

## EXPLORAÇÃO DO USO DE REDES NEURAIS NA PREVISÃO DO COMPORTAMENTO DE ATIVOS FINANCEIROS

*Luiz Antonio Zaneti Jr.*<sup>(\*)</sup>  
*Fernando C. de Almeida*<sup>(\*\*)</sup>

### RESUMO

Este trabalho explora o uso de redes neurais na modelagem do comportamento de ativos financeiros (variação de cotação de ações da Telebrás). Para construção dos modelos, são utilizados dados diários de um período de 3 anos e meio. São utilizadas como variáveis explicativas do comportamento das variações, índices de bolsa de valores bem como o valor das variações das ações de Telebrás em períodos anteriores.

Algumas configurações de rede foram exploradas. As redes criadas não foram capazes de representar o comportamento das variações das ações da Telebrás. Uma das hipóteses de explicação deste resultado é a impossibilidade da rede de representar o alto grau de ruído existente.

---

<sup>(\*)</sup> Pós-graduando em Administração de Empresas na Fundação Armando Álvares Penteado. Engenheiro de Computação pela Escola Politécnica da Universidade de São Paulo. E-mail: luizz@fia.feasp.usp.br.

<sup>(\*\*)</sup> Doutor em Administração pela Ecole Supérieure des Affaires - Grenoble - França; professor no Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e pesquisador da Fundação Instituto de Administração/FEA/USP. E-mail: falmeida@usp.br.

## INTRODUÇÃO

Nos últimos anos uma série de fatores têm permitido um avanço considerável da utilização de técnicas matemáticas e computacionais para a criação de modelos a partir de bases de dados. Muitas destas técnicas apesar de conceitualmente apresentadas a quatro ou cinco décadas, tais como as redes neurais, não podiam ser experimentadas de maneira efetiva, pois não havia capacidade computacional para tanto.

No entanto, atualmente modelos financeiros que exigem forte desempenho computacional, tais como modelos não lineares e técnicas como redes neurais ou algoritmos genéticos podem ser explorados de maneira mais fácil.

Este trabalho apresenta um primeiro esforço de utilização destas técnicas na previsão de retorno de um ativo financeiro na bolsa de valores de São Paulo. A partir dos dados da Telebrás e de alguns outros indicadores procuramos fazer uma previsão dos retornos da Telebrás no dia seguinte.

## Exploração Empírica da Teoria da Arbitragem

### Modelos Financeiros

Diversos modelos financeiros alternativos têm sido explorados a partir do uso de técnicas matemáticas e computacionais. Modelos baseados na teoria do caos partem do princípio que existe uma componente determinística no movimento dos preços dos ativos financeiros e que não segue um processo estocástico (aleatório). A hipótese de que o comportamento dos preços de uma ação pode ser originado a partir de um processo determinístico complexo passa a ser testada (SAVIT 1995).

Já a teoria da arbitragem (APT) proposta por ROSS(1976) pode ser explorada através do uso de redes neurais (AHMADI 1996). A APT assume que o retorno de qualquer papel financeiro é uma função linear de k fatores:

$$R_i = E_i + b_{i1}F_1 + b_{i2}F_2 + \dots + b_{ik}F_k + e_i$$

Onde:

$R_i$  = Taxa de retorno do iésimo ativo

$E_i$  = Retorno esperado do iésimo ativo

$b_{ij}$  = Sensibilidade do retorno do iésimo ativo ao jésimo fator

$F_j$  = jésimo fator comum a todos os ativos tomados em consideração

$e_i$  = termo de erro aleatório para o iésimo ativo

Ross e Roll (1980) exploraram a teoria da arbitragem através do uso da técnica estatística de análise fatorial a fim de estimar a influência de fatores no comportamento do retorno de ativos financeiros.

Em seu trabalho, Ross e Roll utilizam análise fatorial para analisar séries temporais de retornos de ações da NYSE. Foram exploradas 1260 ações em 42 grupos contendo 30 ações cada uma para 2619 períodos. Dos resultados obtidos concluíram que 37 dos 42 grupos tinham no mínimo um fator estatisticamente significativo em cada grupo. Para 24 dos 42 grupos, no mínimo dois fatores eram significativos. De maneira geral, os resultados mostraram que existem dois ou mesmo três fatores que sistematicamente influenciam o retorno das ações.

### Problemas do Uso de Análise Fatorial na Exploração do Modelo de Arbitragem

AHMADI (1996) destaca as limitações da análise fatorial para explorar o modelo de arbitragem. Segundo o autor, a análise fatorial é capaz de dar apoio a modelos ilógicos ou incorretos porque erros de amostragem podem influenciar os resultados obtidos.

Outro aspecto destacado pelo autor é que a habilidade da análise fatorial é bastante dependente da amostra. Isto significa que cada amostra, dos 42 grupos de 30 ações usadas por Ross e Roll, poderia ter fatores distintos. Isto quer dizer que se houvesse 2 fatores em cada uma das 42 amostras, isto poderia indicar a existência de no mínimo 2 fatores ou até 84 fatores distintos. Este procedimento, portanto, não indica nem o número, nem quais os fatores estariam influenciando os retornos dos ativos.

## **Exploração a Partir de Redes Neurais**

A partir da limitação das técnicas estatísticas, Ahmadi apresenta o interesse em explorar técnicas alternativas, em particular redes neurais. Segundo o autor, uma das vantagens é permitir observar todo o padrão de comportamento em vez de olhar seqüencialmente um aspecto de cada vez.

Diversos autores passam a explorar o uso de redes neurais na previsão do retorno de ações cotadas nas bolsas.

Uma vez que é vasto o número de artigos explorando o uso de redes neurais para a previsão de retornos de ações, não vamos entrar em todos os aspectos e detalhes destes artigos. Vamos explorar alguns aspectos que julgamos pertinentes no desenvolvimento de nossa pesquisa.

Muitos dos trabalhos têm como pano de fundo o modelo de arbitragem para o teste de redes neurais. Isto é, assumem a influência de fatores no comportamento dos retornos dos ativos (COLLARD 1996, GATELY 1996, etc.). Outros utilizam apenas o comportamento histórico dos ativos para determinar os retornos futuros (TRIPPI e DeSIEGNO 1992; BERGENSON e WUNSCH 1996; etc.).

Alguns resultados são extremamente positivos. BERGENSON e WUNSCH, por exemplo, através de um sistema híbrido composto de redes neurais e sistema baseado em regras, conseguiram gerar um retorno de 660% em 25 meses a partir de uma aplicação de US\$ 10.000 de 4 de Janeiro de 1989 a 25 de Janeiro de 1991. Um dos aspectos negativos do método, segundo os autores, foi o grande número de horas gasto para a construção do modelo.

DECO, SHUERMAN e TRIPPI (1995) consideraram o comportamento como caótico e desenvolveram um modelo baseado em redes neurais, também com resultados interessantes.

## **Aspectos Relevantes sobre a Aplicação de Redes Neurais**

Destacamos aqui aspectos relevantes sobre a construção de modelos de redes neurais para a previsão de retornos em bolsa de valores.

A utilização de redes neurais exige que se faça uma série de escolhas não triviais na busca de um modelo para o problema. As escolhas envolvem, por exemplo, o modo de implementação, programação de uma rede ou adoção de um *software* voltado para a criação de redes. No caso de uso de pacotes, uma série deles estão disponíveis no mercado, não existindo até o momento um padrão dominante.

Outro aspecto relevante é o método de aprendizado. Diversos autores têm optado pelo uso do método de retropropagação dos erros, popularizado por (RUMELHART e McCELLAND 1986). No entanto, outros métodos como redes probabilísticas são explorados por alguns pesquisadores, ou mesmo variações do método de retropropagação como o método de retropropagação recursivo ou recorrente (GATELY 1996; DECO, SHUERMAN e TRIPPI 1995; etc.). Os resultados têm se mostrado superiores aos obtidos com redes neurais.

A otimização do processo de criação das redes através de algoritmos genéticos (BALAKRISHNAN e HONAVAR 1995), ou a difusão dos dados para melhor representação permitem aumentar a precisão dos modelos (GATELY 1996; VON ALTROCK 1997).

Portanto, a escolha das ferramentas e abordagem adequadas, ou o teste de algumas delas faz parte do processo de desenvolvimento de tais modelos.

A transformação das variáveis é outro aspecto importante do processo de criação dos modelos. A normalização e tratamento adequado de valores faltantes são dois aspectos relevantes. A utilização de médias móveis é outro aspecto importante, utilizado neste trabalho (GATELY 1997).

## **Pesquisa Empírica**

A fim de explorar o uso de redes, selecionamos os valores das cotações das ações da Telebrás de 01/07/94 a 17/03/98, onde procuramos avaliar a capacidade das redes neurais em prever os retornos um dia antes.

### **Seleção das Variáveis**

Com base em trabalhos precedentes, para este trabalho foram selecionados os índices Ibovespa, CDI, Dólar Comercial, Dow Jones bem como os valores de cotação da Telebrás. Outras variáveis poderiam ter sido utilizadas tais como variáveis macro econômicas. No entanto, por indisponibilidade destes dados, para esta pesquisa estes não foram utilizados.

### **Seleção do Período de Coleta de Dados**

O período de coleta de dados foi de 01/07/94 a 17/03/98.

Alguns valores, entretanto, estavam ausentes, em muitos casos, devido a feriados no país. Existem diferentes maneiras de tratamento de dados faltantes ( HAIR Jr. et al. 1998). Neste caso, como grande parte dos dados faltantes eram devido a feriados, estes valores foram preenchidos com o valor anterior, ou seja, foram repetidos.

### **Seleção de Amostras**

A fim de realizar o processo de experimentação através de redes neurais, 3 subconjuntos de dados devem ser selecionados: o de treinamento, o de validação e o de teste. O subconjunto de treinamento serve para treinar a rede, ou seja, é através desses dados que a rede neural é criada. O subconjunto de validação é utilizado para acompanhar o desempenho da rede a partir de uma amostra independente. O aprendizado da rede é interrompido a partir da

observação de seu desempenho sobre os dados de validação. Para testar a rede com um subconjunto de dados independente da criação do modelo, um terceiro subconjunto de teste deve ser selecionado.

Para poder verificar o desempenho da rede em diferentes momentos no tempo, os três subconjuntos foram selecionados em diferentes períodos:

Subconjunto de treinamento: dados de 19/07/94 a 29/01/97

Subconjunto de validação: dados de 30/01/97 a 21/10/97

Subconjunto de teste: 22/10/97 a 17/03/98

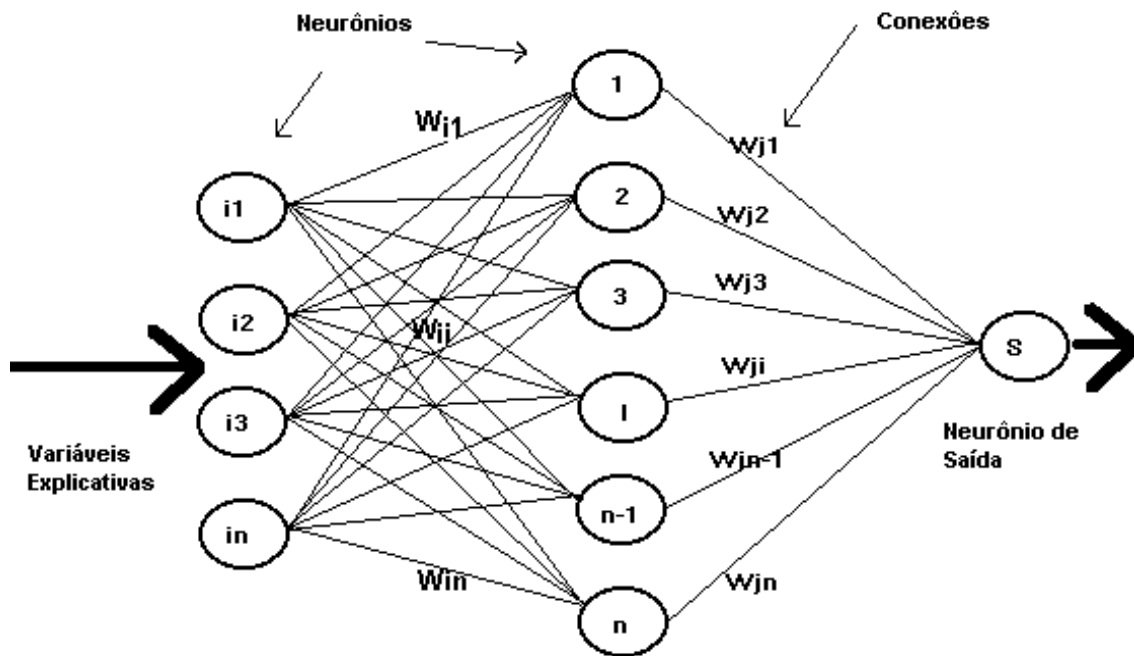
### **Transformação das Variáveis**

Com o objetivo de prever o valor da variação da Telebrás no dia seguinte, foram utilizadas como variáveis explicativas o valor da variação de cada indicador (Ibovespa, CDI, Dólar Comercial, Dow Jones bem como os valores de cotação da Telebrás), um dia, 3 dias e 10 dias antes. Passamos a ter então 15 variáveis de entrada (RAO e RAO 1996).

A fim de melhorar a distribuição dos dados foi utilizado um processo de normalização das variáveis.

### **Criação dos Modelos**

O modelo de redes neurais utilizado foi o de multicamadas, conforme Figura 1.



**Figura 1 - Modelo de Multicamadas**

O método de aprendizado foi o Vanilla Backpropagation (ou Online Backpropagation). Neste método, os pesos da rede são reajustados a cada observação introduzida na rede (ZELL 1995).

Foram utilizadas três camadas nas redes. A primeira, continha 15 neurônios, ou seja, um para cada entrada. O número de neurônios na segunda camada foi alterado ao longo do experimento de forma a testar qual configuração seria a mais adequada. Com isso, algumas redes foram treinadas com 5 neurônios na camada intermediária e outras com apenas 3. Na terceira camada só foi colocado um neurônio, que representava a previsão fornecida pela rede.

## Treinamento

Para comparar duas redes e verificar qual apresentava o melhor desempenho foi utilizado o *Mean Square Error* (MSE) que mede o erro quadrático médio entre os valores observados e os obtidos (previstos) pela rede.

Uma vez inicializada, a rede foi treinada utilizando o conjunto de dados de treinamento. A cada três ciclos de treinamento, o conjunto de

validação era carregado e o MSE da rede para esse conjunto de dados era avaliado. Se o MSE fosse melhor que o obtido anteriormente, a rede era salva como a melhor rede, o conjunto de dados de treinamento era novamente carregado e o treinamento continuava. O treinamento era interrompido caso a melhor rede se mantivesse a mesma após 500 verificações do MSE.

Foram feitos testes com várias configurações de redes, sendo que nos testes os parâmetros variados foram: o número de neurônios na camada intermediária (3 e 5 neurônios) e o parâmetro de aprendizado (entre 0 e 0.2). Os outros parâmetros foram mantidos em todas as configurações ao longo do teste.

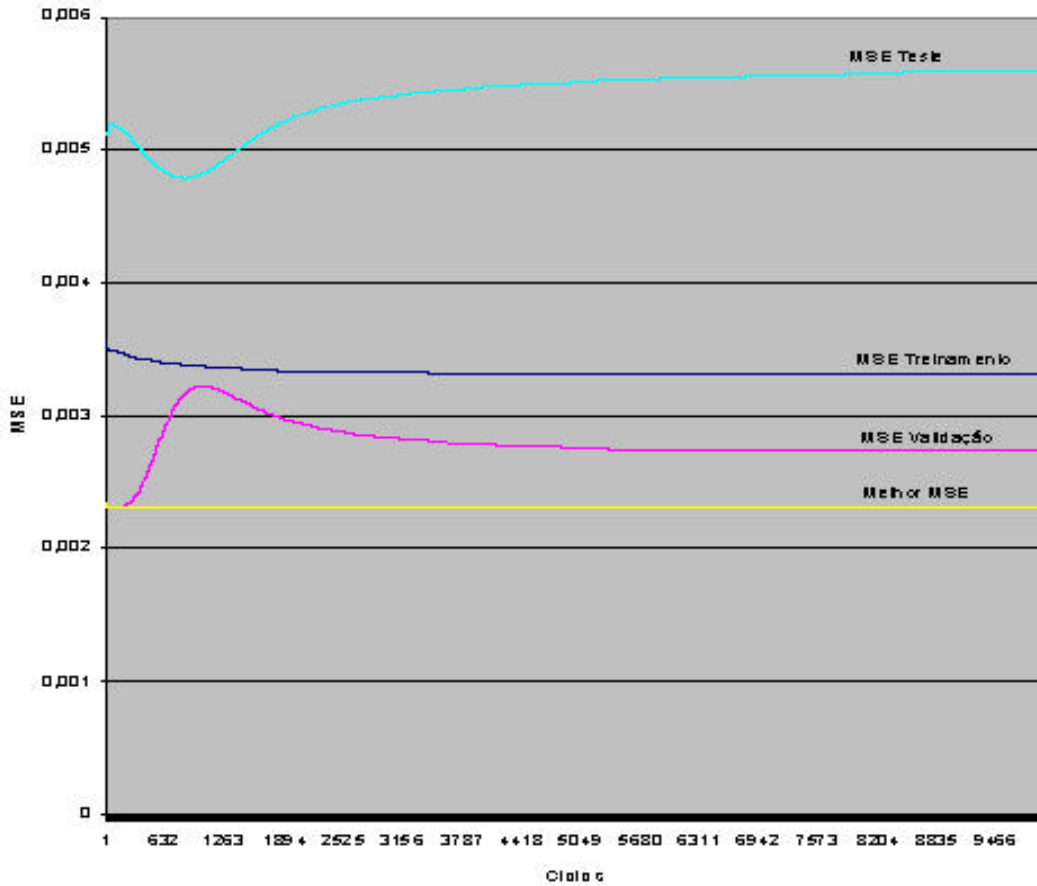
## Resultados

Foi considerada a melhor rede aquela cujo MSE para o conjunto de dados de validação foi o menor. A configuração da melhor rede foi a seguinte: parâmetro de aprendizado = 0.2 e 5 neurônios na camada intermediária.

A figura 2 mostra a curva de MSE ao longo do processo de treinamento. Podemos observar que a curva de validação é menor que a de

treinamento, o que parece, em princípio, incoerente, uma vez que estes dados não participaram do processo de aprendizado. A curva de teste atinge um ponto de erro mínimo e depois volta a crescer.

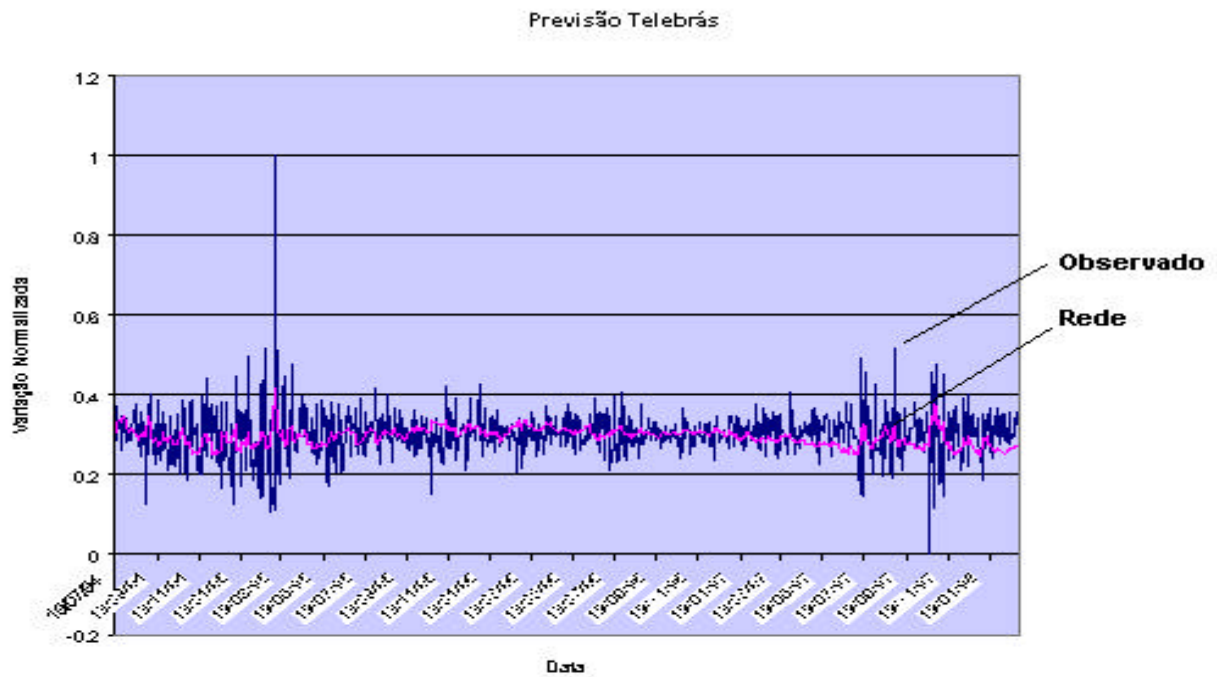
A rede considerada como a melhor, mostrou um menor MSE, para o conjunto de validação, em torno da 300ª iteração, conforme figura 2. Note-se que esta não seria a rede com menor grau de erro para a amostra de teste.



**Figura 2 – Gráfico dos MSE ao Longo do Processo de Treinamento**

A Figura 3 mostra o resultado da melhor rede. Os dados de 19/07/94 até 29/01/97 pertencem ao conjunto de dados de treinamento,

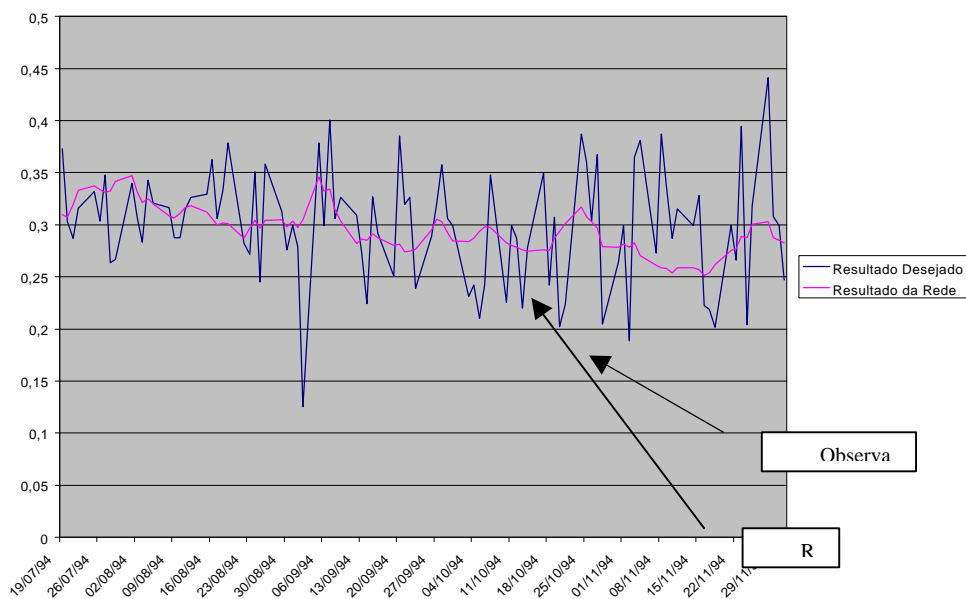
os de 30/01/97 até 21/10/97 são de validação e a partir de 22/10/97 são de teste.



**Figura 3 – Variação Normalizada Observada Versus Obtida pela Rede**

O MSE da melhor rede foi igual a 0,00331447 para os dados de treinamento, 0,00277781 para os de validação e 0,00603721 para os de teste.

Como podemos observar, a rede não foi capaz de acompanhar as variações dos valores da ação. Isto pode ser percebido de maneira mais detalhada na Figura 4 onde um período mais curto é plotado.



**Figura 4 – Variação Normalizada Observada Versus Obtida pela Rede**

Uma das hipóteses de explicação deste resultado é o período muito curto entre as observações. Uma vez que o período de coleta de dados era diária, a intensidade de ruído foi possivelmente muito grande, sendo impossível detectar uma tendência (RAO 1996).

## CONCLUSÕES

A intensidade de oscilação diária é muito acentuada e isso dificultou provavelmente o aprendizado da rede. Seria interessante explorar um intervalo maior entre as observações de modo a minimizar o impacto de ruído, esperando-se que as oscilações sejam menos acentuadas e que seja possível melhorar o desempenho das redes.

Outras variáveis poderiam ser utilizadas, bem como a forma de transformação destas. Ainda um outro aspecto interessante seria fazer uso de métodos econométricos para a transformação de variáveis.

Outro aspecto observado nesta pesquisa é a dificuldade de seleção da melhor arquitetura de rede a partir da seleção de diversos parâmetros. Uma forma de otimizar este processo poderia ser a utilização de algoritmos genéticos para a seleção da melhor rede.

Além da rede *backpropagation* utilizada neste experimento, existem algumas variações desta rede como, por exemplo, *Enhanced Backpropagation* e *BatchBackpropagation* que poderiam ser testadas na tentativa de melhores resultados. Além disso, existem redes com arquiteturas diferentes da *Backpropagation* que também poderiam ser testadas. Como exemplo, podemos citar as *Time Delay Networks* (TDNNs) e *Radial Basis Functions* (RBFs).

O método de avaliação e seleção da melhor rede também poderia ser alterado de forma que a melhor rede fosse aquela que apresentasse o menor MSE para o conjunto de dados de treinamento; e em certos períodos o MSE do conjunto de validação seria avaliado, a fim de determinar quando o treinamento deveria ser interrompido.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AHMADI, H. – “Testability of the Arbitrage Pricing Theory by Neural Networks” .- in *Neural Networks in Finance and Investments*, Robert Trippi and Efrain Turban, Editors, Chicago: Irwin, pp.534-554, 1996.
- BALAKRISHNAN, K. e V. HONAVAR. – *Evolutionary Design of Neural Architectures – A Preliminary Taxonomy and Guide to Literature*. Technical Report, A.I. Research Group, Iowa State University, 1995.
- BERGERSON, K e D.C. WUNSCH II. –“A commodity Trading Model Based on a Neural Network-Expert System Hibrid”. - in *Neural Networks in Finance and Investments*, Robert Trippi and Efrain Turban, Editors, Chicago: Irwin, pp. 581-587, 1996.
- COLLARD, J.E. “Commodity Trading with a Tree Year Old”. - in *Neural Networks in Finance and Investments*, Robert Trippi and Efrain Turban, Editors, Chicago: Irwin, pp. 589-611, 1996.
- DECO, G, B. SHUERMAN e R. TRIPPI. – “Neural Learning of Chaotic Time Series Invariant”.- in *Chaos & Nonlinear Dynamics in the Financial Markets*, Robert Trippi, Editor, Chicago: Irwin, p. 1995.
- GATELY, E. – *Neural Networks for Financial Forecasting* .- N.Y: John Wiley & Sons, 1997
- RAO, VALLURU B. e H. V. RAO – “C++ Neural Networks and Fuzzy Logic”, M & T Books, 2ª Ed., 1996.
- ROLL, R. e S. ROSS. –“An Empirical Investigation of the Arbitrage Pricing Theory”, *Journal of Finance*, Dezembro, pp. 1073-1103, 1980).
- ROSS, S. “The Arbitrage Pricing Theory of Capital Asset Pricing”, *Journal of Economic Theory*, Dezembro, p. 344-360, 1976.
- RUMELHART, D.E., J.C. McCLELLAND, PDP Research Group.- *Parallel Distributed Processing - Exploration in the Microtexture of Cognition*.- Volume 1.-London.- The MIT Press.- 1986.
- VON ALTROCK, C. -Fuzzy Logic and NeuroFuzzy - Applications in Business and Finance .-NJ: Prentice Hall, 1997.
- ZELL, A. et. al. – “Stuttgart Neural Network Simulator User Manual”, v. 4.1, 1995.



## APÊNDICE

### Cálculo da Taxa de Variação

Considerando entrada(i,t) o valor do índice i na data t, a sua taxa de variação ROC (*rate of change*) é (RAO 1996):

$$ROC(i,n) = \frac{\text{entrada}(i,t) - BA(i,t-n)}{\text{entrada}(i,t) + BA(i,t-n)},$$
 onde BA(i,t-n) é a média aritmética de 5 valores do índice i centrada n períodos atrás. Por exemplo:  $BA(i,t-10) = \frac{\text{entrada}(i,t-12) + \text{entrada}(i,t-11) + \text{entrada}(i,t-10) + \text{entrada}(i,t-9) + \text{entrada}(i,t-8)}{5}$ .

### Normalização das Entradas

As entradas foram normalizadas da seguinte forma:

$$\text{entrada\_normalizada}(i,t) = \frac{\text{entrada}(i,t) - \text{média}(i)}{\text{intervalo\_máximo}(i)},$$
 onde média(i) é a média aritmética dos possíveis valores do índice i e intervalo máximo é o maior menos o menor valor possível para o índice i.

Observação: como para o cálculo do ROC(i,10) do índice i são necessários os valores de doze dias antes, as entradas entre 01/07/94 e 18/07/94 foram utilizadas para calcular o ROC(i,10) de outras entradas e então foram descartados. Isto porque não seria possível calcular o ROC(i,10) de tais entradas.

### Saídas

A saída foi calculada da seguinte forma:

$$\text{PrevisãoTelebrás}(t) = 100 \times \frac{\text{Telebrás}(t+1) - \text{Telebrás}(t)}{\text{Telebrás}(t)}$$

Após calcular a saída, ela foi normalizada da seguinte forma:

$$\text{Saída normalizada}(t) = \frac{\text{saída}(t) + \text{deslocamento}}{\text{maior\_valor} + \text{deslocamento}},$$
 onde deslocamento é o valor que somado ao menor valor é igual a zero; e escala é 1 / (maior\_valor + deslocamento).

Com isso, os dados de saída ficaram no intervalo entre 0 e 1.