

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
Rio de Janeiro - RJ - Brasil
ISSN 1413-9022 / v. 25, p. 36 - 53, 2008

PREDIÇÃO DE SÉRIES FINANCEIRAS UTILIZANDO WAVELETS E REDES NEURAIIS: UM MODELO PARA OS FUNDOS DE INVESTIMENTOS IMOBILIÁRIOS

Fabício Soares

Programa de Pós-Grad. em Sistemas e Processos Industriais - Mestrado
Universidade de Santa Cruz do Sul - UNISC
fabriciosoares12@bol.com.br

Rejane Frozza

Programa de Pós-Grad. em Sistemas e Processos Industriais - Mestrado
Universidade de Santa Cruz do Sul - UNISC
frozza@unisc.br

Ruben E. P. Pazos

Programa de Pós-Grad. em Sistemas e Processos Industriais - Mestrado
Universidade de Santa Cruz do Sul - UNISC
rpazos@unisc.br

Resumo

Este artigo apresenta o desenvolvimento de um modelo de predição de séries temporais financeiras com o uso da Rede Neural Artificial TLFN Distribuída (Time Lagged FeedForward Network - Rede Neural Alimentada para frente Atrasada no Tempo), treinada com o algoritmo backpropagation temporal e com o pré-processamento dos sinais de entrada realizado com as Transformadas Wavelets Discretas. A metodologia demonstra como a análise de multirresolução feita com o algoritmo piramidal de Mallat colaborou para o aumento da capacidade de generalização da rede neural, otimizando as previsões feitas pelo modelo implementado. Com a finalidade de demonstrar a eficácia desta metodologia, foi realizado um estudo de caso envolvendo a séries histórica de cotações das cotas, negociadas no mercado secundário, do Fundo de Investimento Imobiliário Almirante Barroso.

Palavras-chave: Rede TLFN distribuída; Wavelets; Predição; Fundos de investimentos imobiliários.

1. Introdução

A expansão do mercado de capitais brasileiro não se deve apenas ao aumento de negócios envolvendo ações de companhias de capital aberto (as Sociedades Anônimas - S.A.), mas também pelo desenvolvimento de novos ativos financeiros como, por exemplo, o caso das cotas dos FIIs (Fundos de Investimentos Imobiliários). Estes fundos têm ampliado a sua participação no mercado de capitais brasileiro nos últimos anos (AMATO *et al.*, 2005). Atualmente, existem aproximadamente 60 fundos, acumulando um patrimônio de 3 bilhões de reais, conforme dados da Comissão de Valores Mobiliários (CVM), porém estes números são muito pequenos se comparados aos fundos imobiliários americanos (*Real Estate Investment Trusts - REITs*) que são importantes fontes de recursos para o financiamento imobiliário nos Estados Unidos.

Os FIIs brasileiros são fundos destinados à aplicação em empreendimentos imobiliários, constituídos na forma de condomínio fechado, ou seja, as cotas que são emitidas e vendidas através de oferta pública não podem ser resgatadas, apenas podem ser revendidas no mercado secundário. As negociações envolvendo este tipo de ativo são realizadas no Mercado de Balcão Organizado (SOMA) da Bolsa de Valores de São Paulo (BOVESPA).

Este estudo vai apresentar um modelo de previsão e análise dos preços das cotas do FII Almirante Barroso, negociadas no mercado secundário, utilizando as Transformadas Wavelets Discretas e as redes neurais artificiais.

O artigo está organizado nas seguintes seções: a seção 2 aborda as características da rede neural artificial desenvolvida; a seção 3 apresenta os conceitos relacionados às *wavelets*; a seção 4 apresenta os resultados alcançados pela pesquisa e a seção 5 apresenta algumas considerações.

2. Redes Neurais Artificiais - RNAs

Muitas áreas do conhecimento têm utilizado a enorme capacidade de identificação de padrões e a capacidade de generalização das redes neurais na solução de diversos problemas. Entre estas está a área de finanças, que busca nas RNAs uma ferramenta para realizar a predição de valores futuros de séries de cotações de opções, *commodities* e, principalmente, de ações.

Em relação à previsão de séries temporais de cotações de ações, Mueller (1996) destaca em seu trabalho que o uso de redes neurais artificiais de múltiplas camadas (MLP - *Multi-Layer Perceptron*), treinadas com o algoritmo *backpropagation*, pode apresentar melhores resultados que outros modelos tradicionais, tais como o método ARIMA e as Médias Móveis Simples e Exponencial.

O desempenho superior das RNAs em relação às médias móveis está ligado a sua capacidade de generalização. Pois, fatores sazonais característicos das séries financeiras, tais como a dificuldade na localização nos espaços tempo-frequência e a presença elevada de ruído no conjunto de dados, exigem esta virtude por parte do modelo empregado para a modelagem.

Vários outros estudos de aplicação de redes neurais na previsão de valores futuros de ativos financeiros já foram realizados no Brasil, dos quais se pode citar, como exemplo, o trabalho de Thomaz e Vellasco (2005) que procurou prever o melhor momento de comprar ou vender as ações da Telebrás (PN – ações preferenciais) entre janeiro de 1995 e maio de 1997. Neste estudo, os autores separaram a série de dados em dois períodos, um para treinar a rede e outro para testá-la. Os resultados apresentados foram satisfatórios, uma vez que a rede conseguiu identificar, em longo prazo, as tendências de alta, baixa e de estabilidade das cotações. O modelo neural utilizado para classificar os vetores de entrada, de acordo com a respectiva tendência, foi o *Self-Organizing* (Auto-Organizável) com o método de aprendizagem por competição (*competitive learning*).

Nos anos noventa, começou o desenvolvimento de modelos de redes adaptados à previsão de séries temporais não-estacionárias, o que tem colaborado para a ampliação do seu uso na análise do mercado financeiro. Moseley (2003), por exemplo, utilizou uma rede TLFN Focada (Rede Neural Alimentada para frente Atrasada no Tempo - Focada) para a predição de valores futuros de quatro séries financeiras obtidas junto ao banco central dos Estados Unidos da América, o *Federal Reserve*. Este tipo de rede apresenta apenas uma diferença importante em relação ao modelo MLP tradicional, visto que, na camada de entrada é adicionado um filtro neural de atraso de tempo, com o objetivo de auxiliar na modelagem da não-estacionaridade dos dados. Por sua vez, o treinamento continua sendo realizado pela RNA *backpropagation* padrão.

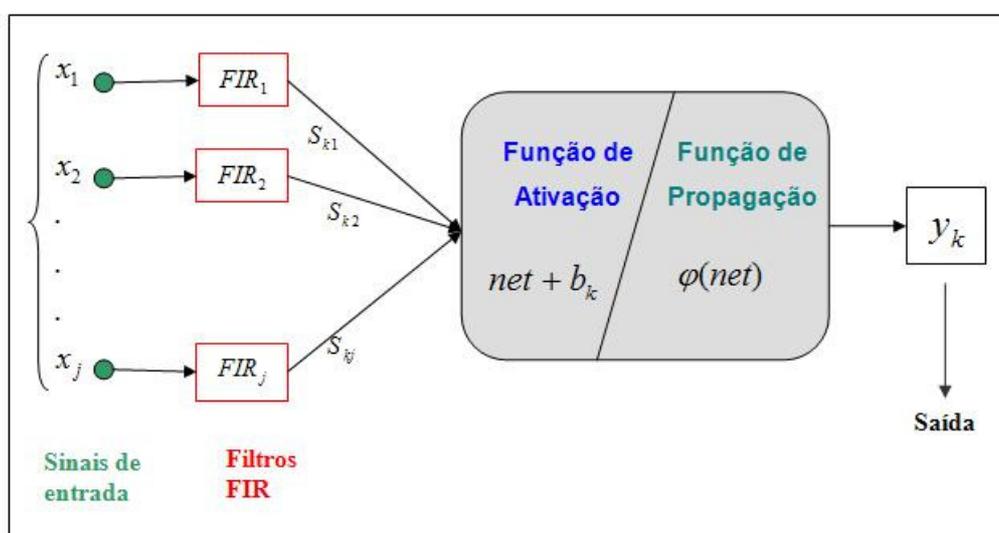
Neste foco, Haykin (2001) apresenta a rede TLFN (*Time Lagged Feedforward Network*, em português, Rede Neural Alimentada para frente Atrasada no Tempo) Distribuída como o modelo mais adequado para o processamento temporal de dados não-estacionários, pois representa uma evolução em relação as redes MLP tradicionais e a TLFN Focada.

2.1 Rede TLFN Distribuída

A diferença fundamental de uma rede TLFN Distribuída para as redes MLP tradicionais é que as redes atrasadas no tempo possuem filtros neurais em cada uma das suas conexões. Desta característica surge a explicação para o uso do termo *Distribuída* que se refere ao fato de que “[...] a influência implícita do tempo é distribuída através da rede.” (HAYKIN, 2001, p. 702).

Na Figura 1, é apresentado o modelo matemático de um neurônio de uma rede TLFN Distribuída. Já os filtros FIR (filtro de resposta a impulso de duração finita) são descritos na Figura 2. Os filtros têm a função de fornecer uma memória de curto prazo para a rede, cada filtro possui um operador de atraso de tempo, assim, valores passados (da iteração $t-1$ até $t-p$, onde p é a ordem do filtro FIR) continuam influenciando diretamente a resposta atual (na iteração t) da rede.

Figura 1 - Modelo matemático de um neurônio com filtro neural de múltiplas entradas



Fonte: Adaptação de Haykin (2001).

Onde:

- x_j : são os sinais de entrada do neurônio k ;
- w_{kj} : são os pesos aplicados a cada sinal de entrada (x_j) do neurônio k ;
- net (função de ativação): é aplicada sobre os sinais de entrada e respectivos pesos;
- b_k (bias): é um parâmetro externo do neurônio artificial k , que tem a função de adicionar ou subtrair um valor da função de ativação (HAYKIN, 2001), com o objetivo de auxiliar no aprendizado;
- φ (função de propagação): é aplicada sobre o resultado da função de ativação;
- y_k : sinal de saída do neurônio k , que pode conectar-se a outras unidades ou ser a própria resposta do sistema, no caso de estar na última camada da rede.

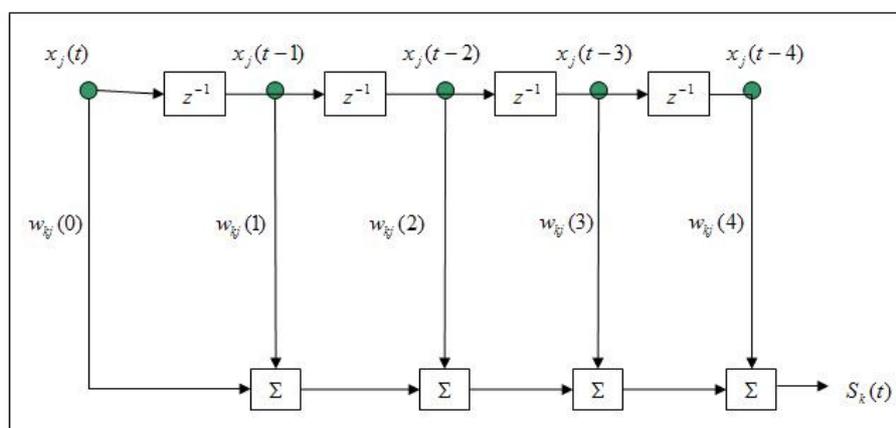
A saída $y_k(t)$ é dada pela equação (1):

$$y_k(t) = \varphi\left(\sum_{j=1}^S \sum_{q=0}^p w_{kj}(q)x_j(t-q) + b_k\right) \quad (1)$$

Onde:

- $w_{kj}(q)$: é o peso da q -ésima sinapse secundária pertencente à j -ésima sinapse primária;
- $x_j(t)$: é o sinal de entrada;
- p : é a ordem do filtro FIR;
- S : é a quantidade de sinais de entrada.

Figura 2 - Filtro FIR de ordem igual a 4 ($p = 4$)



Fonte: Adaptação de Haykin (2001)

Onde z^{-1} é um operador de atraso de tempo unitário, que aplicado sobre $x_j(t)$ resulta na sua versão atrasada $x_j(t-1)$.

A partir do esquema apresentado na Figura 2, pode-se calcular o valor do sinal de entrada $S_k(t)$ do neurônio k pela equação (2):

$$S_k(t) = \sum_{q=0}^p w_{kj}(q)x_j(t-q) \quad (2)$$

Ou através do produto interno dos vetores $w_{kj}(t)$ e $x_j(t)$:

$$S_k(t) = W_{kj}^T X_j(t) \quad (3)$$

Onde $W_{kj}(t) = [w_{kj}(0), w_{kj}(1), \dots, w_{kj}(p)]$ e $X_j(t) = [x_j(t), x_j(t-1), \dots, x_j(t-p)]$. E, portanto, a função de ativação (*net*) pode ser obtida pela equação (4).

$$v_k(t) = \sum_{j=1}^S S_{kj}(t) + b_k \quad (4)$$

2.2 Algoritmo de Aprendizagem *Backpropagation Temporal*

O treinamento de uma rede TLFN Distribuída é realizado com o algoritmo *backpropagation temporal* que é uma derivação do algoritmo padrão. No processamento temporal, há necessidade de que os exemplos do conjunto de treinamento sejam apresentados à rede de forma seqüencial (cronológica) e não aleatória como acontece, normalmente, no treinamento das redes MLP.

Para facilitar a descrição do algoritmo *backpropagation temporal* considere uma rede MLP com uma camada oculta de neurônios, onde a camada de entrada é denominada de i , a intermediária de j e k é a de saída.

No algoritmo original, o cálculo do gradiente local (δ) depende da camada onde se encontra o neurônio. Se for na camada de saída, será obtido pela mesma relação do modelo padrão - $\delta(t) = e_k(t)\phi_k'(v_k(t))$. E, logo, o vetor atualizado dos pesos sinápticos ($W_{kj}(t+1)$) será fornecido pela equação (5), onde δ_k é o gradiente da camada de saída, considerando que os escalares $w_{kj}(t)$ e $y_j(t)$ foram substituídos pelos vetores $W_{kj}(t)$ e $Y_j(t)$ de dimensão $p+1$, onde p é a ordem do filtro FIR (HAYKIN, 2001).

$$W_{kj}(t+1) = W_{kj}(t) + \eta \delta_k(t) Y_j(t) \quad (5)$$

Porém, se o neurônio estiver na camada oculta j (em uma rede com 3 camadas) o cálculo do gradiente local será bem mais complexo e exigirá um custo computacional maior, sendo calculado através da equação (6), cuja demonstração pode ser obtida em Haykin (2001).

$$\delta_j(t) = \varphi'_j(v_j(t)) \sum_{k=1}^C \Delta_k^T(t) W_{kj} \quad (6)$$

Onde:

- $\Delta_k(t) = [\delta_k(t-p), \delta_k(t+1-p), \dots, \delta_k(t)]$;
- C = quantidade de neurônios da camada de saída (camada à direita de j);

E os pesos sinápticos atualizados serão obtidos pela equação (7):

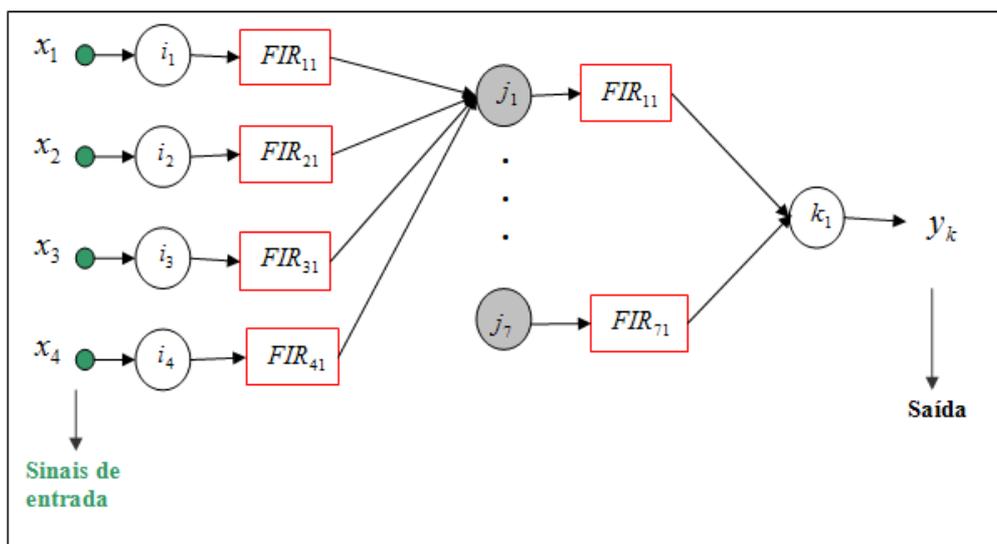
$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_j(t) Y_i(t) \quad (7)$$

Lembrando que o vetor $Y_i(t) = [y_i(t), y_i(t-1), \dots, y_i(t-p)]$ contém os sinais de saída de um neurônio localizado na camada i .

Na Figura 3, é ilustrada a rede TLFN Distribuída que foi implementada no *software Matlab 5.3*. Esta rede tem quatro neurônios na camada de entrada, um para cada variável. Na camada intermediária, a rede possui sete unidades e na de saída apenas uma unidade, que representa o valor que se deseja prever, ou seja, a cotação futura do FII Almirante Barroso. Em cada conexão existente entre as camadas de entrada e a oculta e entre esta e a de saída, haverá um filtro FIR.

Como sinais de entrada serão utilizados os preços de cotação das cotas do FII Almirante Barroso, a taxa básica de juros vigente no país (SELIC- Sistema Especial de Liquidação e de Custódia), a média de pontos do índice IBOVESPA do mês anterior e o índice (provável) de retorno provenientes do aluguel do imóvel, calculado através da divisão do valor distribuído a cada cota no mês anterior a título de aluguel do imóvel pelo valor de fechamento da cota no mercado secundário no instante (dia de negociação) t . Além dos sinais gerados pela aplicação das Transformadas Wavelets Discretas - TWDs sobre estas séries de dados.

Figura 3 - Esquema de uma rede TLFN Distribuída com uma camada oculta



3. Transformadas Wavelets – TWs

Nas diversas áreas da ciência e tecnologia, existem funções que determinam o comportamento de grandezas físicas, tais como as vibrações de uma corda, a distribuição da temperatura numa barra metálica e a velocidade e pressão de um determinado escoamento. Para fazer análises a partir de funções bem mais simples, os cientistas têm obtido decomposições em forma de séries. Com a descoberta do cálculo diferencial, as expansões de Taylor permitiram mostrar a utilidade de expressar em séries de potências as funções que possuem derivadas contínuas até determinada ordem.

Mas isso excluía um amplo leque de funções, por exemplo, as funções periódicas de período T , com valores finitos dos limites laterais nos pontos de descontinuidade, e com derivada contínua no intervalo fundamental de período T , condições que serviram de base para que Fourier desse os primeiros passos para estabelecer a decomposição em séries trigonométricas, satisfazendo determinados critérios, chamados de condições de Dirichlet (MATOS, 2002).

Com o tempo, a análise de Fourier se tornou uma ferramenta muito importante para o estudo dos fenômenos de diversas áreas, especialmente, após a definição da transformada de Fourier, que estabelece uma correspondência entre o espaço temporal e o espaço da frequência. Uma ampla variedade de transformadas de Fourier surgiu, para funções de variável contínua e funções de variável discreta.

Mais tarde, a análise de Fourier tornou-se insuficiente para analisar outros tipos de sinais: eletrocardiogramas, espectrogramas de produtos químicos, de funções com variação abrupta.

Na busca de famílias de funções ortogonais, o matemático húngaro Alfred Haar criou em 1910 a primeira família de *wavelets*, que leva seu nome. Depois seriam definidas famílias de *wavelets* para diferentes aplicações. Nos anos oitenta o geofísico francês Jean Morlet empregou o nome de *wavelet of constant slope*; seus estudos com a transformada janelada de Fourier na prospecção petroléira lhe permitiu concluir que manter uma janela fixa não era bom, então fez o contrário, ele manteve a frequência constante e mudou a janela. Descobriu que esticando a janela, esticava a função, e comprimindo a janela, compactava a função.

Foi com os estudos da belga Ingrid Daubechies (1992), que as funções *wavelets* passariam a ser estudadas a profundidade e seus recursos colocados em prática. Os estudos de Daubechies iniciaram-se em 1988 juntamente com Stephane Mallat. Daubechies desenvolveu estudos direcionados ao processamento de sinais discretos.

As *wavelets* formam uma classe de funções usadas para representar uma função dada, localizando-as no espaço e na escala. Uma família de *wavelets* pode construir-se a partir de uma função $\psi(t)$ de quadrado integrável - isto é, uma função tal que $\int_{\mathfrak{R}} |\psi(t)|^2 dt$ seja um número real, nesse caso se denota $\psi \in L_2(\mathfrak{R})$ - chamada com frequência wavelet geradora (*wavelet mother* em inglês), de forma que gere uma base de funções $\psi_{a,b}(t)$ mediante transformações de dilatação (ou contração) e de translação a partir de $\psi(t)$ (BACHMAN, NARICI & BECKESTEIN, 2000). Desta forma uma família *wavelet* é formada por:

$$\psi_{a,b} = \frac{1}{|a|^{\frac{1}{2}}} \psi\left(\frac{t-b}{a}\right) \quad (8)$$

Onde $a, b \in \mathfrak{R}, a \neq 0$.

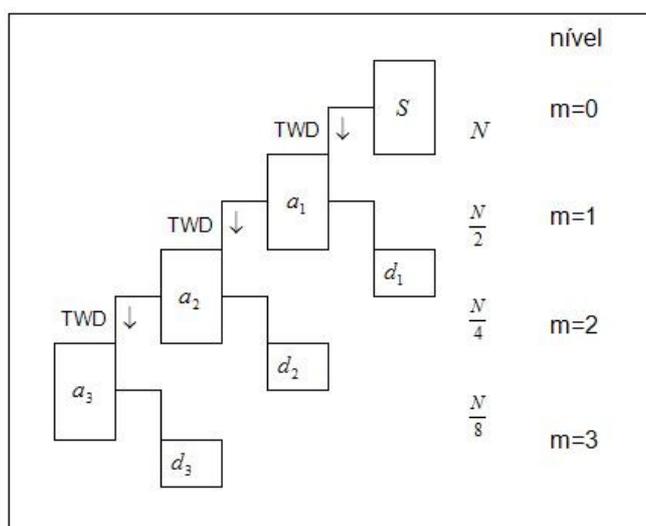
3.1 Implementação das Transformadas *Wavelets* Discretas – TWDs

Uma forma prática de realizar a Análise de Multirresolução – AMR, descrita em detalhes em De Oliveira (2007), aplicando as TWDs é através do algoritmo piramidal de Mallat (REIS & DA SILVA 2004). Na Figura 4, tem-se um esquema do algoritmo

piramidal de Mallat, onde, N é o número de pontos, S é o sinal original, m é o nível de escala (cada vez que se aplica a TWD no sinal: $m = m + 1$), a_m é o sinal suavizado ou tendência geral no nível de escala m e d_m são os “detalhes” do sinal.

Também, é importante ressaltar que para calcular a TWD usando o algoritmo de Mallat é preciso fazer algumas restrições na equação (8): onde $a = 2^m$ e $b = n.2^m$, sendo n (índice de translação) igual ao número de amostras do sinal (GALVÃO *et al.*, 2004).

Figura 4 - Esquema do algoritmo de Mallat



Fonte: Adaptado de REIS & DA SILVA (2004)

Uma forma de implementar, computacionalmente, as TWDs de Haar ou Daubechies, aplicando-se o algoritmo de Mallat, com n coeficientes é através da equação (9):

$$WxS = TWD = a_m + d_m \tag{9}$$

Onde S é a matriz com o sinal original (ou o sinal suavizado a_m , se $m > 0$) e W é a matriz dos coeficientes wavelets, no exemplo da Figura 5 tem-se os coeficientes da Transformada de Haar, onde $c_0 = \frac{\sqrt{2}}{2}$ e $c_1 = \frac{\sqrt{2}}{2}$.

transações com este ativo, pois como se trata de um produto novo no mercado de capitais brasileiro sua liquidez ainda é baixa.

O objetivo foi treinar a rede neural para prever os valores futuros de cotação deste ativo financeiro, um, dez e vinte passos à frente. Além de identificar se variáveis como a taxa básica de juros (SELIC), o índice IBOVESPA e o índice (provável) de retorno provenientes do aluguel do imóvel podem influenciar nos preços das cotas do fundo. Para atingir os objetivos propostos, foram realizadas quatro simulações cujos parâmetros principais estão expostos na Tabela 1.

Tabela 1 - Principais parâmetros dos experimentos realizados

Parâmetro	Simulação 1	Simulação 2	Simulação 3	Simulação 4
Variáveis de entrada	FAMB_a3_h FAMB_d1_h FAMB_d2_h FAMB_d3_h	FAMB_a1_db4 FAMB_d1_db4 FAMB_d2_db4 FAMB_d3_db4	FAMB IND_RENT M_IBOVESPA SELIC	FAMB IND_RENT_a3 M_IBOVESPA_a3 SELIC_a3
Taxa de aprendizagem	0,02	0,02	0,1	0,1
Épocas de treinamento	18	18	18	18
Valor inicial dos pesos	0,0000	0,0000	0,0000	0,0000
Bias dos neurônios das camadas intermediárias e de saída	0,0001	0,0001	0,0001	0,0001
Ordem do filtro FIR das camadas oculta e de saída	1	1	1	1

Já na Tabela 2 é apresentada a descrição de cada variável de entrada, inclusive com a indicação de qual função *wavelet* foi utilizada no pré-processamento dos dados.

Tabela 2 - Descrição das variáveis de entrada

Variável	Descrição
FAMB	Preços de fechamento da cota no Mercado de Balcão Organizado da BOVESPA
FAMB_a3_h	Sinal de aproximação a_3 de FAMB utilizando a função wavelet de Haar
FAMB_d1_h	Sinal de detalhes d_1 de FAMB utilizando a função wavelet de Haar
FAMB_d2_h	Sinal de detalhes d_2 de FAMB utilizando a função wavelet de Haar
FAMB_d3_h	Sinal de detalhes d_3 de FAMB utilizando a função wavelet de Haar
FAMB_a1_db4	Sinal de aproximação a_1 de FAMB utilizando a função wavelet de DAUB4
FAMB_d1_db4	Sinal de detalhes d_1 de FAMB utilizando a função wavelet de DAUB4
FAMB_d2_db4	Sinal de detalhes d_2 de FAMB utilizando a função wavelet de DAUB4

FAMB_d3_db4	Sinal de detalhes d_3 de FAMB utilizando a função wavelet de DAUB4
IND_RENT	Índice de rentabilidade - resultado da divisão do valor pago para cada cota referente aos ganhos de aluguéis do mês anterior pelo preço de fechamento da cota no momento t
M_IBOVESPA	Média de pontos do índice IBOVESPA do mês anterior
SELIC	Taxa básica de juros – valor diário. Fonte: Banco Central do Brasil
IND_RENT_a3	Sinal de aproximação a_3 de IND_RENT utilizando a função wavelet DAUB4
M_IBOVESPA_a3	Sinal de aproximação a_3 de M_IBOVESPA utilizando a função wavelet DAUB4
SELIC_a3	Sinal de aproximação a_3 de SELIC utilizando a função wavelet DAUB4

O conjunto de dados foi dividido em duas partes: uma com 70% dos pontos destinada à fase de treinamento da rede, na qual os pesos sinápticos das conexões dos neurônios foram ajustados. Já os 30% restantes, foram usados na etapa de validação, na qual foi testada a capacidade de generalização do modelo construído. Ainda, é preciso destacar que o ajuste dos pesos sinápticos foi feito a cada padrão do conjunto de treinamento e que eles foram apresentados seqüencialmente à rede. E que foram utilizadas as funções tangente hiperbólica e linear nas camadas intermediária e de saída, respectivamente, da rede.

Como critérios de avaliação dos valores previstos foram utilizados o *MAPE* – Erro Percentual Médio Absoluto – cujo cálculo é feito conforme a equação (12), o *RMSE* - Raiz do Erro Médio Quadrado calculado através da equação (13).

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{d_t - y_t}{d_t} \right| * 100 \quad (12)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (d_t - y_t)^2} \quad (13)$$

Onde: N é o número de exemplos do conjunto de dados, t é o exemplo atual apresentado à rede, d_t é a resposta desejada para o exemplo t e y_t é a resposta calculada pela rede para o exemplo t .

Além destes também foi observado o coeficiente U de Theil, que mede o quanto o resultado obtido pela rede é melhor que uma previsão trivial - aquela em que a melhor estimativa da próxima cotação é o preço atual. A equação (14) apresenta o cálculo do coeficiente U. Destaca-se que se U for menor que 1 (mesmo que em quantidades

pequenas) as previsões feitas pela RNA foram boas, pois alcançaram um desempenho superior às estimativas ingênuas (ABELEM, 1994).

$$U = \frac{\sqrt{\sum_{t=1}^N (d_t - y_t)^2}}{\sqrt{\sum_{t=1}^N (d_t - d_{t-1})^2}} \quad (14)$$

Na Tabela 3 são apresentados os principais resultados alcançados com cada uma das simulações realizadas para as previsões das cotações do FII Almirante Barroso, um e dez passos à frente obtidos pela rede neural. As colunas *Tre* e *Val* referem-se aos valores obtidos para os critérios de avaliação nas fases de treinamento (*Tre*) e validação (*Val*).

Tabela 3 - Principais resultados obtidos

Experimento	Horizonte Previsão *	MAPE (%)		RMSE (R\$)		Coef. de U	
		Tre	Val	Tre	Val	Tre	Val
Simulação 1	1	0,77	0,61	24,00	13,28	0,83	0,99
	10	1,12	0,95	27,30	20,89	0,72	0,90
	20	1,44	1,15	32,30	25,46	0,59	0,93
Simulação 2	1	0,88	0,74	24,50	15,89	0,85	1,19
	10	1,21	1,01	28,70	21,99	0,76	0,95
	20	1,51	1,18	33,60	26,43	0,62	0,96
Simulação 3	1	0,89	1,31	29,06	29,27	1,01	2,19
	10	0,88	1,13	28,99	24,84	0,77	1,08
	20	0,87	1,26	28,18	28,29	0,52	1,03
Simulação 4	1	0,87	1,58	29,45	33,22	1,02	2,49
	10	0,88	1,08	29,20	23,49	0,78	1,02
	20	0,87	1,06	28,40	23,85	0,52	0,87

* passos à frente.

Conforme observado na Tabela 3, as simulação 1 e 2 foram os experimentos que apresentaram os melhores resultados, especialmente, a Simulação 1 que consistiu no desenvolvimento de um modelo de predição híbrido, envolvendo a análise de multirresolução *wavelet* para obtenção dos sinais de entrada (sinal de aproximação a_3 , e sinais de detalhes d nos níveis de resolução 1, 2 e 3 – com a função Haar) que

destacou-se na previsão de curto prazo (1 passo à frente). Já os modelos construídos a partir das simulações 3 e 4, atingiram resultados um pouco inferiores, pois na fase de validação da rede todos os critérios de avaliação mostraram, um declínio na capacidade de generalização da rede. No entanto, a Simulação 4 que utilizou as Transformadas *Wavelets* Discretas (com a função Daubechies de 4 coeficientes –DAUB4) apresentou um resultado um pouco melhor que o experimento 3.

Para mensurar o percentual de influência de cada variável de entrada nas simulações 3 e 4, foi calculada a intensidade das conexões dos neurônios que tinham ligação direta com os sinais de entrada (conexões da camada de entrada com a intermediária), considerando que quanto maior o valor absoluto do peso sináptico, maior sua contribuição para o modelo de previsão (GALETY, 1996). Na Tabela 4 são apresentados os percentuais para cada uma das variáveis.

Tabela 4 - Percentual de contribuição ao modelo de previsão de cada variável de entrada

Experimento	Variável	Valor (%)		
		1	10	20
Simulação 3	FAMB	31,36	30,73	30,29
	IND_RENT	22,53	23,12	23,45
	M_IBOVESPA	21,84	22,05	22,00
	SELIC	24,26	24,10	24,26
Simulação 4	FAMB	30,09	29,76	29,43
	IND_RENT_a3	25,33	25,66	25,75
	M_IBOVESPA_a3	20,17	20,16	20,17
	SELIC_a3	24,41	24,42	24,65

Em virtude dos resultados atingidos não serem do nível esperado (tão bons ou melhores que a Simulação 1), não é possível afirmar, conclusivamente, que os percentuais calculados para cada variável expressem a verdadeira influência que elas têm nos preços das cotas do FII Almirante Barroso, mas são um forte indicativo do que cada uma delas interfere nas cotações deste ativo.

5. Conclusões

Com a metodologia adotada, foi possível desenvolver um modelo de previsão de séries temporais financeiras eficiente. Especialmente aquele em que a análise de

multirresolução, implementada com as TWDs, potencializou a capacidade das RNAs de extrair informações, durante o processo de aprendizagem, e de generalização na fase de validação. Visto que, nas simulações 1 e 2 a decomposição do sinal original em sub-sinais de aproximação (tendências de longo prazo) e de detalhes (curto prazo) resultou nos melhores modelos de previsão. Além de que, a rede TLFN Distribuída mostrou-se adequada para a aplicação em séries temporais não-estacionárias como no caso da previsão das cotações dos FIIs.

Em relação à métrica adotada para mensurar a influência das variáveis SELIC, IBOVESPA e o índice de rentabilidade (IND_RENT) nas cotações do FII Almirante Barroso, os resultados mostram que as variáveis SELIC e índice de rentabilidade tem influência um pouco maior que o IBOVESPA na formação dos preços deste ativo no mercado secundário.

Referências

- ABELEM, A. J. G. **Redes Neurais Artificiais na Previsão de Séries Temporais**. Dissertação de Mestrado. PUC-RIO, Rio de Janeiro, 1994.
- AMATO, F. B. *et al.* Estratégias de Aplicação em Fundos Imobiliários como Diversificação de Investimentos: Uma Análise do Desempenho Recente e seus Fatores de Influência. In: VIII Seminário em Administração, FEA-USP, São Paulo, 2005. **Anais... VIII SEMEAD**.
- BACEN. Desenvolvido pelo Banco Central do Brasil. Apresenta o **Sistema Gerenciador de Séries Temporais – v1.3.2 Módulo Público**. Disponível em: <<https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>>. Acesso em: 01 ago. 2008.
- BACHMAN, G.; NARICI, L, & BECKESTEIN, E. **Fourier and Wavelet Analysis**. New York: Springer, 2000.
- CALADO, L. R.; GIOTTO, R. M. & SECURATO, J. R. Um Estudo Atual Sobre Fundos de Investimentos Imobiliários. In: V Seminário em Administração, FEA-USP, São Paulo, 2001. **Anais...V SEMEAD**.
- DE OLIVEIRA, H. M. **Análise de Sinais para Engenheiros: Uma Abordagem via Wavelets**. Rio de Janeiro: Brasport, 2007.
- GALETY, E. **Neural Networks for Financial Forecasting**. New York: John Wiley & Sons, 1996.
- GALVÃO, R. K. H. *et al.* Estudo Comparativo sobre Filtragem de Sinais Instrumentais usando Transformadas de Fourier e Wavelet. **Revista Química Nova**, n. 6, Vol. 24, p. 874-884. Sociedade Brasileira de Química: São Paulo, 2001.
- HAYKIN, S. **Redes Neurais: Princípios e Práticas**. 2ª ed.. Tradução Paulo Martins Engel. Porto Alegre: Bookman, 2001.
- MATOS, M. P. **Séries e Equações Diferenciais**. São Paulo: Prentice Hall, 2002.

MORSELEY, N. Modeling Economic Time Series Using a Focused Time Lagged FeedForward Neural Network. **Proceeding of Student Research Day**, CSIS, Pace University, New York, 2003.

MUELLER, A. **Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário**. Dissertação de Mestrado. UFSC, Florianópolis/SC, 1996.

REIS, A. J. R. & DA SILVA, A. P. A. Aplicação da Transformada Wavelet Discreta na Previsão de Carga a Curto Prazo via Redes Neurais. **Revista Controle & Automação**, n. 1, Vol. 15, p. 101-108. Sociedade Brasileira de Automática - SBA, Natal/RN, 2004.

THOMAZ, C. E. & VELLASCO, M. M. B. R. Análise de Tendências de Mercado por Redes Neurais Artificiais. **Anais...** 7º Congresso Brasileiro de Redes Neurais, Natal/RN, 2005.

FINANCIAL SERIES PREDICTION USING WAVELETS AND NEURAL NETWORKS: A MODEL FOR REAL ESTATE INVESTMENT FUNDS

Abstract

This paper presents the development of a prediction model of financial time series using the Artificial Neural Network so called Distributed Time Lagged FeedForward (Distributed TLFN). This neural network is trained with the temporal back-propagation algorithm and with preprocessing of input signals performed with discrete wavelet transforms. The methodology shows how the multiresolution analysis of Mallat algorithm has collaborated for the increase of generalization capacity of neural network, optimizing the forecasts developed with the implemented model. In order to demonstrate the effectiveness of this methodology, case study had been realized evolving quotes of quotas, negotiated in the secondary market, of the Real Estate Investment Fund “Almirante Barroso”.

Key-words: *Distributed TLFN network; Wavelets; Prediction; Real estate investment funds.*