

## **PREVISÃO DO ÍNDICE Merval: UMA APLICAÇÃO DE REDES NEURIAS POLINOMIAIS GMDH**

CAPORAL, Bibiana<sup>1</sup>; CAVALHEIRO, Everton<sup>2</sup>; CORRÊA, José Carlos<sup>3</sup>; CUNHA, Carlos<sup>4</sup>

**Palavras-chave:** Econometria; Séries temporais; Co-integração;

### **1. Introdução**

Fundada na década de 20, o Mercado de Valores de Buenos Aires (Merval) chegou a movimentar em torno de US\$ 100 milhões por dia nos anos 90, durante o período da paridade cambial da moeda argentina com o dólar. Constantes crises políticas e econômicas, bem como a escassez crônica de crédito, atingiram todos os segmentos da sociedade portenha, acarretando o encolhimento da Bolsa de Buenos Aires. Por outro lado, a Bolsa de Valores de São Paulo movimenta diariamente cerca de US\$ 3,3 bilhões, sendo, aparentemente, o principal parâmetro de preços da Bolsa de Buenos Aires

O índice Merval é considerado o mais importante indicador do mercado acionário argentino, o que torna relevante as pesquisas relacionadas à previsibilidade deste mercado. Ferson (2007), destaca que o interesse pela previsão do comportamento dos preços das ações é provavelmente tão antigo quanto os próprios mercados, bem como a literatura sobre o assunto é vasta e significativa. Neste sentido, o uso de técnicas pertencentes à grande área de inteligência artificial (IA) para a previsão de séries temporais financeiras tem se mostrado cada vez mais presente nos testes empíricos, havendo um especial destaque para o uso da técnica de Redes Neurais Artificiais (RNAs).

De acordo com Haykin (1999), as pesquisas relacionadas RNAs buscam incessantemente modelar o comportamento do cérebro, neste sentido a aplicação de RNA em modelagens matemáticas tem se mostrado cada vez mais presente na vida do homem, em que o uso de RNA na previsão de séries temporais financeiras tem gerado constante demanda. Sua utilização é representativa, pois conforme cita Boose (1994), a área de inteligência artificial é um campo de estudo multidisciplinar e interdisciplinar, que se apóia no conhecimento e evolução de outras áreas do conhecimento.

---

<sup>1</sup> Co-Autor e discente do Curso de Administração da Universidade de Cruz Alta

<sup>2</sup> Autor e Professor do Curso de Administração da Universidade de Cruz Alta

<sup>3</sup> Co-autor e Professor do Curso de Economia da Universidade de Cruz Alta

<sup>4</sup> Co-autor e Professor do Curso de Economia da Universidade de Cruz Alta

O *Group Method of Data Handling* (GMDH) é um uma derivação do campo das Redes Neurais Artificiais. Esse modelo é caracterizado pela sua habilidade de selecionar a combinação de variáveis exógenas mais significativas para prever uma variável endógena, a partir da manipulação dos dados para encontrar o melhor ajuste. Contudo, raros estudos têm utilizado as modelagens polinomiais, especialmente as do tipo GMDH, em séries temporais financeiras. Desta forma, a pergunta de pesquisa deste artigo é: “*O uso de redes polinomiais GMDH é suficientemente capaz de demonstrar não aleatoriedade no retorno mensal do índice Merval?*”.

## 2. Redes neurais artificiais

As Redes Neurais Artificiais (RNAs) são um conjunto de técnicas que tentam simular, em meio computacional, o funcionamento do cérebro humano de uma maneira simplificada. Elas são capazes de reconhecer padrões, extrair regularidades e detectar relações subjacentes em um conjunto de dados aparentemente desconexos. Além disso, elas apresentam capacidade de lidar com dados ruidosos, incompletos ou imprecisos. Sua capacidade de prever sistemas não lineares torna a sua aplicação no mercado financeiro muito objetiva. Para Haykin (1999), rede neural pode ser definida como um processador maciço e paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. A rede neural assemelha-se ao cérebro humano, em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurônios (pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

Já para Lippmann (1987), as RNAs são sistemas físicos que podem adquirir, armazenar e utilizar conhecimentos experimentais, adquirindo uma boa performance, devido à sua interconexão entre os nós da rede. Basicamente há dois tipos de Redes Neurais Artificiais: *feedforward* e *feedback*. Nas redes *feedforward*, os sinais se propagam em apenas uma direção a partir da unidade de entrada, passando pelas camadas intermediárias até a saída. Nas redes *feedback*, os sinais de entrada podem propagar da saída de qualquer neurônio para a entrada em um outro neurônio.

As RNAs apresentam duas fases distintas em sua metodologia, tais são: a da aprendizagem e a do teste. Na fase de aprendizagem, são apresentados estímulos de entrada, denominados padrões de treinamento que farão com que a rede aprenda com os dados. Na fase de teste é demonstrada a

capacidade de generalização da rede, pois seus resultados deverão ser significativos após seus pesos terem sido ajustados na fase anterior.

A propriedade mais importante das Redes Neurais Artificiais é a habilidade de aprender através de seu ambiente e, com isso melhorar o seu desempenho. Segundo Ackley, Hinton e Sejnowski (1985), a habilidade de aprender ocorre através de um processo interativo de ajustes aplicado aos pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

## 2.1. Redes polinomiais GMDH

O método GMDH (*Group Method of Data Handling*) é composto por um algoritmo proposto pelo russo Alexey Ivakhnenko no final da década de sessenta. Ele consiste em um método matemático para estimar estados de um sistema, saídas de controladores e funções de atuadores (IVAKHNENKO, 1969). O algoritmo proposto inicialmente pode ser considerado auto-organizável e de propagação indutiva para a solução de problemas práticos e complexos. Além disso, é possível obter um modelo matemático do processo a partir de observações de amostras de dados, que será utilizado na identificação e no reconhecimento de padrões, ou até mesmo para descrever o próprio processo.

De acordo com Ahmadi, Mottaghitlab e Nariman-Zadeh (2007), o uso de redes de auto-organizáveis do tipo GMDH tem sido bem sucedido em uma ampla gama de áreas de estudo, complementam Mottaghitlab *et al.* (2010) com os bons resultados deste tipo de rede em áreas mais específicas como a Engenharia e a Economia. A maioria dos algoritmos GMDH utiliza funções de referência polinomiais. Uma conexão genérica entre variáveis de entrada e de saída pode ser expressa pela série funcional de Volterra, análoga do polinômio de Kolmogorov-Gabor.

O conteúdo do algoritmo de Ivakhnenko (1969) se desenvolveu como veículo para identificar relações lineares e não lineares entre *inputs* e *outputs*, gerando uma estrutura que tende a ótima, a partir de um processo sucessivo de várias manipulações dos dados, mediante a incorporação de novas camadas.

De acordo com Schneider e Steiner (2006), os algoritmos da rede GMDH são mais poderosos que os modelos estatísticos tradicionais devido aos seus procedimentos auto-organizáveis. Estes

algoritmos extraem informação dos dados existentes utilizando-se da complexidade obtida através de modelos matemáticos flexíveis, ademais, quebram o conceito de caixa preta, característicos das redes neurais.

### 3. Aspectos metodológicos

Neste estudo foi utilizada como variável exógena o retorno mensal, na forma logarítmica, do índice Merval (Argentina), na forma da primeira diferença em  $t-1$ ,  $t-2$  e  $t-3$ , isto é,  $r_{t-1}$ ,  $r_{t-2}$  e  $r_{t-3}$  que formarão  $x_1$ ,  $x_2$  e  $x_3$ , respectivamente. Foram utilizados, na fase de treinamento, os retornos mensais de janeiro de 2000 até fevereiro de 2007, perfazendo 86 observações. Na fase de teste foram simuladas 47 novas observações, entre março de 2007 até janeiro de 2011.

Para avaliar o sucesso das previsões, conforme orientam Ivakhnenko, Ivakhnenko e Müller (1993), foi utilizado a Equação [12]. Os resultados inferiores, ou iguais, à 0,5 seriam considerados como adequados, os que estiverem entre  $0,5 < \delta^2 < 0,8$  seriam considerados satisfatórios, os maiores que 1 seriam considerados como desinformação e as modelagens seriam consideradas como ineficientes.

$$\delta_i^2 = \frac{\sum_1^N (y_i - \hat{y}_i)^2}{\sum_1^N (y_i - \bar{y})^2} \rightarrow \min. \quad [1]$$

Em [1],  $\delta_i^2$  é o critério de Ivakhnenko;  $y_i$  são os valores tabelados da variável endógena e  $\hat{y}_i$  são os valores calculados de acordo com o modelo.

## 4. Resultados

Foram calculadas até 9 camadas na rede para cada uma das previsões dos índices Ibovespa e Merval.

### 4.1. Resultados da fase de teste

Na Tabela 1 são demonstrados os resultados das previsões do índice argentino Merval.

**Tabela 1:** Resultados das previsões para o índice argentino Merval

Camada	$R^2$	Correlação	Sinais	EMQ	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
Camada 1	0,0008	0,0274	0,5745	0,0118	0,0754	0,0660	0,0459	0,0013	1,2308
Camada 2	0,0000	0,0019	0,4681	0,0114	0,0819	0,0644	0,0259	0,0015	1,1861
Camada 3	0,0479	0,2190	0,5106	0,0095	0,0723	0,0530	0,0441	0,0013	0,9847
Camada 4	0,1736	0,4167	0,5745	0,0083	0,0648	0,0451	0,0903	0,0011	0,8651
Camada 5	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 6	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 7	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 8	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 9	0,2103	0,4586	0,5532	0,0081	0,0637	0,0442	0,1185	0,0011	0,8468

Ao se analisar a Tabela 1 evidencia-se, através do critério de Ivakhnenko demonstrado na equação [1], que houve uma evolução significativa na capacidade de previsão do índice argentino, à medida que foram adicionadas novas camadas à rede. Observa-se que, na primeira camada, os resultados eram considerados como desinformação e, à medida que fora aumentada a complexidade dos modelos, pelo acréscimo de novas camadas, os resultados foram considerados como satisfatórios. Todos os demais indicadores [( $R^2$  ajustado, correlação, acerto de sinais, Erro Médio Quadrático (EMQ), Erro Médio Absoluto (MAE) e de tendenciosidade ( $U$ ,  $U^M$  e  $U^S$ )] melhoraram significativamente, havendo um especial destaque para o coeficiente de determinação amostral que as modelagens explicaram 21,03% da variação do retorno mensal do índice da bolsa de valores de Buenos Aires.

## 5. Considerações finais

O interesse em pesquisar não aleatoriedade nos retornos dos índices representativos do mercado acionário tem se mostrado cada vez mais atual. Neste trabalho procurou-se testar a eficiência das redes neurais GMDH na previsão dos índice Merval. Nesse sentido, a tentativa de utilização do método *Group Method of Data Handling* (GMDH) nesse trabalho apresentou resultados significativos na previsão do, denotando certa ineficiência para o período analisado.

A satisfatória previsibilidade do índice argentino pode ser especialmente explicada pela tendência altista deste mercado. O deslocamento da pontuação do mercado acionário argentino, a partir de janeiro de 2003, é contemporâneo a mudanças políticas e macro-econômicas deste país. Esta tendência altista bem como a força deste tipo de modelagem poderia explicar, o sucesso do método em encontrar não aleatoriedade dos retornos do índice Merval.

Para trabalhos futuros sugere-se o uso de modelos com aprendizagem por correção de erros ou aprendizagem baseada em memória, ou ainda modelos de redes recorrentes, já que essa modelagem foi feita através de redes alimentadas diretamente com múltiplas camadas.

### Referências bibliográficas

- ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKI, T. J. **A learning algorithm for Boltzmann machines.** Cognitive Science, vol. 9, PP. 147-169, 1985.
- AHMADI, H.; MOTTAGHITALAB, M.; NARIMAN-ZADEH, N. **Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine.** Journal of Applied Poultry Research, v.16, p.494-501, 2007.
- BOOSE, J. H. **Personal construct theory and the transfer of human expertise.** Proceedings of AAAI- 84, p. 27-33. California, American Association for Artificial intelligence, 1984.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation.** Prentice Hall, 1999.
- IVAKHNENKO, A. G. **Self-teaching systems of recognition and automatic control.** Moscou: Tekhnika, 392, 1969.
- IVAKHNENKO, A. G. IVAKHNENKO, G. A.; MULLER, J. A. **Self-Organization of Optimum Physical Clustering of the Data Sample for Weakened Description and Forecasting of Fuzzy Objects,** Pattern Recognition and Image Analysis, vol, 3, no, 4, pp, 415-422, 1993.
- LIPPMANN, R. P. **An introduction to computing with neurals nets.** IEEE ASSP, 1987.
- MOTTAGHITALAB, M.; *et all.* **Ahmadi Predicting caloric and feed efficiency in turkeys using the group method of data handling-type neural networks.** Poult. Sci., June 1, 2010; 89(6): 1325 - 1331.
- SCHNEIDER, S.; STEINER, M. **Conditional Asset Pricing: Predicting time varying Beta-Factors with Group Method of Data Handling Methods,** SSRN-ID667468, 2006.