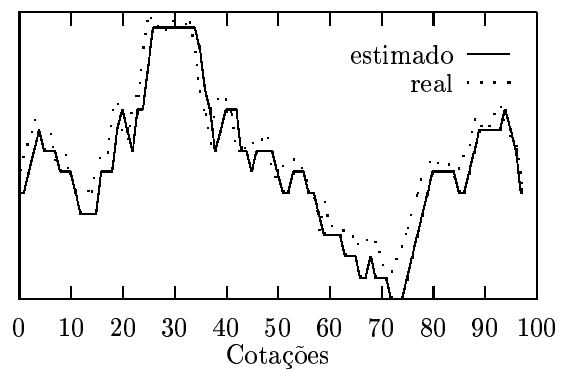


Figura 6: Predição de 01 cotação futura de BEL4

Resultado de LEV4 após 40.000 interações.

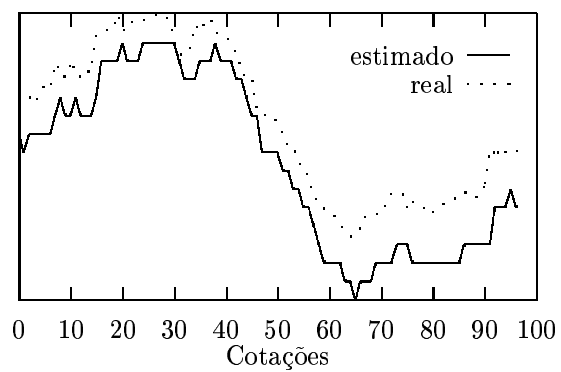


Figura 7: Predição de 01 cotação futura de LEV4

Referências

- [1] L. Marquez, et al, "Neural Networks Models As An Alternative To Regression", *Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences*, 1991.
- [2] B. Sharda, R. B. Patil, *A Connectionist Approach To Time Series Prediction: An Empirical Test*, Journal of Intelligent Manufacturing, 1992.
- [3] B. Freisleben, K. Ripper, *Economic Forecasting using Neural Networks*, Proceedings of the 1995 International Conference on Neural Networks - ICNN' 95, 1995.
- [4] H.White, *Economic Prediction Using Neural Networks: The Case Of IBM Daily Stock Returns*, Proceedings of The IEEE International Conference on Neural Networks, 1988.
- [5] J. Moody, *Economic Forecasting: Challenges and Neural Solutions*, Keynote talk, International Symposium on Artificial Neural Networks, Hsinchu,Taiwan, 1995.
- [6] J. Moody, *Prediction Risk And Architecture Selection For Neural Networks*, Statistics to Neural Networks: Theory and Pattern Recognition Applications, NATO ASI Series F. Springer-Verlag, 1994.
- [7] Sá, Geraldo T., *Mercado de Ações e Bolsa de Valores*, Aplicação Editora Técnica - Rio de Janeiro, 1987.
- [8] Wasserman, Philip D., *Neural Computing: Theory and Praticce*, Van Nostrand Reinhold, 1989.
- [9] Minsky, M., Papert, S., *Perceptrons*, MIT Press, Cambridge, MA, 1969.
- [10] Haikin, Simon, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Co., 1994.
- [11] Leighton, Russell R., *The Aspirin/MIGRAINES Neural Network Software, Release V6.0*, MITRE Corporation, (c) 1992.
- [12] Vishwakarma, Keshav P., *A Neural Network to Predict Multiple Economic Time Series*, Proceedings of The IEEE International Conference on Neural Networks, 1994.
- [13] Araújo, André R., *Smart Analyst-Sistema Inteligente de Apoio à Seleção de Carteiras*, Plano de Negócio apoiado pelo projeto GENESIS/CNPq, 1996.