

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS E PROCESSOS
INDUSTRIAIS – MESTRADO
ÁREA DE CONCENTRAÇÃO EM CONTROLE E OTIMIZAÇÃO DE
PROCESSOS INDUSTRIAIS

Fabício Soares

MODELO DE PREDIÇÃO FINANCEIRA UTILIZANDO
WAVELETS E REDES NEURAS ARTIFICIAIS

Santa Cruz do Sul, abril de 2009

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Fabício Soares

MODELO DE PREDIÇÃO FINANCEIRA UTILIZANDO *WAVELETS* E REDES
NEURAIS ARTIFICIAIS

Dissertação apresentada ao Programa de Pós-Graduação em Sistemas e Processos Industriais – Mestrado, Área de Concentração em Controle e Otimização de Processos Industriais, Universidade de Santa Cruz do Sul – UNISC, como requisito parcial para obtenção do título de Mestre em Sistemas e Processos Industriais.

Orientador: Prof. Dr. Rubén Edgardo Panta Pazos

Co-orientadora: Profa. Dra. Rejane Frozza

Santa Cruz do Sul, abril de 2009

Fabrcio Soares

MODELO DE PREDIÇÃO FINANCEIRA UTILIZANDO WAVELETS E REDES
NEURAS ARTIFICIAIS

Esta Dissertaçāo foi submetida ao Programa de Pōs-Graduaçāo em Sistemas e Processos Industriais – Mestrado, Ārea de Concentraçāo em Controle e Otimizaçāo de Processos Industriais, Universidade de Santa Cruz do Sul – UNISC, como requisito parcial para obtençāo do tītulo de Mestre em Sistemas e Processos Industriais.

Prof. Dr. Rubén Edgardo Panta Pazos (UNISC/RS)
Professor Orientador

Profa. Dra. Rejane Frozza (UNISC/RS)
Professora Co-orientadora

Prof. Dr. Rolf Fred Molz (UNISC/RS)

Prof. Dr. Augusto Vieira Cardona (PUC/RS)

DEDICATÓRIA

Dedico esta conquista aos meus avós José e Maria (*in memoriam*) pelo incentivo, afeto e bondade que marcaram minha vida.

Também dedico a minha namorada, Helenara, pelo seu amor e compreensão.

AGRADECIMENTOS

Agradeço aos meus pais, Paulo e Oreni, pelo apoio e incentivo e tudo mais que fizeram por mim.

Agradeço a minha irmã, Jéssica, pela amizade.

Agradeço a todos os professores do mestrado pelos ensinamentos que contribuíram para ampliar a minha “*visão de mundo*”.

Agradeço as secretárias do mestrado, em especial, a Janaina, pela disponibilidade e dedicação.

Agradeço a todos os colegas do mestrado pela amizade e companheirismo.

Agradeço aos meus supervisores e colegas da REREV/PO e da GICOP/PO (Caixa Econômica Federal) pela compreensão, incentivo e amizade.

Agradeço ao Prof. Dr. Augusto Vieira Cardona e ao Prof. Dr. Rolf Fred Molz pelas contribuições que enriqueceram muito este trabalho.

Agradeço, em especial, o Prof. Dr. Ruben Panta (meu orientador) e a Profa. Dra. Rejane Frozza (minha co-orientadora) pelo profissionalismo com que corrigiram meus erros, pelo conhecimento que compartilharam comigo, pelas suas valiosas orientações, que direcionaram e ampliaram os horizontes do meu estudo, bem como, pela dedicação e amizade que demonstraram ao longo de todo o período de desenvolvimento deste trabalho. O meu *Muito Obrigado*.

Agradeço a todos os meus familiares e amigos que, com maior ou menor intensidade, contribuíram para o sucesso deste trabalho.

"A utopia está lá no horizonte. Me aproximo dois passos, ela se afasta dois passos. Caminho dez passos e o horizonte corre dez passos. Por mais que eu caminhe, jamais alcançarei. Para que serve a utopia? Serve para isso: para que eu não deixe de caminhar".

Eduardo Galeano

RESUMO

Este trabalho 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, foram realizados estudos de caso envolvendo as séries históricas de cotações das ações das empresas Petrobrás (PETR4) e Telemar (TNLP3), além das cotas, negociadas no mercado secundário, do Fundo de Investimento Imobiliário Almirante Barroso (FAMB11B).

Palavras-chave: Predição; Séries temporais; Mercado de capitais; Rede Neural Artificial TLFN Distribuída; Transformadas *Wavelets* Discretas.

ABSTRACT

This work 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 studies had been realized evolving quotes of stock of Petrobras (PETR4) and Telemar (TNLP3), and the quotas, negotiated in the secondary market, of the Real Estate Investment Fund “Almirante Barroso”

Key-words: Prediction; Time series; Stock Market; Time Lagged FeedForward Network; Discrete Wavelet Transforms.

SUMÁRIO

1 INTRODUÇÃO	16
2 ESTADO DA ARTE	19
2.1 Considerações	22
3 ANÁLISE TÉCNICA DO MERCADO DE CAPITAIS	23
3.1 Considerações	31
4 TRANSFORMADAS WAVELETS	32
4.1 Origens.....	32
4.2 Definição	33
4.3 Análise de Multirresolução - AMR	37
4.4 Implementação das Transformadas Wavelets Discretas - TWDs	42
4.5 Remoção de Ruído (<i>denoising</i>) com as TWDs	44
4.6 Considerações	47
5 REDES NEURAIS ARTIFICIAIS - RNAs	48
5.1 Definição e Histórico	48
5.2 O Neurônio Artificial	49
5.3 Funções de Ativação e Propagação.....	52
5.4 Principais Características e Aplicações das RNAs.....	54
5.5 Métodos de Aprendizagem.....	56
5.6 Modelo Perceptron de Múltiplas Camadas (<i>Multi-Layer Perceptron</i>)	60
5.7 Algoritmo de Aprendizagem <i>Backpropagation</i>	63
5.8 Rede TLFN Distribuída e o Algoritmo <i>Backpropagation</i> Temporal.....	68
5.9 Considerações	73
6 METODOLOGIA	74
6.1 Caracterização da Pesquisa.....	74
6.2 Síntese dos Procedimentos Metodológicos.....	75
6.3 Considerações	77
7 MODELO DE PREDIÇÃO FINANCEIRA UTILIZANDO WAVELETS E REDES NEURAIS ARTIFICIAIS	78
7.1 Caracterização do Problema	78
7.2 Pré-processamento dos Sinais de Entrada da Rede Neural	80
7.3 Topologia da Rede Neural Implementada no <i>Software Matlab</i> [®]	81
7.4 Predição de Séries Temporais Financeiras.....	82
7.4.1 Previsão das Cotações da Ação PETR4.....	84

7.4.2 Previsão das Cotações da Ação TNLP3	92
7.4.3 Previsão das Cotações do FII Almirante Barroso.....	97
7.5 Análise dos Resultados.....	102
7.6 Considerações	107
CONCLUSÃO	108
REFERÊNCIAS	112

LISTA DE FIGURAS

Figura 3.1 – Cotação da PETR4: visualização de retas e canais de tendência	26
Figura 3.2 – Gráficos de <i>candlestick</i> (vela) de alta e baixa, respectivamente	27
Figura 4.1 – Função geradora das <i>wavelets</i> Haar.....	34
Figura 4.2(a) – Gráfico do sinal discreto s_k	36
Figura 4.2(b) – Efeito da transformada discreta <i>wavelet</i> Haar sobre um sinal	36
Figura 4.3 – Sinal original (Ações da PETR4).....	40
Figura 4.4 – Sinal aproximado nos níveis de decomposição $m=1, 2$ e 3	40
Figura 4.5(a) – Sinal de detalhes nos níveis de decomposição $m=1, 2$ e 3	41
Figura 4.5(b) – Sinal de detalhes “ampliados”.....	41
Figura 4.6 – Esquema do algoritmo de Mallat.....	42
Figura 4.7 – Demonstração da operação matricial descrita na equação (4.5)	43
Figura 4.8 – Demonstração da operação matricial descrita na equação (4.8)	44
Figura 5.1(a) – Modelo de um neurônio biológico	50
Figura 5.1(b) – Modelo matemático de um neurônio artificial.....	51
Figura 5.2(a) – Topologia de uma rede acíclica de camada única totalmente conectada.....	55
Figura 5.2(b) – Topologia de uma rede acíclica de múltiplas camadas totalmente conectada.....	56
Figura 5.2(c) – Topologia de uma rede cíclica parcialmente conectada	56
Figura 5.3 – Arquitetura de uma rede MLP	61
Figura 5.4 – Esquemas dos fluxos dos sinais em uma rede MLP	64
Figura 5.5 – Ajuste dos pesos sinápticos	68
Figura 5.6 – Modelo matemático de um neurônio com filtro neural de múltiplas entradas	69
Figura 5.7 – Filtro FIR de ordem igual a 4 ($p=4$).....	70
Figura 6.1 – Esquema de execução da técnica proposta.....	76

Figura 7.1 – Pseudocódigo utilizado para obter os sinais <i>wavelets</i>	80
Figura 7.2 – Esquema de uma rede TNFL Distribuída com uma camada oculta ...	82
Figura 7.3 – Gráfico com as cotações da ação PETR4.....	85
Figura 7.4 – Valores previstos pelo modelo – Simulação 1.....	90
Figura 7.5 – Valores Previstos (Simulação 1) X Intervalo de Confiança	91
Figura 7.6 – Gráfico com as cotações da ação TNLP3.....	93
Figura 7.7 – Gráfico com as cotações do FII Almirante Barroso	98
Figura 7.8 – Curva de evolução do MAPE na fase de treinamento da RNA	103
Figura 7.9 – Detalhes da curva de evolução do MAPE	104

LISTA DE TABELAS

Tabela 4.1 – Principais funções de filtragem (<i>wavelets</i> geradoras)	37
Tabela 4.2 – Principais parâmetros de limiarização	47
Tabela 5.1 – Principais funções de propagação usadas nas RNAs	53
Tabela 5.2 – Síntese do processo de aprendizagem supervisionada	58
Tabela 5.3 – Notações utilizadas na demonstração das equações do algoritmo <i>backpropagation</i>	65
Tabela 7.1 – Principais parâmetros dos experimentos realizados	83
Tabela 7.2 – Principais parâmetros dos experimentos – Série PETR4.....	86
Tabela 7.3 – Descrição das variáveis de entrada – Série PETR4.....	87
Tabela 7.4 – Valor dos parâmetros de limiarização para a série PETR4	87
Tabela 7.5 – Principais resultados obtidos – Série PETR4	88
Tabela 7.6 – Percentual de contribuição ao modelo de predição de cada variável de entrada – Série PETR4	91
Tabela 7.7 – Principais parâmetros dos experimentos – Série TNLP3	94
Tabela 7.8 – Descrição das variáveis de entrada – Série TNLP3	94
Tabela 7.9 – Valor dos parâmetros de limiarização para a série TNLP3	95
Tabela 7.10 – Principais resultados obtidos para a série TNLP3.....	95
Tabela 7.11 – Percentual de contribuição ao modelo de predição de cada variável de entrada – Série TNLP3.....	97
Tabela 7.12 – Principais parâmetros dos experimentos – Série FAMB11B	99
Tabela 7.13 – Descrição das variáveis de entrada – Série FAMB11B	99
Tabela 7.14 – Valor dos parâmetros de limiarização para série FAMB11B	100
Tabela 7.15 – Principais resultados obtidos – Série FAMB11B	100
Tabela 7.16 – Percentual de contribuição ao modelo de predição de cada variável de entrada para a série FAMB11B.....	102

LISTA DE ABREVIATURAS

AMR	Análise de Multirresolução
ARIMA	Autoregressivo Integrado de Médias Móveis
BOVESPA	Bolsa de Valores de São de Paulo
CVI	Cumulative Volume Indicador
CVM	Comissão de Valores Mobiliários
DAUB4	Função geradora <i>wavelet</i> Daubechies com 4 coeficientes
FAMB11B	Código de negociação das cotas do FII Almirante Barroso no mercado de balcão da BOVESPA
FII(s)	Fundo(s) de Investimento(s) Imobiliário(s)
FIR Filter	Filtro de Resposta a Impulso de duração Finita (<i>Finite Impulse Response Filter</i>)
FR	Força Relativa
IFR	Índice de Força Relativo
MACD	Convergência e Divergência de Médias Exponenciais (<i>Moving Average Convergence Divergence</i>)
MAPE	Erro Percentual Médio Absoluto
MLP	<i>Perceptron</i> de Múltiplas Camadas (<i>Multi-Layer Perceptron</i>)
OBV	<i>On Balance Volume</i>
PETR4	Código das ações preferências da Petrobrás na BOVESPA
RBF	<i>Radial Basis Function</i> (Funções de Base Radial)
RMSE	Raiz do Erro Médio Quadrado
RNAs	Redes Neurais Artificiais
SURE	<i>Estimation Stein's Unbiased Risk Estimate</i> (Estimador não viesado do risco de <i>Stein</i>)
TLFN	<i>Time Local Filter Neuron</i> (Neurônio com filtro local de tempo)

(RNA) TLFN	Rede Neural Alimentada para frente Atrasada no Tempo (<i>Time Lagged FeedForward Network</i>)
TNLP3	Código das ações ordinárias da Telemar na BOVESPA
TWDs	Transformadas Wavelets Discretas
TWs	Transformada Wavelets

1 INTRODUÇÃO

O mercado de capitais brasileiro tem registrado um grande crescimento nos últimos anos, fruto da estabilidade econômica que atrai investidores estrangeiros e também pelo ingresso de pequenos investidores nacionais, que pela primeira vez, se aventuram nesse tipo de aplicação financeira.

A cada dia, novas empresas buscam na bolsa de valores uma forma mais barata de conseguir recursos para os seus investimentos. Isso tem propiciado o aumento da oferta de ações dos mais variados ramos da economia e de outros ativos financeiros criados após o Plano Real. Um exemplo destes ativos são as cotas dos Fundos de Investimentos Imobiliários (FIIs), fundos estes que tiveram o seu desenvolvimento incentivado como uma forma alternativa de financiamento para empreendimentos imobiliários.

Nessa onda de desenvolvimento também cresce o interesse por métodos que possam auxiliar os analistas a maximizar os lucros dos investidores da bolsa de valores. Corretoras e instituição financeiras procuram atrair o público, oferecendo serviços e rendimentos diferenciados em relação à concorrência.

Desta forma, este trabalho propõe o desenvolvimento de métodos de predição financeira para a cotação de alguns ativos negociados no mercado de capitais brasileiro, utilizando as Transformadas *Wavelets* e as Redes Neurais Artificiais.

Os métodos utilizados para a análise técnica do mercado de capitais brasileiro tiveram pouca evolução neste período de investimentos crescentes. A maioria dos analistas utiliza somente as médias móveis e gráficos de cotações e volumes negociados, como forma de análise de tendência (CAVALCANTE, MISUMI &

RUDGE, 2005). No que se refere aos procedimentos estatísticos de predição, também estão restritos aos modelos de séries temporais como o ARIMA (Autoregressivo Integrado de Médias Móveis) e de regressão múltipla (FREIMAN & PAMPLONA, 2005).

Estes métodos apresentam grandes dificuldades para prever valores futuros de um ativo financeiro, especialmente, quando a previsão refere-se a períodos de tempo pequenos (*intraday*, um dia à frente, uma semana). Pois, são ferramentas pouco flexíveis, incapazes em muitos casos de modelar estas séries temporais que são intrinsecamente não-estacionárias¹.

Mesmo o emprego de ferramentas computacionais como o das redes neurais, destacado em vários trabalhos acadêmicos, teve pouca aplicação prática. As fortes oscilações nas cotações, ou seja, o que indica a presença de ruído nas séries de dados, comprometem a capacidade de aprendizagem das redes neurais artificiais (RNAs).

Com base neste contexto, o uso da Transformadas *Wavelets* (TWs) associado com as RNAs pode trazer inúmeras vantagens, pois permite a remoção do ruído presente nos dados, com filtros implementados com base na TW. E com a aplicação das RNAs nos dados “limpos” (sem ruídos) pode-se obter um modelo (função) capaz de prever com maior eficiência os valores futuros de um ativo financeiro.

A análise gráfica obtida com as TWs também pode apresentar melhores resultados, se comparada às médias móveis, pois permite a visualização das oscilações dos ativos financeiros em curtos períodos de tempo e de tendências para períodos longos. Além de que, ao contrário das médias móveis, que sempre apresentam um atraso na identificação das tendências (MATSURA, 2007) as TWs obtêm uma perfeita localização tempo - frequência.

¹ São sinais não-estacionários, pois apresentam variações acentuadas nos períodos e frequências de oscilações (altas e baixas) dos preços.

Assim, o objetivo principal desta pesquisa é construir um método para análise do mercado de capitais, que seja capaz de prever valores futuros de um ativo financeiro.

Os objetivos específicos da pesquisa podem ser descritos como:

- Desenvolver um método flexível para a predição de séries temporais financeiras, utilizando as redes neurais artificiais.
- Otimizar o modelo neural de predição, utilizando as TDWs no pré-processamento dos sinais de entrada.
- Retirar o “ruído” que existe em séries temporais de cotações de ativos financeiros visando à otimização do modelo de predição, utilizando os métodos de filtragem (*denoising*) implementados com as Transformadas *Wavelets* Discretas.
- Identificar as variáveis que podem influenciar a cotação de um ativo financeiro negociado na Bolsa de Valores de São Paulo – BOVESPA.
- Analisar os resultados atingidos com a implantação desta proposta.

A seguir, no capítulo 2, serão apresentados alguns estudos realizados, recentemente, dedicados à predição de ativos financeiros. Já no capítulo 3 serão detalhados alguns métodos utilizados na Análise Técnica do mercado de ações. Nos capítulos 4 e 5 serão descritos os conceitos principais da teoria das Transformadas *Wavelets* e das Redes Neurais Artificiais, respectivamente. No capítulo 6 será descrita a metodologia utilizada no desenvolvimento do presente trabalho. E, no capítulo 7, serão apresentados os principais resultados obtidos com os estudos de caso realizados.

2 ESTADO DA ARTE

Existem vários trabalhos dedicados ao estudo da Análise Técnica do mercado de capitais, sendo que grande parte destes são patrocinados por corretoras e instituições financeiras que atuam no mercado. Entre as obras, recentes, dedicadas ao grande público e com objetivo educacional, destaca-se o trabalho de MATSURA (2007) que, além de apresentar as principais ferramentas utilizadas pelos analistas profissionais, as quais serão descritas em detalhes no capítulo 3, também descreve a metodologia para a criação de uma estratégia de investimento de sucesso que utiliza como suporte alguns métodos gráficos. O autor, ainda, salienta que no atual estágio de desenvolvimento do mercado de capitais é indispensável a criação de instrumentos computacionais que otimizem os métodos tradicionais, além do desenvolvimento de outros como os projetos envolvendo redes neurais artificiais.

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, 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 tempo-freqüência 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 desenvolvido pela autora De Castro (2001) o qual se concentrou na previsão de séries temporais com o uso da rede RBF (*Radial Basis Function*). Entre as séries analisadas estavam as de algumas ações negociadas na BOVESPA (como, por exemplo, as da Petrobrás e da Eletrobrás) os resultados também foram bons. Uma característica importante dos experimentos realizados no referido estudo diz respeito ao tamanho dos conjuntos de dados que tiveram aproximadamente 350 pontos. Isto demonstrou a grande capacidade de a rede aprender mesmo com poucos exemplos.

Outro estudo relevante foi realizado por THOMAZ & 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*).

Já os autores ZANETTI JR. & ALMEIDA (1998), destacaram os principais problemas que dificultam o uso dos modelos neurais na predição financeira. Esta pesquisa, que também utilizou séries históricas de cotações da ação Telebrás PN, teve como objetivo testar se as RNAs seriam capazes de prever as cotações diárias. Neste estudo, porém, os resultados alcançados foram péssimos, visto que o modelo neural não foi capaz de prever as oscilações diárias das cotações. Sendo que, segundo os autores, o que comprometeu a aprendizagem da RNA foi a constante presença de “ruído” na série de dados originais.

Pois, apesar das redes neurais serem menos sensíveis aos ruídos - que nas séries financeiras está intrinsecamente ligado ao “sobe-desce” do mercado de capitais - do que os modelos tradicionais de regressão, a previsão realizada por elas também é prejudicada. Em alguns casos, extremos, pode até ser totalmente comprometida pelas características oscilatórias (tempo-freqüência) do sinal estudado.

Também sugeriram que se houvesse um pré-processamento nos dados, o modelo neural poderia ter um desempenho melhor. Neste estudo foram utilizadas as redes *perceptron* de múltiplas camadas (MLP- *Multi Layer Perceptron*) treinadas com o algoritmo de aprendizagem *backpropagation online*, onde os pesos sinápticos eram atualizados a cada exemplo apresentado à rede.

Para resolver o problema do ruído, REIS & DA SILVA (2004) propuseram um método no qual a série de dados original passaria primeiro por uma filtragem, utilizando a Transformada *Wavelet* Discreta - TWD através do algoritmo de Mallat. Com este pré-processamento, seria eliminado o ruído e, com a análise de multirresolução, poderiam ser distinguidas as oscilações significativas (que representam eventos que precisam ser previstos) daquelas que são somente aleatórias. Este modelo foi aplicado para a previsão de carga de energia elétrica a curto prazo (01-24 horas) obtendo resultados excelentes, uma vez que conseguiram aumentar a capacidade de aprendizagem da rede neural. Foi usada uma rede MLP treinada com o *backpropagation* padrão.

Nesta mesma linha de estudo, JENSEN (1997) destacou que a análise de multirresolução é uma das principais contribuições que as *wavelets* podem proporcionar para a avaliação do comportamento das cotações de ativos financeiros. Visto que, permitem a localização temporal de picos (que indicam a ocorrência de eventos relevantes) e de tendências de longo prazo.

O desenvolvimento de modelos de RNAs, adaptados à previsão de séries temporais, também tem colaborado para a ampliação do seu uso na análise do mercado financeiro. MOSELEY (2003) 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 *backpropagation*.

2.1 Considerações

Desta forma, a utilização de um modelo de rede neural adequado ao domínio, juntamente com uma técnica de eliminação ou redução de ruídos, poderá contribuir de forma significativa na construção de uma ferramenta para a predição financeira.

3 ANÁLISE TÉCNICA DO MERCADO DE CAPITAIS

As duas principais escolas de análise de investimentos no mercado de capitais, segundo CAVALCANTE, MISUMI & RUDGE (2005), são a fundamentalista e a técnica. Enquanto a primeira se encarrega do estudo do desempenho financeiro e das perspectivas de crescimento do lucro e/ou da produção da empresa emissora do ativo financeiro; a segunda analisa os gráficos das cotações da ação na bolsa de valores, considerando que todas as informações referentes à empresa e as demais variáveis econômicas que possam afetar o valor do ativo já estão incorporadas aos gráficos. Em síntese: “[...] a escola fundamental mostra por que os valores se comportam dessa forma. A escola técnica mostra como os preços se comportam.” (CAVALCANTE, MISUMI & RUDGE, 2005, p. 213).

Ainda, conforme os mesmos autores, a escola técnica tem como principais objetivos mensurar a oferta e procura por determinado ativo, destacar oportunidades de operações atraentes, indicar os limites de oscilação dos preços e auxiliar na concepção de estratégias para minimizar os riscos.

A análise técnica (ou gráfica) do mercado de ações teve início com a “Teoria Dow” criada por Charles Dow no começo do século XX. A idéia central da Teoria Dow é identificar as tendências significativas do mercado (DESCHATRE, 1997). Um outro fator importante diz respeito à forma como as ações se relacionam, pois normalmente, o mercado se comporta em conjunto. Isto foi a base para a criação de diversos índices de ações como, por exemplo, o Dow Jones da Bolsa de Nova Iorque e o IBOVESPA da Bolsa de São Paulo, que mostram o desempenho das principais ações negociadas nestes mercados.

Os pontos essenciais do movimento dos preços, segundo a Teoria Dow (MATSURA, 2007 e DESCHATRE, 1997) são os seguintes:

- *Tendência Primária (ou Movimento Primário)*: É uma tendência de longo prazo (de baixa ou alta), que pode durar de um a dois anos.
- *Tendência Secundária (ou Movimento Secundário)*: É uma tendência de médio prazo que dura de 3 semanas até alguns meses. Este movimento, normalmente, corrige metade da tendência primária da qual faz parte, mas pode chegar até dois terços.
- *Tendência Terciária (ou movimentos menores)*: É uma tendência de curto prazo que pode durar horas, dias ou algumas semanas. Ela está inserida dentro dos movimentos secundários e primários.

MATSURA (2007) salienta que uma tendência primária de alta pode ser dividida em três fases:

- *1ª fase*: Acumulação – início da tendência na qual apenas alguns investidores possuem informações qualificadas para comprar neste momento.
- *2ª fase*: Alta Sensível – momento em que a tendência pode ser confirmada, fase ideal para comprar.
- *3ª fase*: Euforia – momento em que até os leigos entram no mercado, fase em que a tendência está para ser revertida, e que, portanto, os investidores qualificados devem começar a vender.

Conforme o mesmo autor, a tendência de baixa divide-se em três fases também:

- *1ª fase*: Distribuição - é a extensão da fase de euforia, onde os principais investidores vendem os seus ativos.
- *2ª fase*: Baixa Sensível - momento em que a tendência pode ser confirmada, logo é a fase ideal para vender as ações.

- 3ª fase: Pânico - é o final da tendência, quando os leigos vendem e os profissionais qualificados começam a comprar antecipando o movimento de alta que vai começar em seguida.

DESCHATRE (1997) também destaca a relação que existe entre os preços e os volumes negociados, conforme a Teoria Dow:

- A alta simultânea de preços e volume indica que a tendência de alta vai continuar.
- A alta de preços e a redução do volume indicam que a tendência de alta está para ser revertida.
- Já a queda dos preços e o aumento do volume indicam que a tendência é de baixa.
- Porém, quando os preços e o volume estão em queda, indica que a tendência de baixa está no fim.
- Quando o volume está estável não há efeito sobre os preços.

Atualmente, para identificar tendências a maioria dos analistas utiliza retas de suporte/ resistência e de baixa/ alta. A reta de suporte representa o patamar mínimo de preço e a de resistência o nível máximo de preço. Já a reta de tendência de alta é construída a partir dos pontos mais baixos das cotações, que aumentam de valor de forma sistemática, caracterizando a tendência de alta. E a reta de tendência de baixa é construída a partir das cotações de maior valor, que reduzem de valor ao longo do tempo, caracterizando a tendência de baixa.

A partir de uma reta de tendência é possível construir canais de alta ou baixa (dependendo da tendência), por onde os preços oscilam durante uma determinada tendência. Para traçar um canal, basta construir uma reta paralela a de tendência que abrigue os topos das cotações, no caso de uma tendência de alta, ou de fundos se for uma tendência de baixa.

No gráfico da figura 3.1 está a evolução das cotações, de janeiro a dezembro de 2001, das ações preferenciais da Petrobrás negociadas na BOVESPA com o

código PETR4. Neste gráfico, foram traçadas as retas de suporte e resistência e identificadas, através das retas e canais respectivos, uma tendência de alta e outra de baixa. Também, pode ser visto que as retas de suporte e resistências delimitam o começo e fim dos canais de baixa e alta.

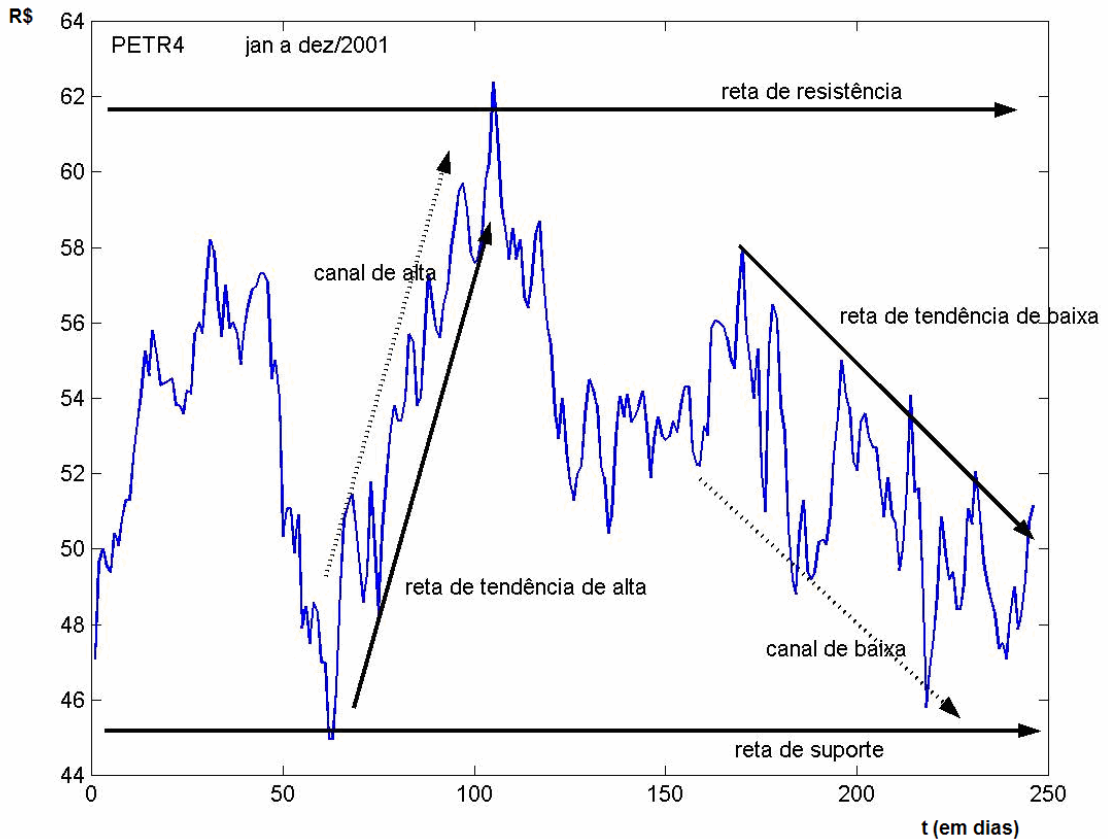


Figura 3.1 – Cotações da PETR4: visualização de retas e canais de tendência

A questão importante que se apresenta é: Como identificar o começo e o fim de uma tendência? A solução deste problema não é uma tarefa fácil, mas a partir da Teoria Dow surgiram vários métodos que buscam respostas. Entre os métodos mais utilizados estão as Ondas de Elliott, que foram criadas por Ralph Nelson Elliott na década de trinta.

A teoria das Ondas de Elliott diz que os movimentos de preços seguem um padrão em que a tendência principal é composta de cinco ondas e a sua correção por três ondas e que, portanto, um ciclo completo será composto por oito ondas (MATSURA, 2007).

Elliott também foi o pioneiro na aplicação dos conceitos matemáticos da Seqüência de Fibonacci no mercado de capitais. Segundo Elliott, as relações entre as ondas seguem as razões de 62% (resultado da divisão de um número qualquer da Seqüência de Fibonacci pelo seu antecessor) e de 38% (divisão de um número de Fibonacci pelo seu segundo antecessor).

MATSURA (2007) e DESCHATRE (1997) ainda apresentam vários tipos de figuras que são construídas a partir de certos padrões existentes nos gráficos de preços. Estes métodos são pouco objetivos e exigem uma grande experiência por parte do analista. Alguns em certas situações, até alcançam bons resultados na identificação de tendências e de momentos de reversão, mas em outros podem resultar em prejuízos.

No momento, o tipo de gráfico mais utilizado pelos analistas técnicos, especialmente, os operadores de curto prazo é o *candlestick*, ou gráfico de vela, (MATSURA, 2007). Este gráfico representa como se comportou o preço ao longo do dia de negociação, ele é formado por um corpo e por sombras superiores e inferiores, conforme se pode observar na figura 3.2.

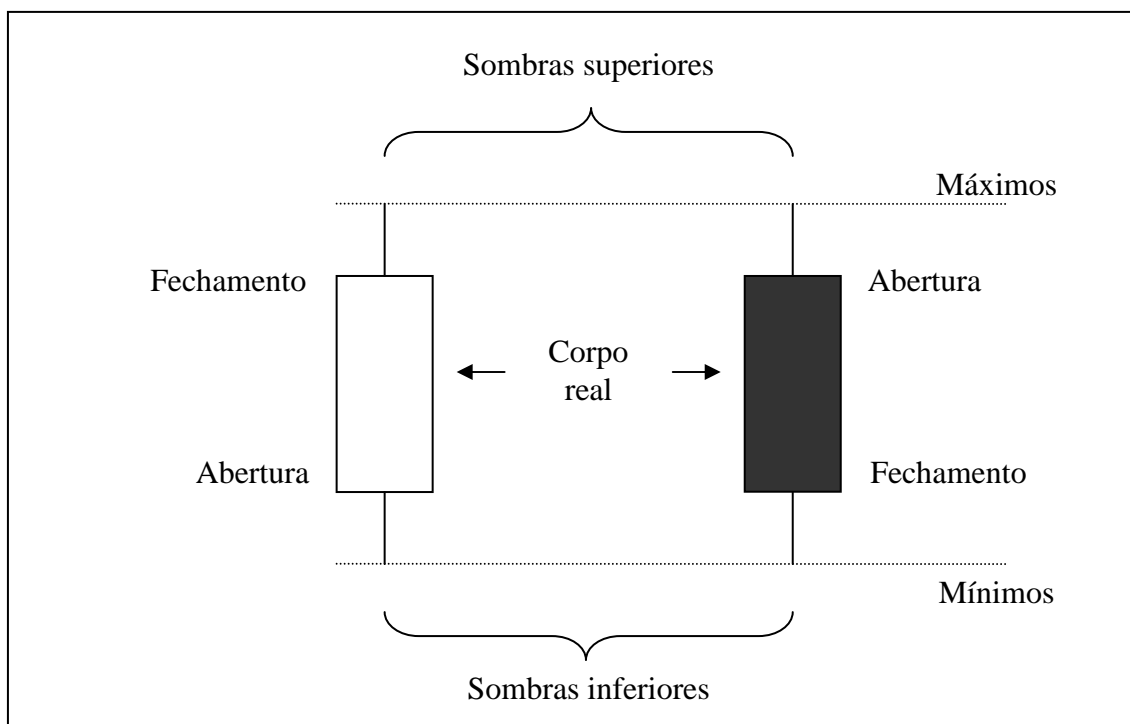


Figura 3.2 – Gráficos de *candlestick* (vela) de alta e baixa, respectivamente.
Fonte: Adaptação de MATSURA (2007).

Onde:

- *Corpo real*: é a região entre a abertura e o fechamento; quando ele é vazado indica que o preço fechou acima do valor de abertura, quando é preenchido indica que ocorreu o contrário.
- *Sombras*: indicam os preços mínimos e máximos pelos quais foi negociada uma ação ao longo de um dia de pregão.

A análise dos gráficos de *candlestick* tem permitido a obtenção de vários padrões que auxiliam na determinação do início e do término de tendências. Porém, apresenta o inconveniente de que as figuras, usadas como indicadores, são caracterizadas de forma subjetiva, não permitindo a utilização de métodos matemáticos de análise e validação.

Já os principais indicadores estatísticos de tendências (ou rastreadores de tendências) são os seguintes:

- *Médias Móveis*: A principal característica de uma série temporal de cotações de ativos financeiros é a sua constante volatilidade. As oscilações sofridas pelos preços dificultam a identificação da verdadeira tendência de evolução das cotações. Para superar estes obstáculos as principais ferramentas utilizadas são as médias móveis. As médias móveis mais comuns são as simples, ponderadas e exponenciais. O cruzamento de médias é um método que vem sendo adotado por muitos profissionais para prever os pontos de início e término de tendências. Este método consiste do uso de gráficos de médias de períodos diferentes (10 e 50 dias, por exemplo), porém possui um grave problema, pois como se tratam de médias sempre há um atraso na identificação dos pontos de reversão das tendências, o que resulta em diminuição dos lucros tanto nas operações de compra como de venda.
- *MACD (Moving Average Convergence Divergence) convergência e divergência de médias exponenciais*: É um método gráfico construído a partir de duas médias móveis exponenciais de períodos diferentes; é o resultado da diferença da média de curto prazo e da de maior período. Segundo Gerald Appel apud MATSURA (2007) os períodos ideais são 12 e 26 dias para construir a curva MACD.

Quando este gráfico passa de negativo para positivo tem-se um indicativo de que está começando uma tendência de alta. Já se passar de positivo para negativo tem-se o início de um período de baixa nos preços. Para corrigir os problemas de falsos sinais de compra e venda produzidos pela volatilidade das cotações foi criada uma média móvel dos próprios valores de MACD, que foi denominada de Sinal (DECHATRE, 1997). Plotando-se estas duas curvas em um mesmo gráfico têm-se um novo indicador formado pelo cruzamento delas. Onde, quando o *Sinal* cruza de baixo para cima, a linha MACD tem um indicativo de venda (começo da tendência de baixa) E, quando passa de cima para baixo tem-se o início de uma tendência de alta.

- *Bandas de Bollinger*: É um indicador criado por John Bollinger, formado por três linhas: uma central que é uma média móvel simples; uma superior que é o resultado da linha central mais duas vezes o seu desvio padrão e uma inferior que é igual a (linha) central menos duas vezes o desvio padrão. O estreitamento desta banda sinaliza que poderá ocorrer uma reversão na tendência.

Conforme MATSURA (2007), estes rastreadores de tendência devem ser usados quando o movimento de preços apresenta uma tendência definida, caso contrário, os seus resultados poderão não ser satisfatórios. Pois, quando o mercado está “de lado” (sem tendência) o ideal é o uso de outros indicadores: os *osciladores* que sinalizam melhor as mudanças que ocorrem nos preços. Entre os principais osciladores estão:

- O Índice de Força Relativo (IFR) calculado de acordo com a equação (3.1):

$$IFR = 100 - \left(\frac{100}{1 + FR} \right) \quad (3.1)$$

Onde: $FR = \frac{\text{média das variações de fechamento em ALTA}}{\text{média das variações de fechamento em BAIXA}}$. Assim, os valores do IFR

variam de zero a 100. Desta forma, quando maior for o *FR*, mais o *IFR* irá se aproximar de 100, indicando que a tendência de compra é superior a de venda. Já a divergência entre o *IFR* e o gráfico de preços será um sinal de que a tendência atual está perdendo força.

- *On Balance Volume* (OBV) ou *Cumulative Volume Indicador* (CVI) que combina a evolução dos preços com o volume negociado (MATSURA, 2007). Ele é calculado da seguinte maneira:

- se o mercado fechou em alta: $OBV_t = OBV_{t-1} + V_t$

- se o mercado fechou em baixa: $OBV_t = OBV_{t-1} - V_t$

- se o mercado fechou estável: $OBV_t = OBV_{t-1}$

Onde:

OBV_t = índice acumulado no dia t .

OBV_{t-1} = índice acumulado no dia $t - 1$.

V_t = volume negociado no dia t .

Segundo DECHATRE (1997), o OBV procura “[...] detectar os processos de acumulação ou distribuição [...]” da Teoria Dow, fazendo, por exemplo, as seguintes relações com o gráfico de preços:

- Fase de Distribuição: preços estáveis e OBV caindo.
- Fase de Acumulação: preços estáveis e OBV aumentando.
- Fase de Euforia: preços em alta, mas o OBV em queda.

Em relação ao gerenciamento dos riscos, na atualidade, a Análise Técnica carece de métodos gráficos que possam atender as demandas dos profissionais. Especialmente, para a determinação dos percentuais de lucro e prejuízos (MATSURA, 2007), os chamados *Stops* de Lucro ou Objetivo - preço que o investidor estipulou como mínimo para a venda do ativo financeiro obtendo lucro sobre o valor de compra - e os *Stops* de Perda ou *Stops Loss* - valor que o investidor determinou como nível máximo de prejuízo. Este é um valor abaixo do preço de compra da ação, mas que impede uma perda que comprometa a capacidade de atuação do investidor no mercado. Uma vez que, deve-se ter em mente que o prejuízo também faz parte de qualquer investimento no mercado de capitais.

Estes foram alguns dos métodos incorporados, ao longo dos anos, à ‘ciência’ da Análise Técnica no mercado de capitais brasileiro. Existem, ainda, muitos outros que, assim como estes, possuem suas virtudes e falhas. De onde se pode concluir

que não existe um método completo que possa ser usado de forma exclusiva. Sendo prática, entre os investidores profissionais, a combinação destes instrumentos de análise, almejando a otimização das suas decisões e, em consequência, o aumento dos lucros durante as operações de compra e venda de ações.

3.1 Considerações

Desta forma, justifica-se a escolha deste domínio por apresentar a oportunidade de desenvolver uma pesquisa que, ao mesmo tempo, pode atrair o interesse do meio acadêmico (especialmente das áreas de redes neurais, processamento de sinais e econometria) e dos profissionais que atuam diretamente na bolsa de valores. Pois, hoje, existe uma expectativa de grande crescimento para o mercado de capitais brasileiro: como fonte de captação de recursos para a expansão das empresas e como alternativa de investimento para pequenos e médios investidores, antes restritos às aplicações em renda fixa e na caderneta de poupança.

4 TRANSFORMADAS *WAVELETS*

Neste capítulo serão apresentados, inicialmente, os aspectos históricos e as principais definições das Transformadas *Wavelets* – TWs. Já, na seção 4.3, serão descritos os conceitos principais da análise de multirresolução e, em seguida, nas seções 4.4 e 4.5, serão detalhados os algoritmos e as equações que são utilizados no processo de multirresolução e de limiarização de sinais discretos.

4.1 Origens

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, a velocidade e pressão de um determinado escoamento, para citar apenas alguns exemplos. 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 ².

² Trata-se de uma função satisfazendo as condições de Dirichlet.
Fonte: <mathworld.wolfram.com/DirichletFourierSeriesConditions.html>

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 foram 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 petrolífera lhe permitiu concluir que manter uma janela³ 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.

Os estudos da belga Ingrid Daubechies (1992) contribuíram para as pesquisas em *wavelets* e o uso de seus recursos na prática. Os estudos de Daubechies iniciaram-se em 1988, juntamente com Stephane Mallat, relacionando-se ao processamento de sinais discretos.

4.2 Definição

As *wavelets* são uma classe de funções que se empregam para localizar uma função dada no espaço e na escala. Uma família de *wavelets* pode construir-se a partir de uma função $\psi(t)$ de quadrado integrável⁴, chamada com frequência

³ Janela: intervalo gráfico de um sinal no qual são analisados o seu período e frequência.

⁴ Isto é uma função tal que $\int_{\mathfrak{R}} |\psi(t)|^2 dt$ seja um número real. Nesse caso se denota $\psi \in L_2(\mathfrak{R})$.

wavelet geradora, ou *wavelet* semente⁵, 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 pela equação (4.1).

$$\psi_{a,b}(t) = \frac{1}{|a|^{1/2}} \psi\left(\frac{t-b}{a}\right), \text{ com } a, b \in \mathbb{R}, a \neq 0, \quad (4.1)$$

Onde: o termo $\frac{1}{|a|^{1/2}}$ é um fator de normalização da energia do sinal.

Exemplo 1. Considere-se a seguinte função $\psi(t) = \chi_{[0,1)}(2t) - \chi_{[0,1)}(2t-1)$, sendo $\chi_{[0,1)}$ a função característica do intervalo $[0,1)$, conforme gráfico da figura 4.1 que apresenta a função geradora das *wavelets* de Haar.

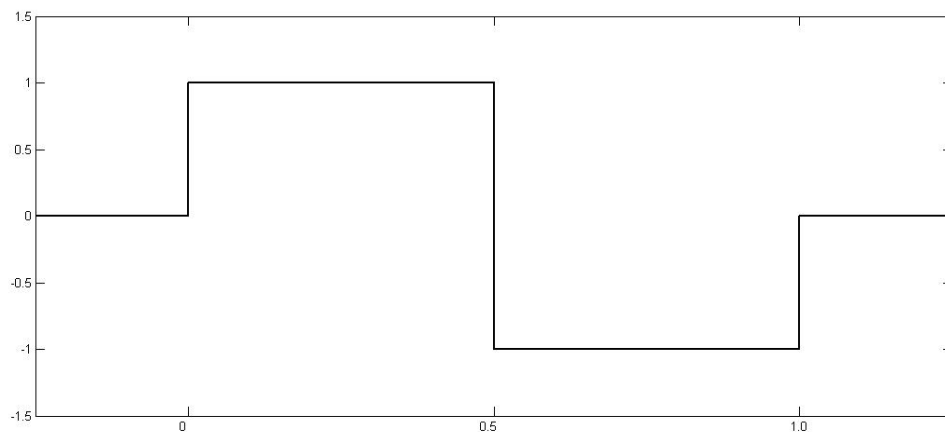


Figura 4.1 - Função geradora das *wavelets* Haar

As *wavelets* Haar $\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$, com $j, k \in \mathbb{Z}$, formam uma base ortogonal para $L_2(\mathbb{R})$. Para $k = 0$, considere-se a família $\{\psi_{j,0} \mid j \in \mathbb{Z}\}$; quanto maior é j , menor será o conjunto cozero⁶ $[0, 1/2^j]$.

⁵ Na língua inglesa se diz *mother wavelet*.

⁶ O conjunto cozero de uma função é o conjunto de todos os valores x tais que o valor $f(x)$ é diferente de zero. Simbolicamente representa-se: $\{x \mid f(x) \neq 0\}$.

É importante ressaltar que para calcular as TWDs (de forma computacionalmente eficiente) é preciso fazer algumas restrições na equação (4.1): onde deve-se ter $a = 2^m$ e $b = n \cdot 2^m$ - sendo n (índice de translação) igual ao número de amostras do sinal. Este processo resulta em uma estrutura de escalas e translações denominada de “diádica”, em que potências de 2 estão relacionadas com intervalos (oitavas) e durações das notas, em um procedimento semelhante às notações musicais. (GALVÃO et al, 2001).

Salienta-se, ainda, que para um cálculo adequado da TWD de um sinal, utilizando a *wavelet* de Haar, é preciso que este possua uma quantidade de amostras igual a uma potência de 2, conforme ilustrado no exemplo 2, onde o sinal discreto s_k possui 64 pontos.

A correspondente transformada discreta Haar é definida da seguinte forma sobre um sinal discreto $s = [s_1 \ s_2 \ \dots \ s_n]$: geram-se dois sub-sinais de comprimento $\frac{n}{2}$, a_1 e d_1 da seguinte maneira:

$$a_1 = \left[\frac{s_1 + s_2}{\sqrt{2}} \quad \frac{s_3 + s_4}{\sqrt{2}} \quad \dots \quad \frac{s_{n-1} + s_n}{\sqrt{2}} \right]$$

$$d_1 = \left[\frac{s_1 - s_2}{\sqrt{2}} \quad \frac{s_3 - s_4}{\sqrt{2}} \quad \dots \quad \frac{s_{n-1} - s_n}{\sqrt{2}} \right]$$

Exemplo 2. Dada a função $f : [0,1] \rightarrow \mathfrak{R}$ definida por $f(t) = 4t(1-t)$, define-se por amostragem ($n = 64$) o correspondente sinal discreto $s_k = 4 \frac{k}{64} \left(1 - \frac{k}{64}\right)$, com $k = 0, 1, \dots, 64$. Na figura 4.2 é mostrado o gráfico do sinal da função quadrática $f(t)$ tomado por amostragem.

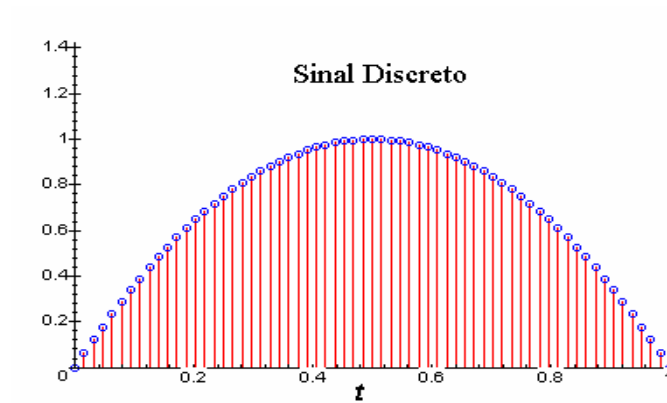


Figura 4.2(a) – Gráfico do sinal discreto s_k

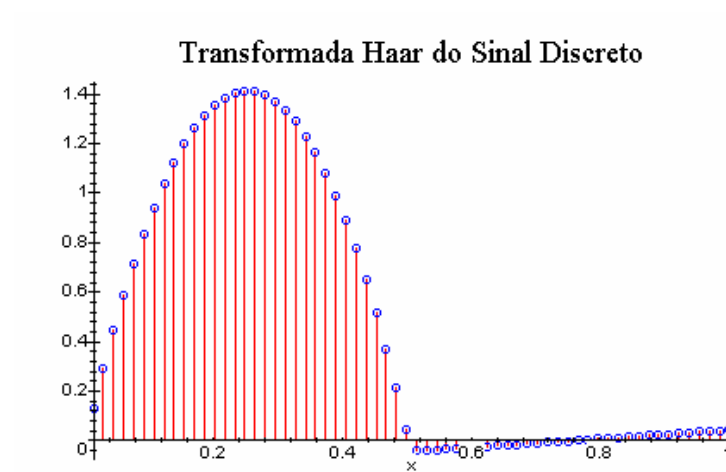
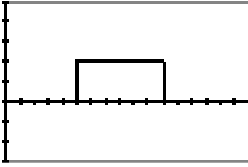
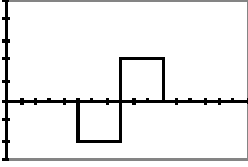
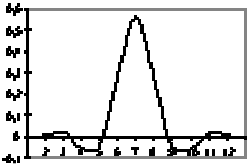
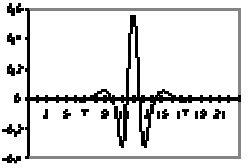
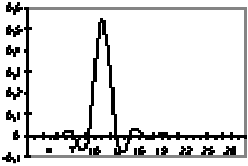
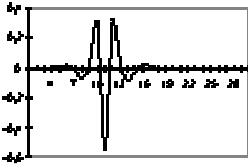
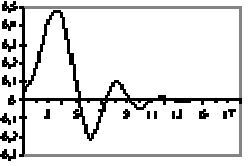
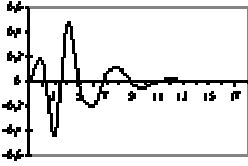
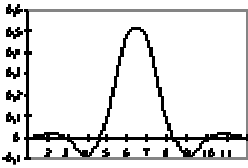
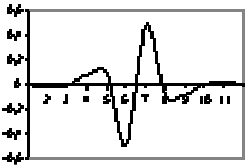
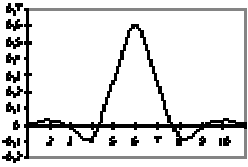
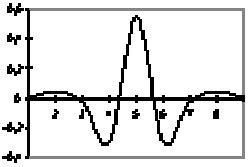


Figura 4.2(b) - Efeito da transformação discreta *wavelet* Haar sobre um sinal

O efeito gráfico é de compactação principal do sinal original no sub-sinal da esquerda. Na verdade, o que acontece é que a energia se acumula no sub-sinal a_1 .

A seguir, a Tabela 4.1 apresenta as funções de filtragem que podem ser utilizadas:

Tabela 4.1 – Principais funções de filtragem (*wavelets* geradoras)

Classe de filtro	Função geradora	Wavelet	Descrição
Haar			A classe de filtro mais simples, filtro que utiliza somas e diferenças médias.
Quase Coiflet			
Coiflet			Classe de <i>wavelet</i> da 1ª geração, desenvolvida por R.Coifman
Daubechies			<i>Wavelet</i> não simétrica, estritamente ortogonal, estritamente com próprio suporte compacto.
Johnston-Barnard			
Spline bi-ortogonal			Simétrica, se calcula a partir dos coeficientes binomiais.

Fonte: <www.fh-jena.de/contrib/fb/et/personal/ansorg/ftp/wavelet/wavelet.htm>

4.3 Análise de Multirresolução - AMR

A AMR foi criada por Stéphane Mallat em 1989 e consiste de um processo pelo qual um sinal é analisado em relação às baixas e às altas frequências que o compõe. É resultado da combinação de uma função de escala $\phi(t)$ e de uma função

wavelet $\psi(t)$ (DE OLIVEIRA, 2007). Através desta abordagem é possível extrair informações importantes de um sinal, como por exemplo, identificar as tendências de longo prazo e localizar os momentos de oscilações relevantes.

Conforme DE OLIVEIRA (2007, p. 82) pode-se definir uma AMR como uma seqüência de subespaços fechados V_m onde $V_m \subset L_2(\mathfrak{R})$ e $m \in Z$, cujo objetivo é decompor um sinal em pedaços. Cada um dos subespaços V_m contém uma projeção (parte) do sinal original.

Uma análise de multirresolução em $L_2(\mathfrak{R})$ deve satisfazer as cinco propriedades a seguir apresentadas (DE OLIVEIRA, 2007):

i) $V_m \subset V_{m-1}$ ($\forall m$): um subespaço na escala de resolução m sempre estará contido em um de resolução menor: $V_m \dots \subset V_3 \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \dots$

ii) $f(t) \in V_m \subset L_2(\mathfrak{R}) \Leftrightarrow f(2t) \in V_{m-1}$: invariância em escala.

iii) $\text{clos} \bigcup_{m \in Z} V_m = L_2(\mathfrak{R})$: a união de todos os subespaços é igual ao espaço de sinal de energia finita, em outras palavras, a soma dos subespaços V_m (vetores que contêm os sinais de aproximação e de detalhes do sinal original) é igual a um espaço (vetor) de energia finita $L_2(\mathfrak{R})$.

iv) $\bigcap_{m \in Z} V_m = \{0\}$: “A função toda nula é o único sinal de $L_2(\mathfrak{R})$ que pode ser representado em qualquer escala V_m ” (DE OLIVEIRA, 2007, p. 83).

v) $\exists \phi(t) \in V_0$ tal que $\{\phi(t-n)\}_{n \in Z}$ é uma base ortogonal para V_0 e que a função de

escala $\left\{ \phi_{m,n}(t) = 2^{-\frac{m}{2}} \phi(2^{-m}t - n) \right\}_{m,n \in Z}$ é uma base de V_m .

Na prática a AMR procura representar um sinal como uma soma de aproximações e detalhes, resultado da aplicação de dois filtros: um passa-baixa que é uma aproximação grosseira do sinal e outro denominado de passa-alta que captura os detalhes do sinal. Nas equações 4.2 e 4.3 têm-se uma forma de obter os coeficientes de aproximação (ou de escala) e os de detalhes (ou *wavelets*) a partir de uma filtragem discreta (DE OLIVEIRA, 2007):

$$a_{m+1} = c_{m+1}[n] = \sum_{k \in Z} h[k - 2n] c_m[k] \quad (a_{m+1} \text{ é o sinal suavizado}) \quad (4.2)$$

$$d_{m+1} = d_{m+1}[n] = \sum_{k \in Z} g[k - 2n] c_m[k] \quad (d_{m+1} \text{ é o sinal de detalhe}) \quad (4.3)$$

Ainda, é preciso destacar que entre as condições necessárias para a análise de multirresolução, no que se refere aos coeficientes dos filtros discretos, estão as seguintes:

$$\text{i) } \sum_{n \in Z} h_n = \sqrt{2} \quad (\text{condição passa-baixa})$$

$$\text{ii) } \sum_{n \in Z} g_n = 0 \quad (\text{condição passa-alta})$$

A partir dos coeficientes de aproximação e de detalhes é possível reconstruir integralmente o sinal original, conforme a equação 4.4:

$$S = a_m \oplus d_m \oplus d_{m-1} \oplus \dots \oplus d_1 \quad (4.4)$$

Onde:

S = sinal original

m = nível de decomposição

\oplus = indica uma soma ortogonal

Nas figuras 4.4 e 4.5(a e b), respectivamente, é possível observar como agem os filtros de aproximação (h) e de detalhes (g) no sinal apresentado na figura 4.3. Utilizou-se a função *wavelet* geradora (ou *mother-wavelet*) de Haar.

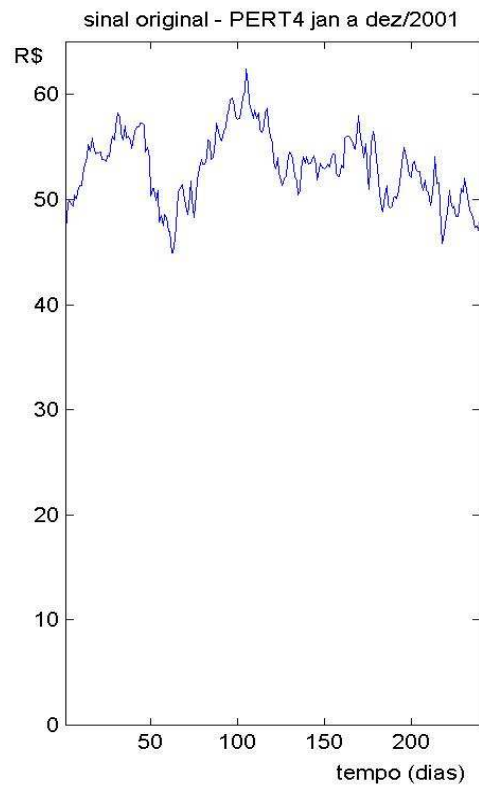
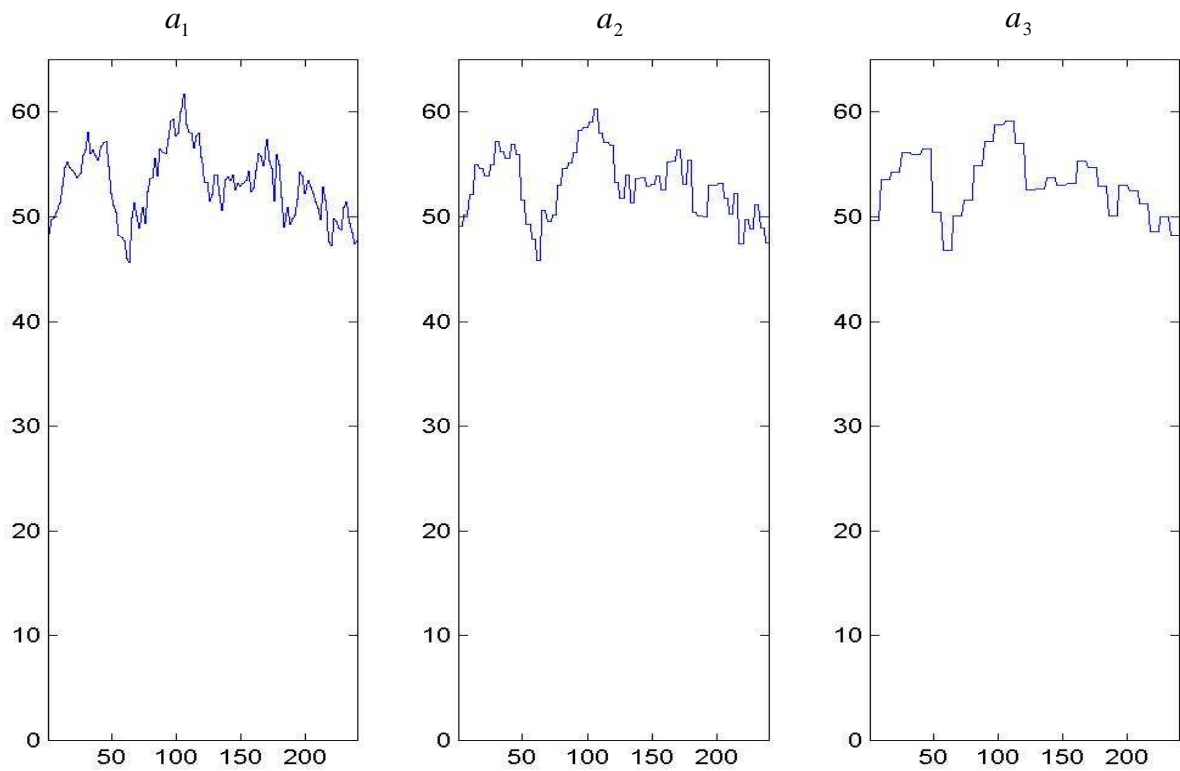


Figura 4.3 – Sinal original (Ações da PETR4)

Figura 4.4 – Sinal aproximado nos níveis de decomposição $m = 1, 2$ e 3

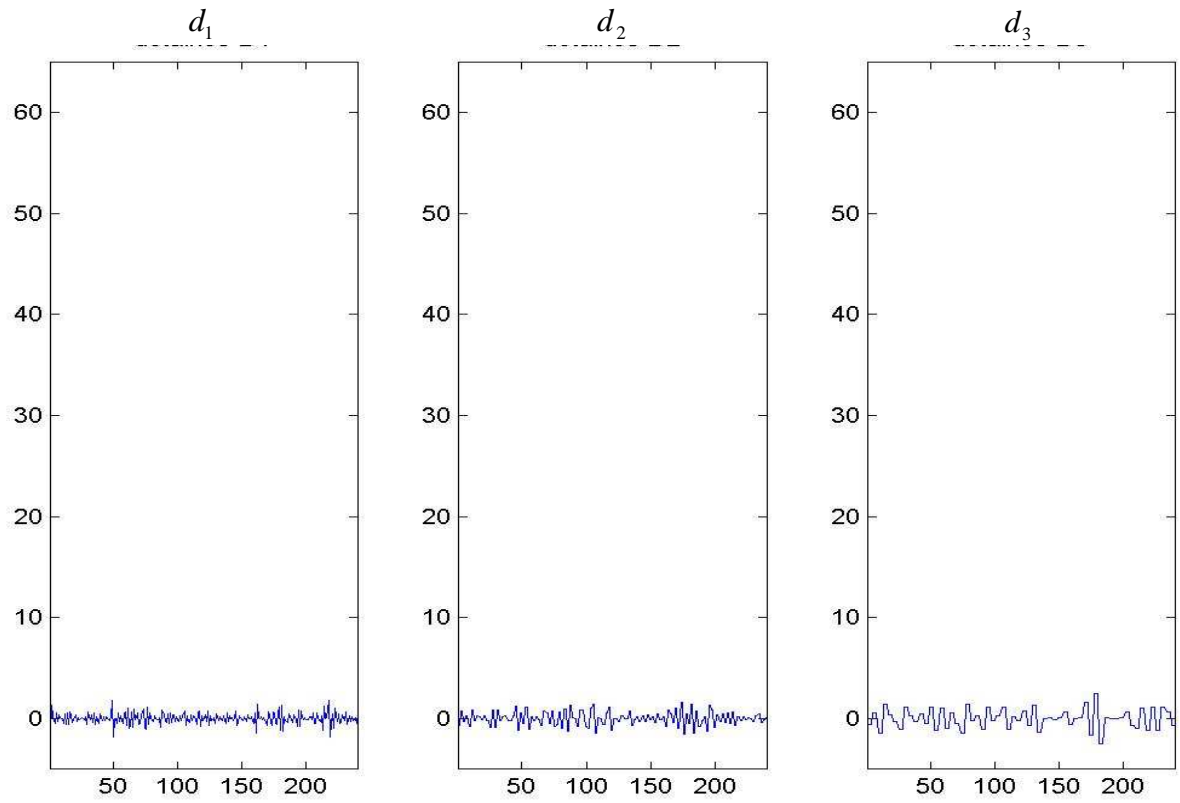


Figura 4.5 (a) – Sinal de detalhes nos níveis de decomposição $m = 1, 2$ e 3

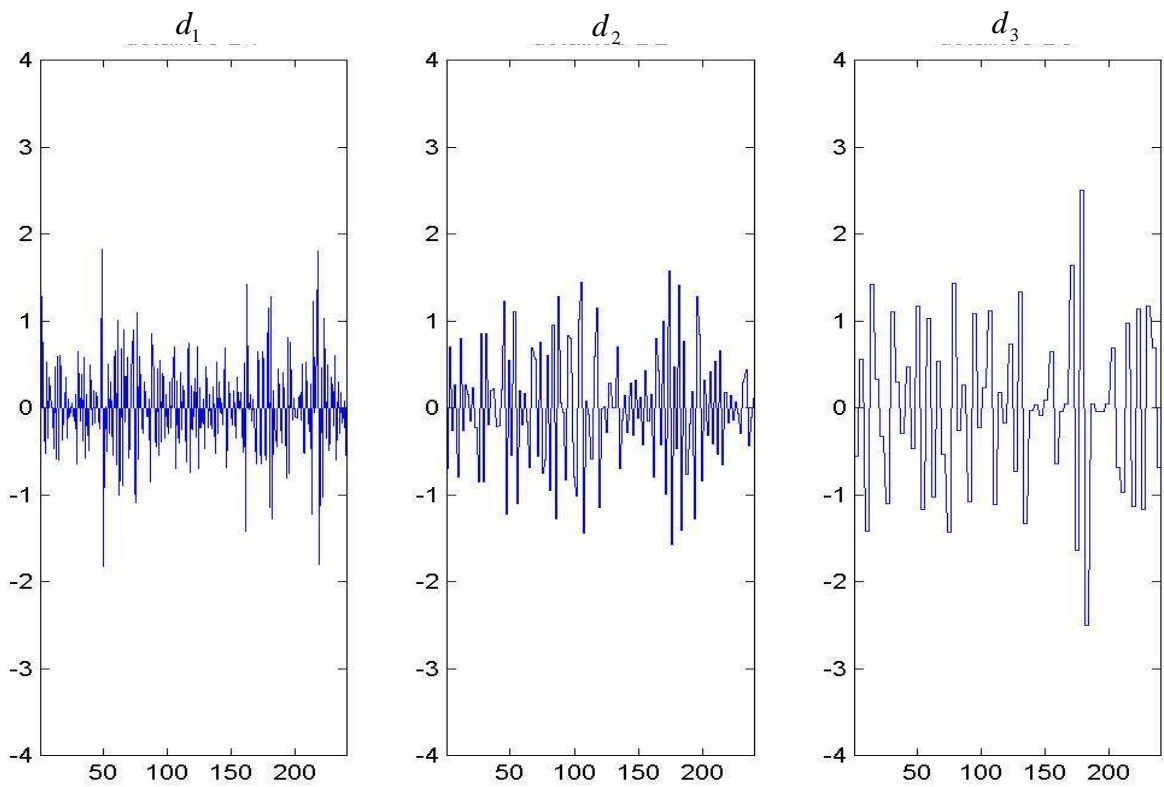


Figura 4.5 (b) – Sinal de detalhes “ampliados”

4.4 Implementação das Transformadas *Wavelets* Discretas – TWDs

Uma forma prática de realizar a AMR, segundo os autores REIS & DA SILVA (2004), aplicando as Transformadas *Wavelet* Discretas (TWDs) é através do algoritmo piramidal de *Mallat*, que consiste em dividir o sinal original em dois: um é o sinal original suavizado e outro uma ampliação das oscilações ou “ruídos” do sinal analisado. Realiza-se, assim, uma análise de multirresolução: encontrando uma linha de tendência para períodos longos e fazendo um estudo em detalhes dos períodos curtos. Na figura 4.6, tem-se um esquema do algoritmo piramidal de *Mallat*, onde, N é o número de pontos, S_0 é o sinal original, m é o nível de escala (cada vez que se aplica a TWD no sinal $a: m = m + 1$), a_m é o sinal suavizado ou tendência geral no nível de escala m e d_m são os “detalhes” do sinal.

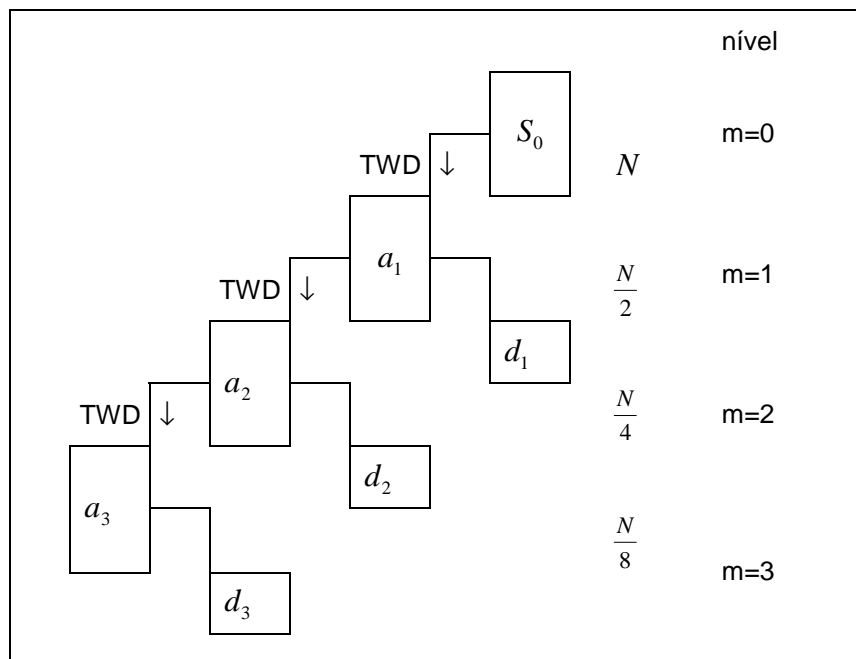


Figura 4.6 – Esquema do algoritmo de Mallat
Fonte: Adaptado de REIS & DA SILVA (2004)

Aplicando-se o algoritmo de Mallat para calcular a TWD e usando-se a função Haar como *wavelet-mãe*, tem-se o apresentado na equação (4.5).

$$S_0 \cdot xW = TWD \rightarrow TWD = a_m + d_m \quad (4.5)$$

onde S_0 = matriz com o sinal original (ou suavizado no nível $m-1$) e W = matriz dos coeficientes *wavelets*. Na figura 4.7, tem-se um esquema da multiplicação matricial que pode ser implementada no *software Matlab*[®] para realizar a análise de multirresolução usando a Transformada de Haar.

S_0	W	TWD
$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \\ \cdot & \cdot \\ k_{n-1} & k_n \end{bmatrix}$	$\cdot x \begin{bmatrix} c_0 & c_0 \\ c_1 & -c_1 \end{bmatrix}$	$= \begin{bmatrix} a_m & d_m \\ a_m & d_m \\ \cdot & \cdot \\ a_m & d_m \end{bmatrix}$

Figura 4.7 - Demonstração da operação matricial descrita na equação (4.5)

Onde: $c_0 = \frac{\sqrt{2}}{2}$ e $c_1 = \frac{\sqrt{2}}{2}$ (coeficientes *wavelets* de Haar). Para restaurar o sinal original basta multiplicar a matriz TWD pela matriz inversa de W que é igual à transposta (W^T). Visto que, trata-se de uma matriz ortogonal, onde sua matriz inversa coincide com a transposta, isso pelo fato que cada filtro é de norma unitária. Por exemplo, para o filtro passa-baixa de Haar a norma é apresentada na equação (4.6).

$$Norma = \sqrt{\left(\frac{\sqrt{2}}{2}\right)^2 + \left(\frac{\sqrt{2}}{2}\right)^2} = 1 \quad (4.6)$$

E também para o passa-alta, a norma é apresentada na equação (4.7).

$$Norma = \sqrt{\left(\frac{\sqrt{2}}{2}\right)^2 + \left(-\frac{\sqrt{2}}{2}\right)^2} = 1 \quad (4.7)$$

Uma outra forma de implementar, não só a TWD de Haar, mas outras, como as *wavelets* de Daubechies com n coeficientes é a apresentada na equação (4.8) (adaptado de: DE SOUZA, 2004).

$$WxS_0 = TWD \quad (4.8)$$

W								S_0	TWD
c_0	c_1	0	0	0	0	0	0	k_1	a_m
0	0	c_0	c_1	0	0	0	0	k_2	a_m
0	0	0	0	c_0	c_1	0	0	k_3	a_m
0	0	0	0	0	0	c_0	c_1	k_4	a_m
c_0	$-c_1$	0	0	0	0	0	0	k_5	d_m
0	0	c_0	$-c_1$	0	0	0	0	k_6	d_m
0	0	0	0	c_0	$-c_1$	0	0	k_7	d_m
0	0	0	0	0	0	c_0	$-c_1$	k_8	d_m

Figura 4.8 - Demonstração da operação matricial descrita na equação (4.8)

Os coeficientes das linhas ímpares correspondem ao filtro passa-baixa que resulta no sinal suavizado (S_{m+1}) e os das linhas pares correspondem ao filtro passa-alta responsável pelo sinal de detalhe (d_{m+1}) (DE SOUZA, 2004). A restauração do sinal também é realizada através da transposta da matriz W da figura 4.8.

4.5 Remoção de Ruído (*denoising*) com as TWDs

Conforme foi destacado as TWDs permitem que se faça uma redução (*denoising*) do ruído presente no conjunto de dados. Esta filtragem é realizada através de um processo denominado de limiarização (*thresholding*) que consiste na eliminação ou diminuição da magnitude de alguns coeficientes *wavelets* (detalhes d).

Parte-se do princípio que junto com os detalhes do sinal, o vetor d contenha também o ruído aleatório intrínseco da série de dados em análise. No caso das cotações dos ativos financeiros acredita-se que os detalhes do sinal sejam frutos de dois “fenômenos” distintos: as variações aleatórias nos preços, característica dos mercados de capitais (ruído), e de eventos oscilatórios que retratam algum acontecimento relevante que afeta os valores das ações.

Em termos matemáticos, pode-se modelar este fenômeno com a equação (4.9):

$$S_k = \widehat{S}_k + \sigma n_k \quad k = 1, 2, \dots, N \quad (4.9)$$

Onde:

S_k = sinal contaminado por um ruído aditivo.

\widehat{S}_k = sinal sem ruído.

σ = estimativa do ruído.

n_k = são variáveis aleatórias de um ruído branco gaussiano⁷.

Considera-se que S_k é o resultado da soma do sinal original (sem ruído) mais um componente de ruído aditivo (DE OLIVEIRA, 2007).

Aplicando-se a TWD sobre o sinal S_k pode-se obter o sinal \widehat{S}_k (sinal filtrado) pela equação (4.10):

$$\widehat{S}_k = S_m \oplus d_m^T \oplus d_{m-1}^T \oplus \dots \oplus d_1^T \quad (4.10)$$

Onde:

m = nível de decomposição.

⁷ Ruído aleatório que segue uma distribuição (estatística) gaussiana.

d_m^T = são os coeficientes de detalhes (ou coeficientes *wavelets*) no nível de decomposição m após o processo de limiarização.

Os coeficientes d_m^T são obtidos após o processo de limiarização (*thresholding*) que pode ser realizado por dois métodos básicos:

- O limiar rígido (*threshold-hard*) onde todos os coeficientes que, em módulo, são menores que um certo limiar T são eliminados (zerados) e os outros coeficientes mantidos com o mesmo valor, conforme equação (4.11).

$$L_T^R(d_k) = \begin{cases} 0, & \text{se } |d_k| < T \\ d_k, & \text{se } |d_k| \geq T \end{cases} \quad (4.11)$$

Onde $L_T^R(d_k)$ é o limiar rígido aplicado sobre os coeficientes *wavelets*.

- O limiar suave (*threshold soft*) que assim como no limiar rígido todos os coeficientes menores que T , em módulo, são zerados, porém os demais são reduzidos do valor T , conforme a equação 4.12 (DE SOUZA, 2004).

$$L_T^S(d_k) = \text{sgn}(d_k)(|d_k| - T) = \begin{cases} d_k - T, & \text{se } d_k > T \\ 0, & \text{se } |d_k| \leq T \\ d_k + T, & \text{se } d_k < -T \end{cases} \quad (4.12)$$

Onde $T \in \mathfrak{R}^+$ e

$$\text{sgn}(d_k) = \begin{cases} 1, & \text{se } d_k > T \\ 0, & \text{se } |d_k| \leq T \\ -1, & \text{se } d_k < -T \end{cases} \quad (4.13)$$

Sendo que $L_T^S(d_k)$ é o limiar suave aplicado sobre os coeficientes *wavelets*.

A definição do parâmetro T é de grande importância para o sucesso do processo de filtragem. Na Tabela 4.2 são apresentados os principais tipos de

parâmetros de limiarização, conforme GENÇAY, SELÇUK & WHITCHER (2002) e GUTIERREZ (2002).

Tabela 4.2 – Principais parâmetros de limiarização

Tipo de Limiar	Descrição*	Função no <i>Matlab</i>**
Universal	Parâmetro global – não depende do nível da TWD	'sqtwolog'
Minimax	Parâmetro global – não depende do nível da TWD	'minimaxi'
SURE (<i>Estimation Stein's Unbiased Risk Estimate</i>)	Estimador não viesado do risco de (SURE) – depende do nível da TWD	'rigrsure'
SURE - híbrido	Parâmetro híbrido – usa-se um teste de esparsidade em cada nível da TWD para escolher entre o limiar Universal ou o SURE	'heursure'

* A descrição completa de cada tipo de limiar pode ser obtida em GENÇAY, SELÇUK & WHITCHER (2002) e GUTIERREZ (2002).

** Função do *Matlab*[®] utilizada para obter o valor do limiar a partir de uma série temporal.

4.6 Considerações

Neste capítulo, foram apresentados os fundamentos principais do processo de redução de ruído de um conjunto de dados com o objetivo de otimizar o método de predição de séries temporais via RNAs que será detalhado no capítulo 5, a seguir.

5 REDES NEURAIIS ARTIFICIAIS

Este capítulo aborda os principais conceitos das Redes Neurais Artificiais (RNAs), tais como: o modelo de neurônio artificial, suas funções matemáticas de ativação e propagação, as topologias mais comuns e os métodos de aprendizagem utilizados para treiná-las. Também, apresenta o modelo Perceptron de Múltiplas Camadas (*Multi-Layer Perceptron* – MLP) e seu algoritmo de treinamento *backpropagation*.

5.1 Definição e Histórico

Pode-se definir rede neural artificial como sendo um modelo matemático implementado através de programas computacionais, que procura imitar o funcionamento do cérebro humano, no que se refere à capacidade de aprender e generalizar (BRAGA, CARVALHO & LUDEMIR, 2000). Possui como elementos os neurônios artificiais que se comunicam entre si de forma não linear, recebendo os “estímulos” de diversos outros, constituindo-se em um sistema de processamento paralelo.

Suas principais funções são o aprendizado e o reconhecimento de padrões, tais como imagens, sons e caracteres, mesmo que estes apresentem ruídos (falta ou distorção de informações).

Os primeiros modelos neurais foram criados na década de 40 pelo psiquiatra McCulloch e o matemático Pitts que realizaram um trabalho de unificação da neurofisiologia e da lógica matemática. A partir disso, uma série de trabalhos de

diferentes pesquisadores de várias áreas do conhecimento contribuíram para o desenvolvimento das RNAs. Em HAYKIN (2001) encontra-se um detalhado histórico desta evolução, do qual se podem destacar alguns momentos:

- Em 1949, Donald Hebb apresenta a formulação de uma regra de aprendizagem para a modificação dos pesos sinápticos. Este trabalho serviu de inspiração para a criação de sistemas adaptativos e de aprendizagem (HAYKIN, 2001).
- Em 1959, paralelamente, foram desenvolvidos dois trabalhos importantes: Bernard Wildrow criou o modelo ADALINE que foi utilizado na primeira aplicação prática de um sistema neural (THOMÉ, 2007) e o modelo Perceptron, criado por Franck Rosembat.
- Nas décadas de 60 e 70, o estudo das RNAs passou por um período de “trevas”. As sérias limitações conceituais do Perceptron e a falta de recursos computacionais eficientes desestimularam os investimentos.
- Somente na década de 80, com o desenvolvimento da informática (surgimento de processadores mais baratos e velozes) e com os trabalhos de John Hopfield (1982) - que introduziu pela primeira vez a idéia de função de energia aos sistemas neurais – e a criação do algoritmo de aprendizagem *backpropagation*, em 1986, pelos pesquisadores Janes McClelland e David Rumelhart; que foi retomado o interesse pelo desenvolvimento e aplicação das RNAs nas mais diversas áreas do conhecimento, passando elas a ser objeto de pesquisas multidisciplinares.

5.2 O Neurônio Artificial

Ao longo do desenvolvimento da teoria das redes neurais sempre se buscou criar um neurônio artificial que se aproximasse o máximo possível do humano, que tivesse a capacidade de se comunicar com muitos outros ao mesmo tempo, assim como acontece com os neurônios do cérebro humano. Com este objetivo, realizou-se uma analogia onde: os dendritos (do modelo biológico) são os terminais de entrada do modelo matemático, o axônio é o terminal de saída ou resposta da RNA e as sinapses são “*emuladas*” (BRAGA, CARVALHO & LUDEMIR, 2000, pg. 08) através da associação de pesos às conexões.

Os pesos são valores numéricos aplicados a cada uma das conexões da rede e representam o conhecimento adquirido pelo sistema (são a “memória” das RNAs). Eles têm como função ponderar os sinais de entrada de cada um dos neurônios, a fim de ajustar a importância de cada conexão para o aprendizado de um determinado padrão (BRAGA, CARVALHO & LUDEMIR, 2000).

Nas figuras 5.1(a) e 5.1(b) pode-se verificar uma comparação entre o modelo simplificado de um neurônio biológico e o artificial utilizado nas RNAs.

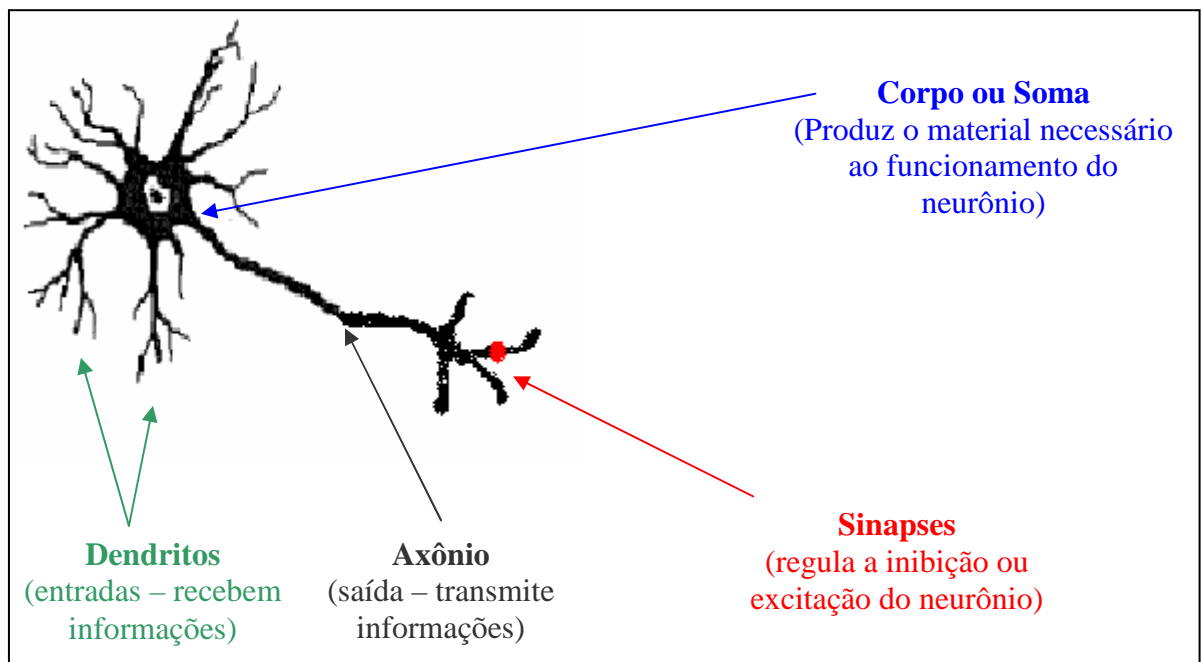


Figura 5.1(a) – Modelo de um neurônio biológico
Fonte: Adaptação de FALCÃO (2003) apud FREIMAN & PAMPLONA (2005).

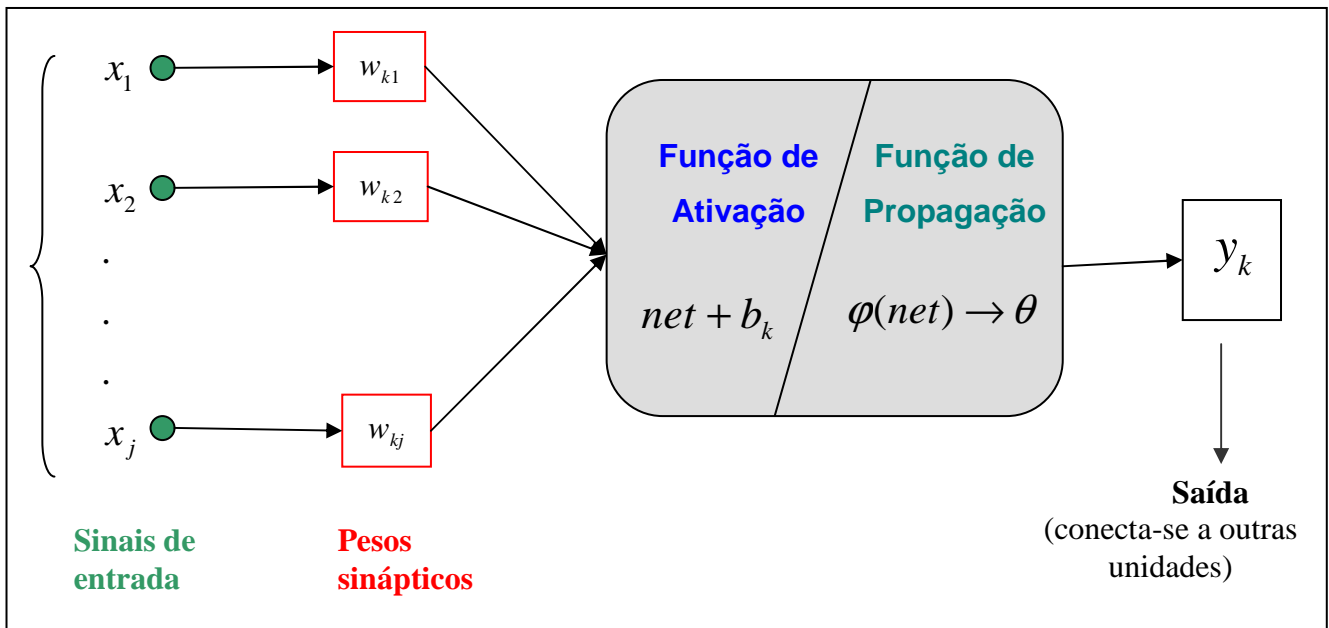


Figura 5.1(b) – Modelo matemático de um neurônio artificial
 Fonte: Adaptação de HAYKIN (2001).

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.
- θ (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.

Com base nos esquemas apresentados (figuras 5.1(a) e 5.1(b)) pode-se traçar um paralelo entre o neurônio humano e o artificial, uma vez que os sinais de entrada x_j do modelo artificial são uma versão dos sinais de estímulos recebidos pelos dendritos do modelo natural. As sinapses, responsáveis por regular os estímulos de saída do neurônio humano, são modeladas através dos pesos

sinápticos que aplicados sobre os sinais de entrada intensificam ou atenuam estes sinais. As funções de ativação e propagação, que são uma espécie de núcleo de processamento do neurônio artificial, regulam o valor final obtido para a saída (OSÓRIO & BITTENCOURT, 2000). Já a saída y_k equipara-se ao axônio que é responsável pela transmissão de estímulos para outros neurônios.

Na seção seguinte, são detalhadas as características das funções de ativação e propagação que juntas formam a unidade de processamento do neurônio artificial.

5.3 Funções de Ativação e Propagação

Conforme THOMÉ (2007, pg. 21): “A função de ativação é aplicada sobre os sinais de entrada e gera uma saída intermediária, normalmente chamada de *net* [...]”. Esta função executa um processamento sobre estes estímulos ponderados pelos respectivos pesos (w_i), sendo responsável pelo sinal de excitação ou inibição emitido pelo neurônio para os que estão imediatamente a sua frente. A principal função de ativação presente em 99% das redes (THOMÉ, 2007) é a apresentada na equação (5.1).

$$a_i = (x, w) = w^T x = \sum_{j=1}^n w_{ij} x_j \quad (5.1)$$

Trata-se de uma função linear, resultado do somatório dos sinais de entrada multiplicados pelos seus respectivos pesos sinápticos, sendo w^T a matriz transposta dos pesos.

Já a função de propagação ou limiar lógico (θ), é aplicada sobre o estado de ativação (*net*), gerando um sinal de saída para o neurônio, sendo, portanto, responsável pela transmissão dos estímulos do estado de ativação para os outros neurônios (THOMÉ, 2007). Em relação aos tipos de funções usadas, existe um amplo leque de opções, cuja escolha depende, dentre outros fatores, do algoritmo de aprendizagem utilizado e da topologia da rede, os quais devem estar sempre

atrelados à aplicação a que se destinam. A Tabela 5.1 apresenta as principais funções utilizadas.

Tabela 5.1 – Principais funções de propagação usadas nas RNAs

Funções	Equações	Características/ Aplicações
Linear	$y = \alpha x$	x é o valor de entrada e α é um número real que define a saída linear (y) de x . Utilizada, normalmente, em problemas que sejam linearmente separáveis.
Rampa	$y = \begin{cases} +b, & x \geq +b \\ x, & x < +b \\ -b, & x \leq -b \end{cases}$	é uma derivação da função linear, onde os valores de saída são restritos a uma faixa constante $[-b, b]$.
Degrau (passo)	$y = \begin{cases} +b, & x > 0 \\ -b, & x \leq 0 \end{cases}$	tem-se apenas dois valores possíveis para a saída do neurônio $[-b, b]$. Normalmente, é aplicada em problemas de classificação binária (apenas dois estados possíveis). Não é uma função diferenciável, pois não é contínua no ponto $x = 0$.
Logística Sigmoidal	$y = \frac{1}{1 + e^{\frac{-x}{T}}}$	é utilizada em redes de múltiplas camadas, sendo particularmente eficiente para a solução de problemas não-lineares. O parâmetro T determina a suavidade da curva. Quanto maior o valor de T menor é taxa de crescimento da função em direção a sua assíntota (+1), ou seja, mais suave é o gráfico da função.
Tangente Sigmoidal	$y = \frac{1 - e^{\frac{-x}{T}}}{1 + e^{\frac{-x}{T}}}$	é um modelo originado da função logística sigmoidal, com a diferença que os valores de saída são simétricos e restritos ao intervalo $(-1, +1)$.

A composição das funções de ativação e propagação constitui a função de transferência que realiza o mapeamento das características/ padrões existentes nos sinais de entrada.

A função logística sigmoidal, que também é chamada de *S-shape*, é uma função limitada, monótona e diferenciável em qualquer ponto (BRAGA, CARVALHO

& LUDEMIR, 2000). Ela é utilizada na maioria dos modelos neurais, pois tem duas características importantes: permite que a entrada possua qualquer valor $(-\infty, +\infty)$ e é capaz de converter os valores de saída para o intervalo $[0, +1]$ (THOMÉ, 2007).

5.4 Principais Características e Aplicações das RNAs

Uma RNA é um sistema de processamento de informações paralelo e distribuído (várias unidades de processamento – neurônios – atuando em conjunto, e o padrão de entrada é distribuído nos vários neurônios da RNA), que apesar de ser constituído de unidades (neurônios) relativamente simples torna-se uma ferramenta complexa e poderosa, modelando problemas a partir das interações de suas unidades de processamento. Sua estrutura e funcionamento são inspirados no cérebro humano, especialmente na capacidade de generalização, visto que o conhecimento adquirido e armazenado durante a fase de treinamento pode ser extrapolado para um conjunto de dados diferente, mas semelhante ao aprendido pela RNA.

Segundo os autores (BRAGA, CARVALHO & LUDEMIR, 2003, p. 145): “As RNAs são capazes de resolver, basicamente, problemas de aproximação, predição, classificação, categorização e otimização.” Muitas áreas do conhecimento têm aplicado as RNAs para a solução de questões que envolvam pelo menos um desses problemas. Algumas aplicações são: classificação de padrões utilizada na análise de crédito (gestão de riscos) e de sinais biomédicos (medicina); predição de valores futuros de séries temporais (finanças); modelagem de processos e controle (indústria, eletrônica e automação), reconhecimento de caracteres (processamento de imagens), entre outras.

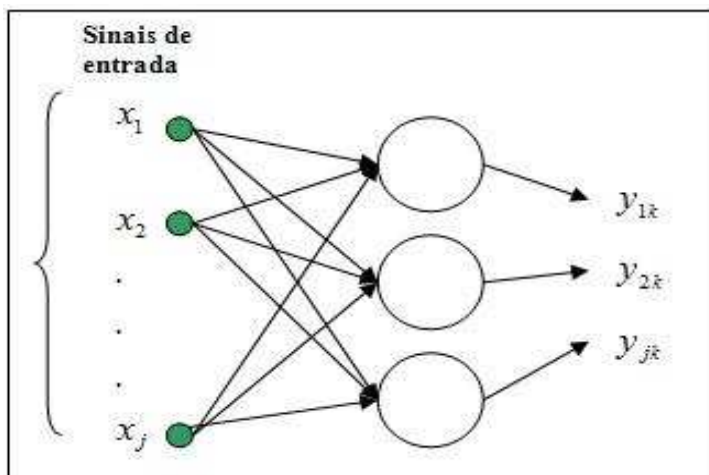
Quanto à arquitetura (ou topologia) as redes neurais podem ser classificadas segundo os seguintes critérios (BRAGA, CARVALHO & LUDEMIR, 2000):

- *Quanto ao número de camadas:* Podem ser formadas por uma única camada, como se pode observar na figura 5.2(a) ou de múltiplas camadas, como visto na figura 5.2(b). Redes multicamadas possuem pelos menos duas camadas de

neurônios entre os sinais de entrada e de saída. Quando existirem uma ou mais camadas entre a de entrada e saída, estas serão denominadas de camadas ocultas ou intermediárias.

- *Quanto aos tipos de conexões:* Podem ser acíclicas (*feedforward*⁸) onde a saída de um neurônio só pode ser usada como entrada de um outro que esteja em uma camada à frente da atual (figuras 5.2(a) e 5.2(b)) e cíclicas ou recorrentes (*feedback*⁹) onde a saída de um neurônio pode ser usada como entrada de outro de camada atual ou anterior (figura 5.2(c)).
- *Quanto à conectividade:* Pode ser uma rede completamente conectada (figuras 5.2(a) e 5.2(b)) onde cada um dos neurônios de uma camada da rede está conectado a todos os outros da camada seguinte (HAYKIN, 2001). Ou, caso falte alguma conexão entre estas unidades, tem-se uma rede fracamente (ou parcialmente) conectada (figura 5.2(c)).

Nas figuras 5.2 (a), (b) e (c) são apresentados alguns exemplos de topologias de RNAs.



Esta é uma rede formada por apenas uma camada de neurônios; ela é capaz de resolver apenas problemas que sejam linearmente separáveis.

Figura 5.2(a) – Topologia de uma rede acíclica de camada única totalmente conectada

⁸ Tradução: alimentada para frente

⁹ Tradução: retro-alimentada

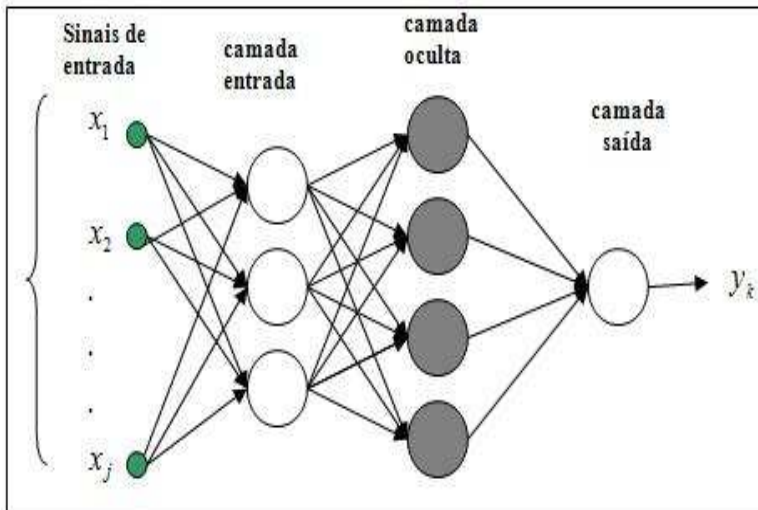


Figura 5.2 (b) – Topologia de uma rede acíclica de múltiplas camadas totalmente conectada

Esta é uma rede MLP formada por uma camada de entrada, uma camada oculta e uma de saída [3-4-1]¹⁰. Este modelo de rede é amplamente utilizado na resolução de diversos tipos de problemas tais como: classificação e aproximação de funções.

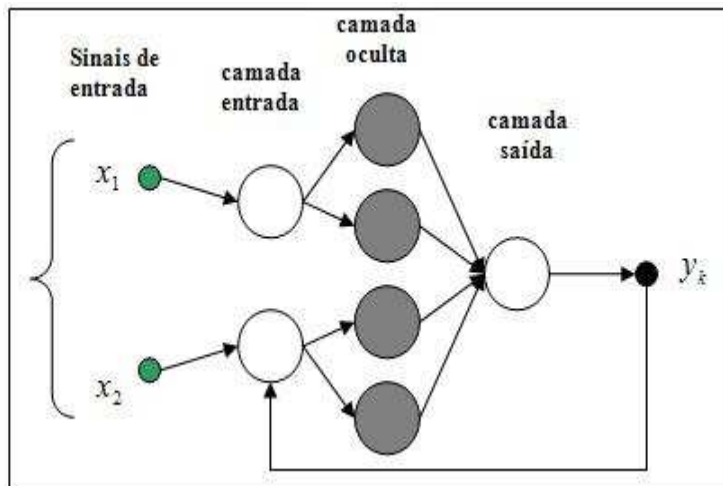


Figura 5.2 (c) – Topologia de uma rede cíclica parcialmente conectada

Esta é uma rede recorrente formada por três camadas, onde a resposta obtida pela rede é usada como sinal de entrada. Entre as aplicações das redes recorrentes está o processamento temporal de dados.

5.5 Métodos de Aprendizagem

A capacidade de aprender, a partir de exemplos, é uma das características que difere as RNAs de outras ferramentas estudadas na área de Inteligência Artificial (IA) ou mesmo da abordagem estatística tradicional. No processo de aprendizagem de uma RNA, não são definidas regras ou parâmetros fixos, mas procura-se encontrar a intensidade das conexões que existem na rede (BRAGA, CARVALHO & LUDEMIR, 2000).

¹⁰ Número de neurônios em cada camada: entrada, intermediária e de saída, respectivamente.

É na fase de aprendizagem que as RNAs realizam a modelagem dos problemas que lhes são apresentados. Por meio da extração de padrões (informações relevantes do conjunto de dados) a rede constrói o modelo matemático que será processado na etapa de reconhecimento. Estes padrões (os modelos) são armazenados nos pesos sinápticos que estão acoplados em todas as conexões da rede. Assim, pode-se dizer que os pesos sinápticos são a memória da RNA, onde está contido todo o conhecimento adquirido por ela.

Segundo os autores BRAGA, CARVALHO & LUDEMIR (2000) os métodos de aprendizagem são divididos, basicamente, em dois grupos principais - *aprendizado supervisionado* e *não-supervisionado* - e dois secundários – *aprendizado por reforço* e *por competição*.

O aprendizado supervisionado consiste de um processo de treinamento no qual são apresentados para a rede valores de entrada e de respostas (saídas) desejadas. A partir da comparação entre a saída atual calculada pela rede e a desejada, são ajustados os pesos sinápticos. O objetivo é minimizar a diferença existente entre os valores calculados pelo sistema e a resposta desejada. Trata-se de um processo incremental, visto que, os pesos das sinapses sofrem pequenas alterações a cada etapa do treinamento até que seja obtida uma resposta que satisfaça algum critério de parada, que pode ser um valor máximo aceitável para o erro ou um determinado número de ciclos de treinamento. O critério mais utilizado como medida de desempenho de uma RNA é a soma dos erros quadráticos de todas as saídas (BRAGA, CARVALHO & LUDEMIR, 2000).

Este ajuste por correção de erros procura minimizar a diferença entre a soma ponderada dos sinais de entrada pelos pesos ($\sum x_j w_{kj}^T = y_k(t)$) e a saída desejada (d_k) para a unidade k . A equação genérica para a alteração dos pesos sinápticos pelo método da correção dos erros é apresentada na equação (5.3) (BRAGA, CARVALHO & LUDEMIR, 2000):

$$w_k(t+1) = w_k(t) + \eta ex_k(t) \quad (5.3)$$

onde:

$w_k(t+1)$ = pesos sinápticos no instante de tempo (t+1)

$w_k(t)$ = pesos sinápticos no instante de tempo (t)

η = taxa de aprendizagem

x_k = sinal de entrada

$e(t) = d_k(t) - y_k(t)$ = erro no instante de tempo (t), lembrando que: $d_k(t)$ é a resposta desejada e $y_k(t)$ é a resposta calculada pela rede.

A equação (5.3) é utilizada tanto no algoritmo de treinamento do modelo perceptron de camada única quanto no *backpropagation* usado para treinar as redes de múltiplas camadas. Este algoritmo de aprendizagem supervisionado é muito utilizado em projetos de RNAs (BRAGA, CARVALHO & LUDEMIR, 2000).

Na Tabela 5.2, tem-se um resumo, passo a passo, do processo de aprendizagem supervisionada de RNAs.

Tabela 5.2 – Síntese do processo de aprendizagem supervisionada

Etapa	Ação
1	Escolha dos valores iniciais dos pesos sinápticos, que normalmente é feita de forma aleatória.
2	Apresentação de um (novo) sinal de entrada cuja saída correspondente é conhecida.
3	Cálculo da saída gerada pela rede (y_k).
4	Cálculo do erro: $e(t) = d_k(t) - y_k(t)$.
5	Verifica-se o valor do erro e o número de ciclos. Se um dos critérios for verdadeiro encerra-se o treinamento.
6	Atualização do contador para a próxima observação a ser apresentada à rede. Caso seja a última observação do conjunto treinamento, reinicializa o contador.
7	Atualização dos pesos: $w_k(t+1) = w_k(t) + \eta e x_k(t)$.
8	Retorna à etapa 2.

Fonte: Adaptado de THOMÉ (2007, p. 27).

Um caso particular do aprendizado supervisionado é o por reforço. Este aprendizado difere do modelo tradicional por ser “[...] **baseado em qualquer medida que possa ser fornecida ao sistema.**” (BRAGA, CARVALHO & LUDEMIR, 2000, p. 25). Nele a única informação fornecida para a aprendizagem da rede é se a resposta calculada está correta ou não. Trata-se de um processo *on-line* que realiza um mapeamento de entrada-saída, com o objetivo de maximizar o índice de desempenho escalar que é conhecido como sinal de reforço.

O método de aprendizagem por reforço é estruturado sobre a idéia de que quanto maior for a satisfação (do sistema) com determinada ação, maior será a chance de ele repeti-la, assim como um sinal de reprovação vai fazer com que o sistema procure ter um outro comportamento. Como nesse tipo de aprendizagem não é fornecida informação (além de certo ou errado) de como a rede pode melhorar o seu desempenho, o seu uso fica restrito a apenas algumas classes de problemas.

Em relação ao método de aprendizado não-supervisionado, pode-se afirmar que sua principal característica é o fato de que neste processo são apresentados a rede neural apenas os sinais de entrada. Como consequência disto, tem-se que somente ocorrerá aprendizado se houver alguma redundância no conjunto de dados de treinamento. O sistema busca, a partir destes dados, encontrar regularidades que permitam construir grupos (automaticamente) para classificar os padrões existentes.

Também, pode-se salientar que o aprendizado por competição é um exemplo importante do método não-supervisionado. Tal processo consiste no fato de que para uma determinada entrada, as unidades de saída (que estão diretamente ligadas a elas) disputam entre si para decidir qual será ativada e, conseqüentemente, terá os seus pesos atualizados no ciclo seguinte de treinamento. A idéia básica é sempre fortalecer as unidades mais fortes, aumentando o seu poder inibidor sobre as outras até que exista apenas uma unidade ativa na rede (a vencedora).

5.6 Modelo *Perceptron* de Múltiplas Camadas (*Multi-Layer Perceptron-MLP*)

As redes MLP foram responsáveis pela retomada do desenvolvimento da teoria das RNAs nos anos oitenta. Estas redes conseguiram superar as limitações do modelo perceptron de camada única que, como destacado anteriormente, são capazes de resolver apenas problemas linearmente separáveis. Para a solução de problemas não-lineares, ficou provado que há necessidade da presença de pelo menos uma camada intermediária (ou também chamada de oculta) na estrutura da RNA.

Ainda, segundo CYBENKO (1989) apud BRAGA, CARVALHO & LUDEMIR (2000) a presença de uma camada oculta garante que a rede pode aproximar qualquer função contínua e, se possuir duas camadas, pode modelar qualquer função.

Quanto à arquitetura, as redes multicamadas são implementadas com uma função de ativação linear e uma função de propagação sigmoideal (logística ou tangente hiperbólica). A escolha das funções sigmoideais deve-se ao fato de que para a utilização do método de treinamento baseado no gradiente descendente é preciso que esta função seja contínua, diferenciável e não decrescente (BRAGA, CARVALHO & LUDEMIR, 2000).

Na figura 5.3, é apresentada a topologia de uma rede MLP, composta por uma camada de entrada, duas intermediárias e uma de saída.

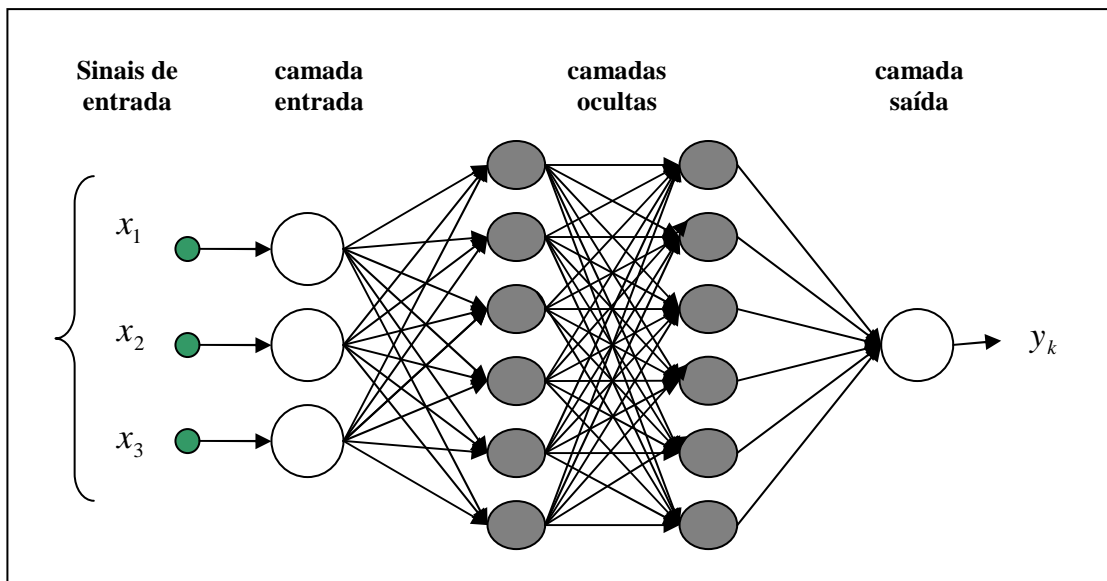


Figura 5.3 – Arquitetura de uma rede MLP

A camada de entrada da rede recebe os sinais externos e os transmite para as camadas seguintes, seus nodos são conhecidos como unidades sensoriais do sistema. Normalmente, define-se o número de neurônios desta camada de acordo com a quantidade de variáveis (sinais de entrada) envolvidas no problema.

Já as camadas intermediárias são responsáveis pela extração das características presentes no conjunto de dados (BRAGA, CARVALHO & LUDEMIR, 2000). O número de unidades destas camadas é uma questão difícil de definir, mas de grande importância para o sucesso da rede. Segundo BRAGA, CARVALHO & LUDEMIR (2000, p. 55) alguns fatores devem ser observados para a definição deste parâmetro, tais como:

- Número de unidades de entrada e saída.
- Tamanho do conjunto de treinamento.
- Quantidade de ruído presente nos dados.
- Complexidade da função a ser modelada.
- Distribuição estatística dos dados de treinamento.

Com base na análise destas informações, os mesmos autores sugerem que uma rede deve possuir um número de conexões dez vezes menor que a quantidade de exemplos do conjunto de treinamento.

Uma outra metodologia é proposta por Hecht-Nielsen/Kolmogov (FREIMAN & PAMPLONA, 2005) na qual o número de unidades das camadas intermediárias deve ser igual à da equação (5.4).

$$k_l = 2j + 1 \quad (5.4)$$

Onde:

k_l = número de neurônios da camada l

j = número de variáveis de entrada.

Uma quantidade grande de nodos nas camadas intermediárias da rede pode comprometer a sua capacidade de generalização, pois além de modelar as características relevantes do problema, também estará absorvendo as informações do ruído. Por outro lado, se o número de neurônios for menor que o necessário, a rede pode não encontrar resposta para o problema.

Em relação ao treinamento das redes MLP, normalmente é dividido o conjunto de exemplos em duas partes: uma que deve possuir, aproximadamente, 70% dos dados (conforme THOMÉ, 2007) que será utilizada para o treinamento propriamente dito da rede. Esta parte é denominada de *conjunto de treinamento*. Com estes exemplos, será feito o ajuste dos pesos sinápticos. O restante, que é chamado de *conjunto de validação*, tem a finalidade de verificar a capacidade de generalização da RNA.

Para a avaliação do desempenho da rede podem ser utilizados os seguintes métodos para mensuração dos erros de predição realizados:

- Erro Percentual Médio Absoluto (*MAPE*): é o valor médio do erro percentual das previsões sobre o conjunto de teste, conforme equação (5.5).

$$MAPE = \frac{1}{N} \sum_{j=1}^N \left| \frac{d_j - y_j}{d_j} \right| * 100 \quad (5.5)$$

onde:

N = número de exemplos do conjunto de teste

j = exemplo atual apresentado à rede

d_j = resposta desejada para o exemplo j

y_j = resposta calculada pela rede para o exemplo j

- Raiz do Erro Médio Quadrado (*RMSE*): a diferença fundamental para o primeiro método é que este ressalta os erros maiores, conforme equação (5.6).

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (d_j - y_j)^2} \quad (5.6)$$

- Coeficiente U de Theil: 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 (5.7) 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).

$$U = \frac{\sqrt{\sum_{j=1}^N (d_j - y_j)^2}}{\sqrt{\sum_{j=1}^N (d_j - d_{j-1})^2}} \quad (5.7)$$

Já para o treinamento das redes MLP, tem-se como principal algoritmo de aprendizagem o *backpropagation* que será detalhado na seção 5.7.

5.7 Algoritmo de Aprendizagem *Backpropagation*

A aprendizagem de uma rede MLP com o algoritmo *backpropagation* consiste em um processo de duas fases distintas, conforme BRAGA, CARVALHO &

LUDEMIR (2000): a primeira denominada de *fase forward*¹¹, é onde os sinais funcionais (HAYKIN, 2001) ou sinais de entrada são propagados da camada de entrada (mais a esquerda) até a última camada (saída). A resposta produzida pelos neurônios de saída é comparada com a resposta desejada, gerando, assim, o sinal de erro. Na segunda fase, conhecida como *fase backward*¹², este sinal de erro é propagado da camada de saída (mais a direita) até a camada de entrada da rede. Nesta fase, é realizado o ajuste dos pesos sinápticos de cada dos nodos das camadas, a fim de minimizar o erro.

Na figura 5.4, tem-se uma ilustração de como os sinais de entrada (estímulos externos) e de erros (calculados pelo sistema) propagam-se ao longo da rede neural.

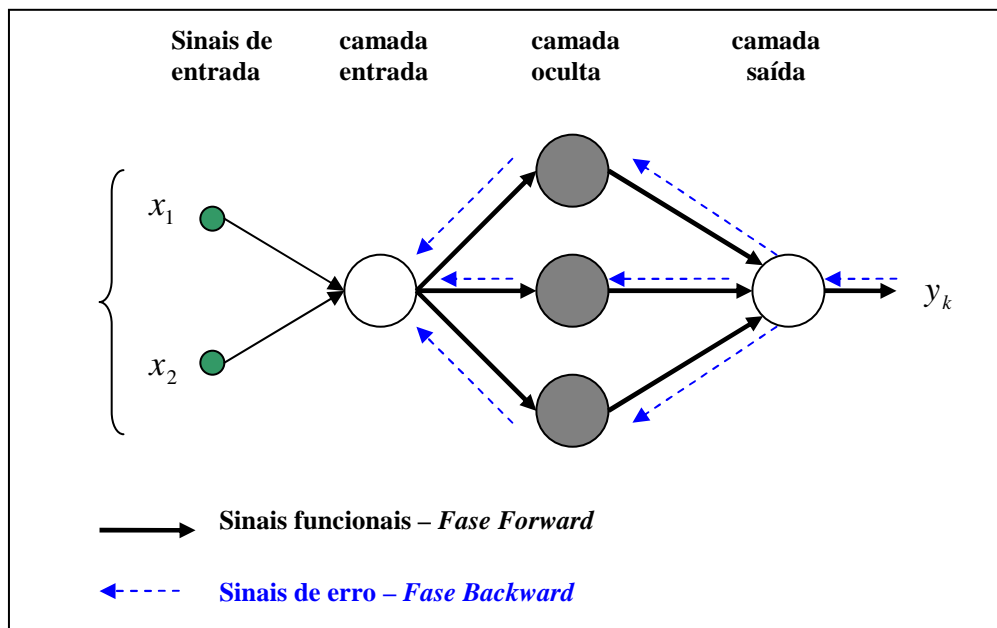


Figura 5.4 – Esquema dos fluxos dos sinais em uma rede MLP
Fonte: Adaptação de HAYKIN (p. 186, 2000).

Na Tabela 5.3, estão algumas informações e notações que serão importantes para facilitar a compreensão das equações que descrevem o funcionamento do *backpropagation*.

¹¹ Tradução: para frente, adiante.

¹² Tradução: para trás.

Tabela 5.3 – Notações utilizadas na demonstração das equações do algoritmo *backpropagation*

Notação	Descrição
Índices i , j e k	Referem-se a neurônios diferentes, onde o neurônio i pertence à camada de entrada, o j à camada oculta e o k à camada de saída.
t	Iteração (tempo), na iteração t , o t -ésimo exemplo é apresentado à rede.
$\xi(t)$	Soma dos erros quadráticos médios ou energia do erro, na interação t .
$d_k(t)$	Resposta desejada para o neurônio k .
$y_k(t)$	Resposta calculada pela rede na saída do neurônio k .
$e_k(t)$	Erro do sinal na saída do neurônio k .
$w_{kj}(t)$	Peso sináptico conectando à saída do neurônio j e à entrada do neurônio k .
$\Delta w_{kj}(t)$	Correção aplicada ao peso sináptico.
b_k	Bias aplicado ao neurônio k .
$v_k(t)$	Igual a função de ativação mais o bias ($net + b_k$); constitui o sinal que será aplicado à função de propagação.
$\varphi_k(\cdot)$	Função de propagação do neurônio k .
$\varphi'_k(\cdot)$	Derivada da função de propagação.
η	Taxa de aprendizagem.
δ	Gradiente local.

O processo realizado pelo algoritmo *backpropagation* pode ser resumido em 4 etapas a seguir descritas:

- **1ª etapa: Cálculo do erro** - Apresenta-se um sinal de entrada e sua respectiva saída (desejada) para a rede e, em seguida, calcula-se o sinal de erro para o(s) neurônio(s) da camada de saída, conforme a equação (5.8):

$$e_k(t) = d_k(t) - y_k(t) \quad (5.8)$$

- **2ª etapa: Definição da função de custo** – Calcula-se a energia total do erro dos neurônios da camada de saída, segundo a equação (5.9):

$$\xi(t) = \frac{1}{2} \sum_{k=1}^C e_k^2(t) \quad (5.9)$$

Onde: C é a quantidade de unidades de saída.

Logo, a energia média do conjunto de treinamento é dada pela equação (5.10).

$$\xi_{med} = \frac{1}{N} \sum_{t=1}^N \xi(t) \quad (5.10)$$

Onde: N é a quantidade de amostras (pontos) do conjunto de treinamento.

Esta é a função de custo que se deseja minimizar na fase de ajuste dos pesos sinápticos.

- *3ª etapa: Cálculo do valor de correção dos pesos sinápticos* – Para minimizar o erro de resposta da rede é aplicada uma correção (Δw_{kj}) aos pesos sinápticos, que tem valor proporcional à derivada parcial de $\frac{\partial \xi(t)}{\partial w_{kj}(t)}$. Assim, a correção dos pesos é feita de acordo com a equação (5.11) (HAYKIN, 2001):

$$\Delta w_{kj}(t) = -\eta \frac{\partial \xi(t)}{\partial w_{kj}(t)} \quad (5.11)$$

Calculando esta derivada parcial pela regra da cadeia tem-se que, conforme equação (5.12):

$$\frac{\partial \xi(t)}{\partial w_{kj}(t)} = -e_k(t) \phi'_k(v_k(t)) y_j(t) \quad (5.12)$$

Lembrando que:

$y_j(t) = \varphi_j(v_j(t))$ é a saída do neurônio j .

$v_k(t) = \sum_{j=1}^A w_{kj}(t) y_j(t)$, onde A é a quantidade de unidades da camada j .

Portanto, pode-se escrever a equação de correção dos pesos, conforme equação (5.13).

$$\Delta w_{kj}(t) = \eta \delta_k(t) y_j(t) \quad (5.13)$$

Onde gradiente local é dado pela equação (5.14):

$$\delta(t) = e_k(t) \phi'_k(v_k(t)) \quad (5.14)$$

Para efetivar o ajuste dos pesos, ainda é preciso considerar em qual camada o neurônio encontra-se. Caso seja na camada de saída, pode-se utilizar as equações (5.13) e (5.14). Porém, se o neurônio pertence à camada oculta, não é possível calcular diretamente o valor do erro, pois não há uma resposta desejada específica para cada unidade. Para calcular o sinal de erro dos neurônios desta camada é utilizado um processo recursivo, que considera os sinais de erro de todos os neurônios que estão conectados diretamente a ele (HAYKIN, 2001). Na verdade, este valor é uma estimativa do erro, por isso, não é recomendável construir RNAs com mais de uma camada oculta, uma vez que esta estimativa contém imprecisões que podem comprometer o sucesso do sistema.

Em HAYKIN (2001) é apresentada a demonstração completa da equação para o cálculo do gradiente local de neurônio oculto, que é expresso pela equação (5.15).

$$\delta_j(t) = \phi'_j(v_j(t)) \sum_{k=1}^C \delta_k(t) w_{kj}(t) \quad (5.15)$$

Logo, a correção dos pesos de um neurônio oculto é dada pela equação (5.16).

$$\Delta_{ji}(t) = \eta \delta_j(t) y_i(t) \quad (5.16)$$

Em síntese, o ajuste dos pesos sinápticos pelo algoritmo *backpropagation* pode ser representado pelo esquema da figura 5.5, onde o cálculo do gradiente local depende da camada da qual faz parte o neurônio.

$$\begin{pmatrix} \text{Correção} \\ \text{do peso} \\ \Delta w(t) \end{pmatrix} = \begin{pmatrix} \text{Taxa de} \\ \text{aprendizagem} \\ \eta \end{pmatrix} \cdot \begin{pmatrix} \text{Gradiente} \\ \text{local} \\ \delta(t) \end{pmatrix} \cdot \begin{pmatrix} \text{Sinal de entrada} \\ \text{do neurônio} \\ y(t) \end{pmatrix}$$

Figura 5.5 – Ajuste dos pesos sinápticos
Fonte: Adaptação de HAYKIN (2001).

- *4ª etapa: Ajuste dos pesos sinápticos* - O vetor dos pesos sinápticos na iteração $(t + 1)$ é fornecido pela equação (5.17).

$$w(t + 1) = w(t) + \eta \delta(t) y(t) \quad (5.17)$$

Este processo de atualização dos pesos pode ser implementado de duas formas: Pode ser por *modo padrão* – onde os pesos são ajustados a cada exemplo do conjunto de treinamento ou por *modo bath* – onde apenas um ajuste é feito a cada época¹³ de treinamento.

O processo de aprendizagem da rede é repetido até que a magnitude do erro seja menor que determinado valor, calculado conforme os critérios apresentados na seção 5.6, ou que seja atingido o número máximo de épocas de treinamento previamente estabelecido.

5.8 Rede TLFN Distribuída e o Algoritmo *Backpropagation* Temporal

As redes MLP treinadas com o algoritmo *backpropagation* têm sido aplicadas com sucesso na resolução de diversos tipos de problemas, tanto de classificação quanto de predição. Porém, quando o problema envolve a modelagem de séries temporais não-estacionárias, os resultados, muitas vezes, não são satisfatórios.

¹³ A apresentação completa do conjunto de treinamento à rede forma uma época. (HAYKIN, 2001)

O autor HAYKIN (2001) destaca que para o processamento de séries de dados não-estacionários o tipo de rede neural mais apropriado é a Rede Alimentada para frente Atrasada no Tempo (TLFN – *Time Lagged Feedforward Network*) Distribuída. Onde o termo *Distribuída* refere-se ao fato de que “[...] a influência implícita do tempo é distribuída através da rede.” (HAYKIN, 2001, p. 702)

A diferença 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.

Na figura 5.6 é apresentado o modelo matemático de um neurônio de uma rede TLFN Distribuída. Já os filtros FIR (filtro de resposta à impulso de duração finita) são descritos na figura 5.7. Eles 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.

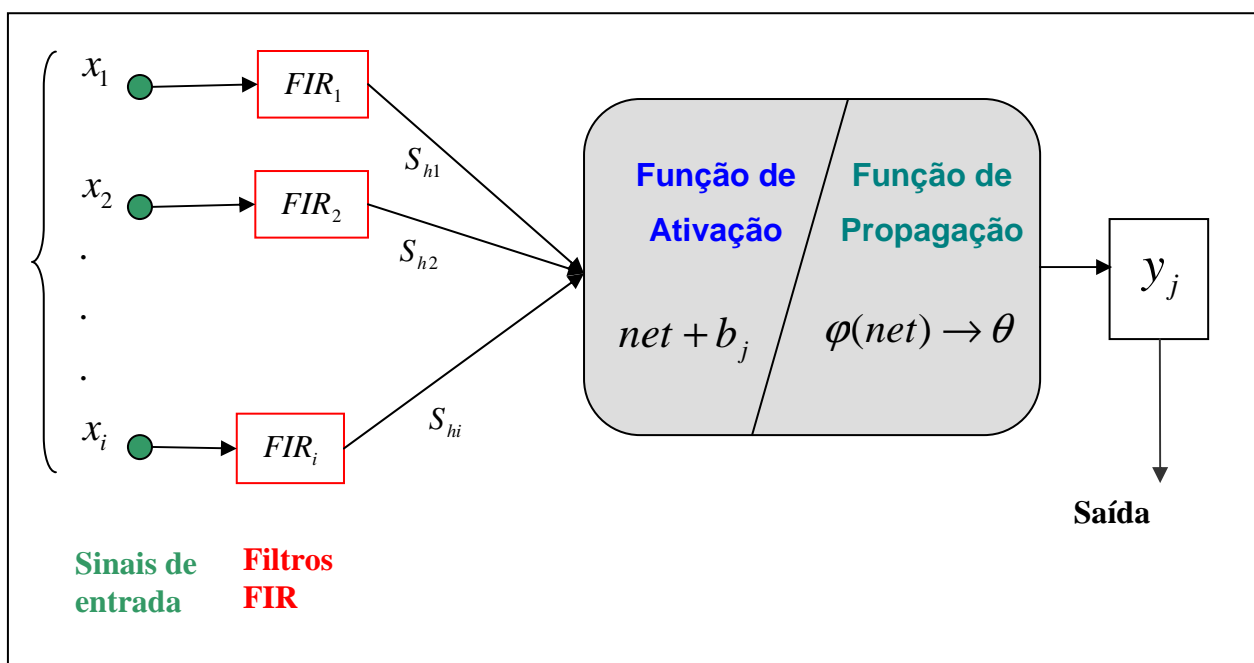


Figura 5.6 – Modelo matemático de um neurônio com filtro neural de múltiplas entradas.

Fonte: Adaptação de HAYKIN (2001)

A saída $y_j(t)$ é dada pela equação (5.18):

$$y_j(t) = \varphi\left(\sum_{i=1}^S \sum_{q=0}^p w_{hi}(q)x_i(t-q) + b_h\right) \quad (5.18)$$

Onde:

$w_{hi}(q)$ = é o peso da q -ésima sinapse secundária pertencente à i -ésima sinapse primária.

$x_i(t)$ = é o sinal de entrada.

p = é a ordem do filtro FIR.

S = é a quantidade de sinais de entrada.

É preciso destacar que em uma rede neural TLFN Distribuída cada sinapse primária é desdobrada em $p + 1$ sinapses secundárias. Assim, as sinapses primárias que em uma rede MLP tradicional são, na prática, um valor escalar (que será multiplicado pelo sinal de entrada na iteração t), passam a ser um vetor de tamanho $p + 1$ em uma rede TLFN. Logo, os valores pertencentes a este vetor serão as sinapses secundárias que serão multiplicadas pelo vetor (também de tamanho $p + 1$) que contém os sinais de entrada nos instantes t até $t - p$.

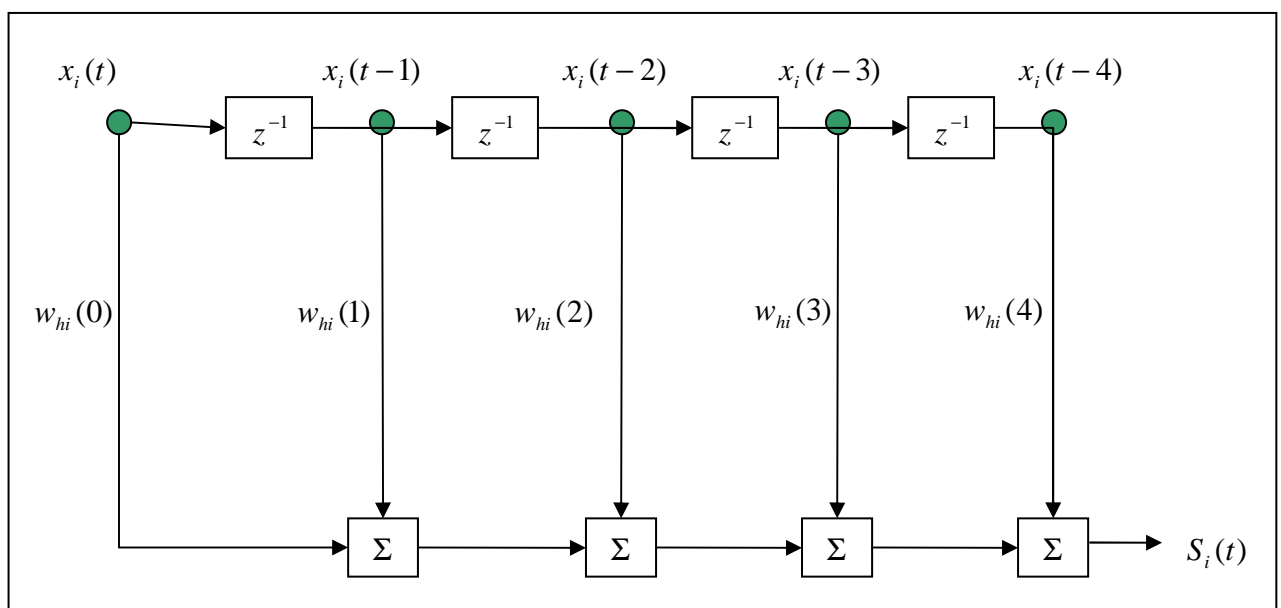


Figura 5.7 – Filtro FIR de ordem igual a 4 ($p = 4$)

Fonte: Adaptação de HAYKIN (2001)

Onde:

z^{-1} = é um operador de atraso de tempo unitário, que aplicado sobre $x_i(t)$ resulta na sua versão atrasada $x_i(t-1)$ ¹⁴.

A partir do esquema apresentado na figura 5.7 pode-se calcular o valor do sinal de entrada $S_j(t)$ do neurônio j pela equação (5.19):

$$S_j(t) = \sum_{q=0}^p w_{hi}(q)x_i(t-q) \quad (5.19)$$

Ou através do produto interno dos vetores $w_{hi}(t)$ e $x_i(t)$, conforme equação (5.20):

$$S_j(t) = W_{hi}^T X_i(t) \quad (5.20)$$

Onde:

$$W_{hi}(t) = [w_{hi}(0), w_{hi}(1), \dots, w_{hi}(p)]$$

$$X_i(t) = [x_i(t), x_i(t-1), \dots, x_i(t-p)]$$

E, portanto, a função de ativação (*net*) pode ser obtida pela equação (5.21):

$$v_j(t) = \sum_{i=1}^S S_i(t) + b_j \quad (5.21)$$

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.

No algoritmo original, o cálculo do gradiente local depende da camada onde se encontra o neurônio. Se for na camada de saída, será obtido pela mesma relação

¹⁴ z^{-1} resulta da aplicação da transformada z , pois para $z\{x(t)\}$ tem-se $z^{-1} = z\{x(t-1)\}$.

do modelo padrão (Equação 5.14), 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. E, logo, o vetor atualizado dos pesos sinápticos ($w_{kj}(t+1)$) será fornecido pela Equação 5.17 fazendo-se as devidas substituições.

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 (5.22), cuja demonstração pode ser obtida em HAYKIN (2001).

$$\delta_j(t) = \varphi'_j(v_j(t)) \sum_{k=1}^C \Delta_k^T(t) W_{kj} \quad (5.22)$$

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)

p = ordem do filtro FIR

E os pesos sinápticos atualizados serão obtidos pela equação (5.23):

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_j(t) Y_i(t) \quad (5.23)$$

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 .

5.9 Considerações

Neste capítulo, foram apresentados os conceitos principais da teoria das Redes Neurais Artificiais, assim como, os métodos de aprendizagem utilizados para o seu treinamento. Também, foram detalhadas as características das redes TLFN Distribuída e do algoritmo *backpropagation temporal*, como sugestão para desenvolvimento do trabalho. Esta rede terá como objetivo prever os valores futuros de alguns ativos financeiros.

O próximo capítulo apresentará a descrição da metodologia utilizada ao longo deste estudo, descrevendo as etapas e procedimentos que foram realizados.

6 METODOLOGIA

Este capítulo apresenta a caracterização da pesquisa e dos seus procedimentos metodológicos. Em termos gerais, segundo PARRA FILHO & SANTOS (2002), pode-se classificar a presente pesquisa como aplicada, uma vez que visa atingir resultados imediatos e tendo como finalidade principal ser utilizada na prática. Contudo, conforme o autor SANTOS (2000), pode-se caracterizar a metodologia empregada nos trabalhos científicos, focando em três critérios principais: objetivos, procedimentos de coleta de dados e fonte das informações.

Destacando estes aspectos será descrita a metodologia adotada nas próximas seções, apresentando a caracterização da pesquisa na seção 6.1 e a síntese dos procedimentos implementados na seção 6.2.

6.1 Caracterização da Pesquisa

O início deste trabalho deu-se pela pesquisa bibliográfica que resgatou alguns dos modelos matemáticos que são utilizados na análise técnica do mercado de capitais brasileiro. Esta etapa pode ser caracterizada, segundo o critério objetivo, como exploratória. Porém, como na etapa final é proposta uma nova metodologia para aperfeiçoar a predição de séries temporais financeiras, pode-se classificar esta fase como explicativa, uma vez que envolve a tentativa de contribuir para a evolução do conhecimento existente sobre o assunto em estudo.

Quanto aos procedimentos de coleta de dados o tipo que melhor se enquadra é o estudo de caso. Este que por sua vez pode ser descrito como um “*Estudo*

aprofundado e exaustivo de um ou de poucos objetos, de maneira a permitir o seu conhecimento amplo e detalhado” (TOGNETTI, 2006).

Em relação às fontes de informações, foi realizada uma pesquisa de campo, caracterizada pela obtenção das diversas séries temporais de ativos financeiros e de índices relevantes da economia brasileira. Além do levantamento bibliográfico, que conforme SANTOS (2000) deve ser a base do processo de investigação científica.

6.2 Síntese dos Procedimentos Metodológicos

Na figura 6.1 é apresentado um esquema de execução da proposta de pesquisa apresentada neste estudo.

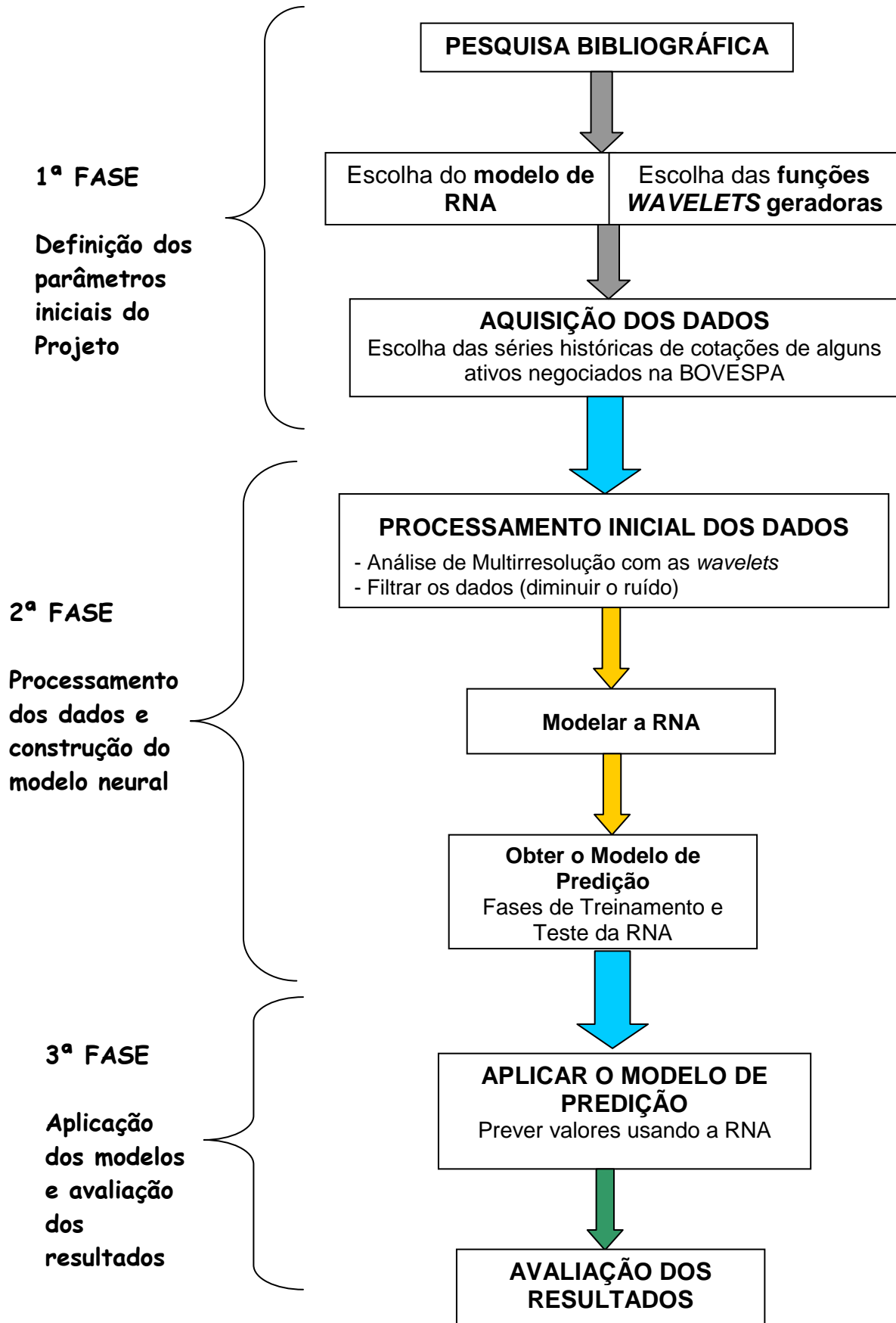


Figura 6.1 – Esquema de execução da técnica proposta.

Na 1ª fase do desenvolvimento do trabalho, foi escolhido o modelo neural a ser implementado. Esta escolha foi feita com base na literatura existente sobre a predição de séries temporais com RNAs. A rede escolhida foi a TLFN Distribuída e o algoritmo *backpropagation temporal*, que já foram utilizadas em outros trabalhos do mesmo domínio. Também, foram coletados dados das cotações de alguns ativos financeiros (ações preferenciais, ordinárias e de cotas de FIIs) negociados na BOVESPA. Assim, como a definição de quais *wavelets* geradoras foram utilizadas durante a fase de processamento dos dados.

Na 2ª fase foi feito o processamento inicial dos dados o qual implicou na realização das seguintes etapas:

- Aplicação da Análise de Multiresolução (AMR) com as *wavelets*, conforme será descrito em detalhes na seção 4.3.
- Filtragem do ruído usando as técnicas de limiarização descritas na seção 4.5 do capítulo 4 deste trabalho.

Na seqüência, foi realizado o processo de treinamento e teste da rede neural com o objetivo de obter o modelo de predição. Também, é preciso salientar que os dados antes das fases de aprendizagem e validação da RNA foram normalizados linearmente no intervalo de zero a um, com o objetivo de facilitar estes processos.

Na 3ª fase do trabalho (fase final) foi utilizado o modelo neural para a predição de valores 'futuros' dos ativos financeiros. No caso das ações da Petrobrás (PETR4) e Telemar (TNLP3) foram previstos 1, 3 e 5 passos à frente, já para as cotas do FII Almirante Barroso foram 1, 10 e 20 passos à frente. E na etapa de encerramento foram avaliados os resultados obtidos com a metodologia apresentada neste estudo.

6.3 Considerações

No próximo capítulo a proposta referente à metodologia, descrita nesta seção, será detalhada. Também, serão apresentados os resultados referentes aos estudos de caso realizados.

7 MODELO DE PREDIÇÃO FINANCEIRA UTILIZANDO *WAVELETS* E REDES NEURAS ARTIFICIAIS

Este capítulo apresenta um modelo de predição financeira direcionado para o mercado de capitais brasileiro.

7.1 Caracterização do Problema

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 FIs (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 quase 3 bilhões de reais, 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. Entre os motivos, que inibem o aumento no volume de investimentos nestes fundos, estão: a falta de informações e de instrumentos de avaliação e a previsão da rentabilidade que poderá ser obtida com tais aplicações.

Este novo mercado de capitais carece de modelos de análise adaptados a sua realidade, pois cada país tem características específicas que influenciam as cotações dos ativos financeiros negociados na bolsa de valores. Por sua vez, as oscilações observadas nestas cotações são extremamente irregulares, pois são

afetadas por diversos fatores (variáveis) que dificultam a construção de modelos matemáticos para a previsão de valores futuros.

Por exemplo, o valor de cada ativo financeiro negociado na bolsa de valores é resultado de uma série de fatores, tais como: potencialidade/credibilidade da empresa (da qual o acionista se tornou sócio ao comprar a ação), relação entre o valor pago pelo ativo e o recebido em dividendos (taxa de retorno do investimento), taxa de juros básica praticada no país, produtividade e capacidade de expansão da empresa, taxa de crescimento do setor em que a empresa atua, liquidez do ativo, problemas econômicos e políticos que podem ser momentâneos ou constantes, especulações aleatórias de investidores, entre outros. Enfim, tudo isso transforma a modelagem do problema em um processo complexo.

Outro aspecto que tem dificultado o desenvolvimento da análise técnica do mercado de capitais é que a maioria dos analistas de investimentos tem uma formação deficiente em matemática e computação, o que os leva a desprezar o uso destas ciências na análise da evolução das séries temporais de preços das ações.

A presente pesquisa buscará analisar as variáveis envolvidas na obtenção das respostas das seguintes questões: Como a aplicação de filtros de sinais construídos com as Transformadas *Wavelets* pode auxiliar na análise do mercado de ações? Como aprimorar a predição financeira usando a matemática e a computação, especialmente, com o uso da inteligência artificial (sistemas de redes neurais)? Quais as variáveis que podem influenciar o valor de uma ação ou cota de um FII?

A aquisição dos dados foi feita diretamente dos *sites* da BOVESPA, do BACEN (Banco Central do Brasil) e da ABECIP (Associação Brasileira das Entidades de Crédito Imobiliário e Poupança). Foram coletados dados referentes a séries históricas de cotações de ativos financeiros, além de índices e taxas de juros.

7.2 Pré-processamento dos Sinais de Entrada da Rede Neural

Foi realizado sobre alguns dos sinais de entrada um pré-processamento utilizando as TWDs. Na figura 7.1 é apresentado o pseudocódigo implementado no software *Matlab*[®] para obter algumas das variáveis de entrada da rede neural. Este código demonstra como foram obtidos os sinais *wavelets* de forma recursiva, onde as previsões para $t+1$ (ou para $t+3$, $t+5$, $t+10$, $t+20$) são feitas com base nas informações conhecidas até o momento t .

```

nw2=2 % nível em que é calculada a TWD
PETR4 % sinal original

for t=1:512
    Xt2=[PETR4(t-3),PETR4(t-2),PETR4(t-1),PETR4(t)]
    Tt=wavedec(Xt2,nw2,'Haar')
    [C2,L2]=wavedec(Xt2,nw2,'Haar')
    A2=wrcoef('a',C2,L2,'Haar',2)
    D2=wrcoef('d',C2,L2,'Haar',2)
    PETR4_H_A2t(t)=A2(1)
    PETR4_H_D2t(t)=D2(1)
End

```

Figura 7.1 – Pseudocódigo utilizado para obter os sinais *wavelets*

Onde:

t = contador, iteração.

$Xt2$ = vetor utilizado para calcular os coeficientes *wavelets* de detalhes e aproximação na iteração t .

wavedec = função do *Matlab* (*toolbox wavelet*) usada para obter os coeficientes *wavelets*.

'Haar' = função geradora wavelet escolhida.

$C2$ = vetor que contém os coeficientes.

$L2$ = tamanho do vetor $C2$.

wrcoef = função (*Matlab*) usada para reconstruir o sinal original somente com os coeficientes de aproximação ou detalhes.

'a' e 'd' = possíveis escolhas para o primeiro parâmetro da função *wrcoef*.

$A2$ = sinal reconstruído somente com coeficientes de aproximação.

$D2$ = sinal reconstruído somente com coeficientes de detalhes.

$PETR4_H_A2t(t)=A2(1)$ = vetor com os sinais de aproximação, nível 2, da série de cotações das Ações da Petrobrás, calculados de forma recursiva.

$PETR4_H_D2t(t)=D2(1)$ = vetor com os sinais de detalhes, nível 2, da série de cotações das Ações da Petrobrás, calculados de forma recursiva.

Salienta-se que o tamanho do vetor $Xt2$ está diretamente ligado ao nível, para qual se quer calcular a TWD, e a quantidade de coeficiente da função geradora *wavelet* escolhida. Por exemplo, para o cálculo da TWD no nível 2, a função Haar (2 coeficientes) necessita 4 valores ou, em outras palavras, é o resultado da média de quatro valores do sinal original, já a DAUB4 (Daubechies de 4 coeficientes) precisa de 10 pontos, conforme WALKER (1999).

7.3 Topologia da Rede Neural Implementada no *Software Matlab*®

Na figura 7.2 é apresentada um modelo de rede TLFN Distribuída. 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 representará o valor que se deseja prever, ou seja, a cotação futura do ativo financeiro. Em cada conexão existente entre as camadas de entrada e a oculta e entre esta e a de saída, há um filtro FIR.

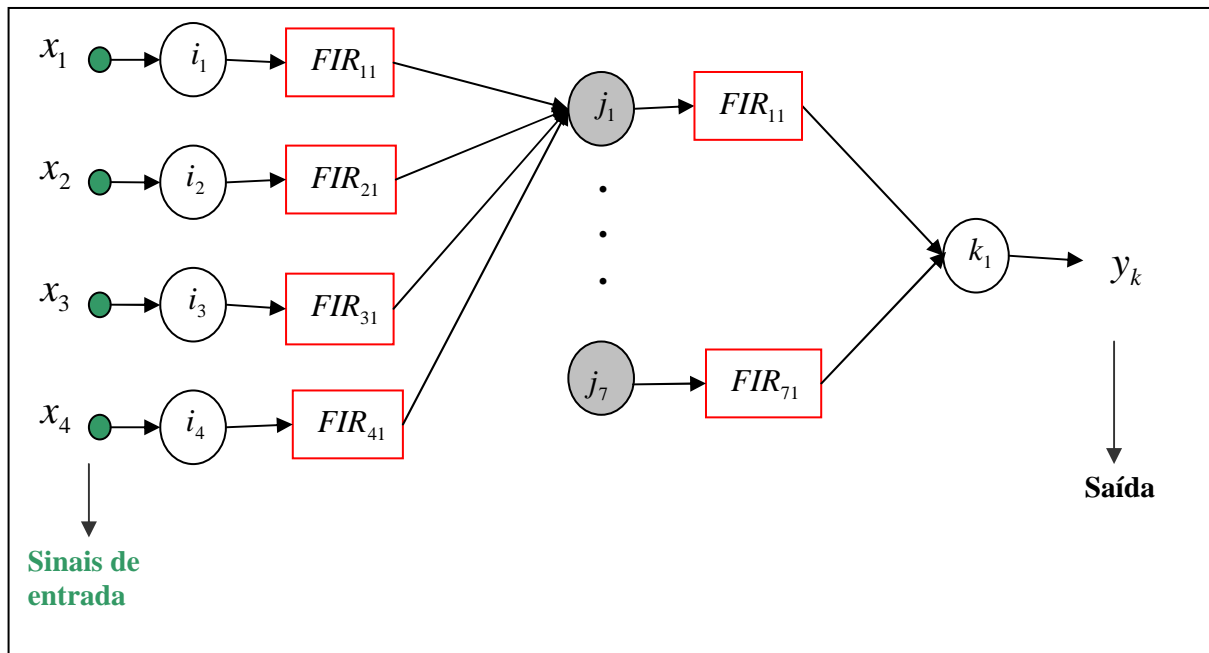


Figura 7.2 – Esquema de uma rede TLFN Distribuída com uma camada oculta

Em síntese, a metodologia apresentada justifica-se por dois fatores principais: o grande potencial que as RNAs possuem para realizar a predição de valores em séries temporais, especialmente, quando estas são influenciadas por múltiplas variáveis. E pela contribuição fornecida pela AMR feita com as TWDs que possibilitam a diminuição do ruído nos dados e a localização temporal das oscilações relevantes, além da criação e/ou aprimoramento dos indicadores gráficos de tendências e reversão utilizados na análise técnica do mercado de capitais.

Para implementar a RNA TLFN Distribuída foi utilizado o *software Matlab*[®], assim como, para a aplicação das Transformadas *Wavelets* Discretas nos sinais de entrada da rede.

7.4 Predição de Séries Temporais Financeiras

Para aplicar a metodologia apresentada foram utilizadas três series temporais de ativos financeiros negociados na BOVESPA. A série de cotações da ação preferencial da Petrobrás, negociada com o código PETR4, além das cotações da

ação ordinária da Telemar (TNLP3) e das cotas do fundo de investimento imobiliário FII Almirante Barroso (FAMB11B).

O objetivo foi treinar a rede neural para prever os valores futuros das cotações destes ativos financeiros. Além de identificar quais variáveis podem influenciar os seus preços. Para atingir os objetivos propostos, foram realizadas varias simulações. Na Tabela 7.1 estão expostos os valores dos parâmetros que foram comuns a todas as simulações.

Tabela 7.1 - Principais parâmetros dos experimentos realizados

Parâmetro	Valor/ Tipo
Normalização dos dados	Linear [0,1]
Função de ativação	Linear
Função de propagação	Tangente hiperbólica
Valor do parâmetro T da função de propagação	1
Valor inicial dos pesos sinápticos	0,0000
<i>Bias</i> dos neurônios das camadas intermediárias e de saída	0,0001
Ordem do filtro FIR da camada oculta	2
Ordem do filtro FIR da camada de saída	1

Inicialmente, foi realizada uma fase de pré-teste com o objetivo de definir os melhores valores para os parâmetros constantes na Tabela 7.1. Estas escolhas foram feitas de forma empírica com base em várias simulações. Por exemplo, para a normalização dos dados foram experimentadas a linear no intervalo [0,1] e no intervalo [-1,1]. Para a função de propagação, foram testadas a logística e a tangente hiperbólica que apresentou em todas as simulações resultados superiores à logística. Quanto ao valor de T foi observada a relação direta que existe com a taxa de aprendizagem, que por sua vez influencia o processo de aprendizado da rede; foi feita a opção por fixar o valor T e variar a taxa de aprendizagem. Para o valor inicial dos pesos sinápticos foram testadas algumas das metodologias descritas em Pavelka & Prochálka (2004), tais como as funções $rand$, $randn$ (disponíveis no *Matlab*) e os algoritmos $2.(rand)-1$ e $A.(rand)-\frac{A}{2}$, onde $A = 0,72$. Porém, os resultados não foram superiores ao obtido quando se iniciava a rede com

os valores iguais a zero. Tendo ainda, a desvantagem de que, em algumas simulações, o número de épocas de treinamento foi maior.

Para avaliação dos resultados além dos critérios apresentados na seção 5.3 do capítulo 5 foi também utilizada a volatilidade histórica, definida pela equação (7.1) (abaixo), das séries temporais analisadas.

$$\text{Volatilidade histórica} = \sigma_x \sqrt{K} \quad (7.1)$$

Onde σ_x é o desvio padrão das variações de preços calculadas conforme a equação (7.2) e K é o número de dias úteis - por exemplo, para a volatilidade histórica diária tem-se $K = 1$, já para a volatilidade semanal tem-se $K = 5$.

$$X_t = \ln\left(\frac{d_{t+1}}{d_t}\right) \quad (7.2)$$

Onde: d_t é o preço de fechamento do ativo financeiro no dia de negociação t e X_t é a variação ocorrida nos preços entre o dia j e o dia seguinte de negociação $t + 1$.

7.4.1 Previsão das Cotações da Ação PETR4

Para realizar a previsão das cotações da ação PETR4 foi utilizada uma série histórica de suas cotações¹⁵, contendo os preços de fechamento do mercado nos dias que efetivamente houve negociação, do período compreendido entre maio de 2006 até junho de 2008, num total de 512 pontos. Esta série foi dividida em duas partes, uma com 358 pontos (70% do total) e outra com 154. A primeira foi usada para o treinamento da rede e a outra para a fase de teste da rede neural. A figura 7.3 apresenta o gráfico desta série financeira.

¹⁵ Foram feitos ajuste na série histórica a fim de preservar o valor real da ação que sofreu um *split* (desdobramento de ações, onde a porcentagem possuída por cada acionista permanece a mesma, mas o número de ações é duplicado e, conseqüentemente, o valor de cada ação passa a ser metade do que era antes) durante o período analisado.

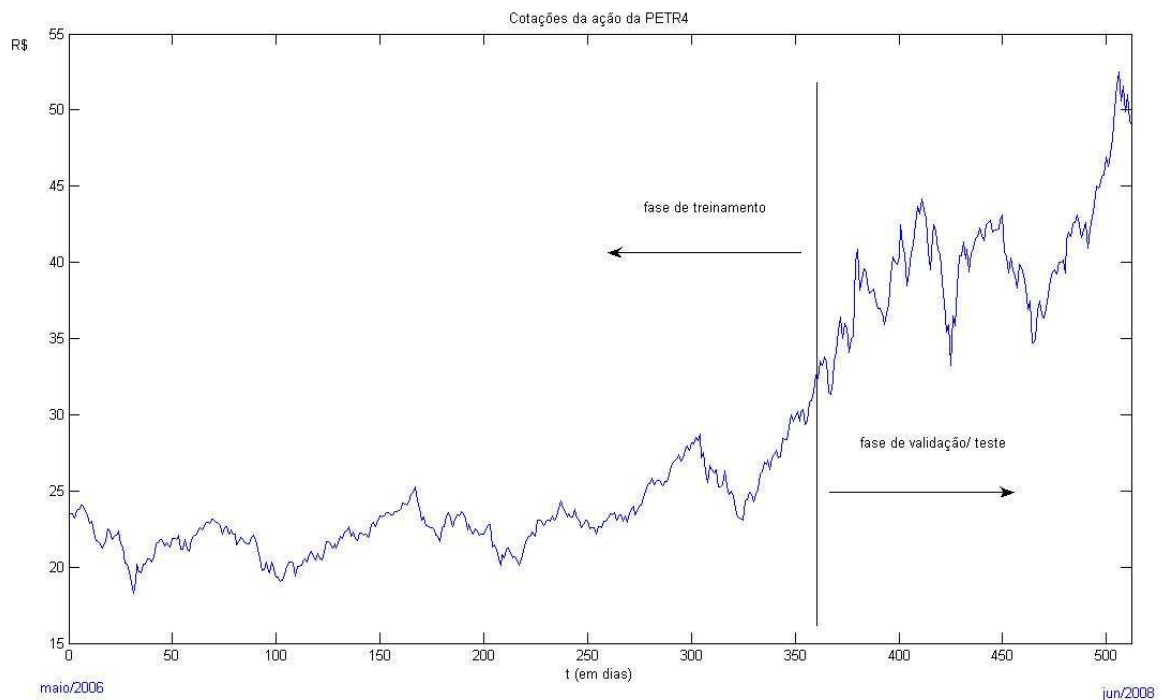


Figura 7.3 – Gráfico com as cotações da ação PETR4

Analisando o gráfico da figura 7.3 é facilmente observável que existe uma grande diferença entre os dados da fase de treinamento e teste. Na parte usada para o treinamento da rede a volatilidade dos preços é igual a 1,88% bem inferior a da fase de teste, onde se tem uma volatilidade histórica igual a 3,00%. O início da fase de teste coincide com a divulgação de fatores que influenciaram profundamente o valor das ações da Petrobrás, tanto positivamente como o caso da descoberta das reservas de petróleo do pré-sal, como negativamente no caso do surgimento dos primeiros dados negativos da economia americana, como resultado da crise financeira desencadeada pelo mercado imobiliário americano.

A princípio, utilizar uma série com características mais homogênea levaria a resultados melhores, porém em condições reais isto talvez nunca seja possível, pois fatores inesperados sempre afetarão o mercado de capitais. Assim, com esta série temporal pode-se testar a eficiência da metodologia em condições críticas.

A seguir, na tabela 7.2 são apresentados os principais parâmetros das simulações realizadas. A primeira coluna contém a denominação para cada um dos experimentos, cada simulação (de 1 a 6) foi repetida para a previsão no horizonte de

1, 3 e 5 passos à frente. Nas demais colunas encontram-se, respectivamente, as informações referentes às variáveis de entrada, função geradora *wavelet* utilizada no pré-processamento dos dados, taxa de aprendizagem da RNA, quantidade de épocas de treinamento – momento em que foi paralisado o processo de aprendizagem da rede e o número de neurônios da camada oculta. Sendo que a quantidade de unidades desta camada foi estabelecida com base em testes realizados na fase inicial de implementação do modelo neural de predição.

Tabela 7.2 - Principais parâmetros dos experimentos - Série PETR4*

Experimento	Variáveis de Entrada	Função Wavelet	Taxa de Aprendizagem	Épocas de Treinamento	No. Neurônios da Camada Oculta
Simulação 1	PETR4_a2 PETR4_d1 PETR4_d2 PETR4_d3	Haar	0,0415	20	07
Simulação 2	PETR4_a1 PETR4_d1 PETR4_d2 PETR4_d3	DAUB4**	0,0415	20	07
Simulação 3	PETR4 Petroleo SELIC IPCA	-	0,16	20	07
Simulação 4	PETR4_a2_haar Petroleo_a2_haar SELIC_a3_db4 IPCA_a3_db4	Haar/ DAUB4	0,16	20	07
Simulação 5	PETR4	-	0,12	03	25
Simulação 6	PETR4_max4_h	Haar	0,12	03	25

* Além destas simulações foram realizadas outras, porém estas foram omitidas, visto que, os seus resultados foram pouco significativos.

** Daubechies com 4 coeficientes.

Já na Tabela 7.3 é apresentada a descrição de cada uma das variáveis de entrada utilizadas nos experimentos constantes da Tabela 7.2. A segunda coluna desta tabela, além da descrição dos códigos, apresenta também a fonte da qual foram obtidos os dados.

Tabela 7.3 - Descrição das variáveis de entrada – Série PETR4

Variável	Descrição
PETR4	Preços de fechamento da ação na BOVESPA
PETR4_a1	Sinal de aproximação a_1 de PETR4
PETR4_a2	Sinal de aproximação a_2 de PETR4
PETR4_d1	Sinal de detalhes d_1 de PETR4
PETR4_d2	Sinal de detalhes d_2 de PETR4
PETR4_d3	Sinal de detalhes d_3 de PETR4
Petroleo	Preços de fechamento do barril de petróleo <i>brent</i> . Fonte: Petrobras S.A.
SELIC	Taxa básica de juros – valor diário. Fonte: Banco Central do Brasil
IPCA	Valor do IPCA (Índice de Preços ao Consumidor Amplo – calculado pelo IBGE), índice utilizado pelo BACEN para fixar as metas de inflação. Fonte: ABECIP
Petroleo_a2_haar	Sinal de aproximação a_2 de Petroleo utilizando a função wavelet Haar
SELIC_a3_db4	Sinal de aproximação a_3 de SELIC utilizando a função wavelet DAUB4
IPCA_a3_db4	Sinal de aproximação a_3 de IPCA utilizando a função wavelet DAUB4
PETR4_max4_h	Sinal reconstruído após a aplicação do processo de limiarização rígido, sendo utilizado como valor para o parâmetro T o <i>minimax</i> dividido por 4.

Na simulação 6, foi feita uma redução no valor do parâmetro de limiarização, uma vez que, com o valor original do *minimax* (que já é o tipo de limiar de menor valor para esta série temporal), todos os coeficientes *wavelets* eram zerados, resultando no sinal PETR4_a1 que teve um desempenho insatisfatório na construção do modelo de previsão. Na Tabela 7.4 são apresentados os valores dos parâmetros de limiarização calculados para a série temporal PETR4 e testados na fase de pré-teste da rede.

Tabela 7.4 – Valor dos parâmetros de limiarização para a série PETR4

Parâmetro	Valor
Universal	3,3349
Minimax	1,8609
SURE	2,2910
SURE--híbrido	3,3349
Minimax (dividido por 4)	0,4652

Na Tabela 7.5 são apresentados os principais resultados alcançados com cada uma das simulações realizadas para as previsões das cotações da ação PETR4, um, três e cinco 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 7.5 - Principais resultados obtidos – Série PETR4

Experimento	Horizonte Previsão *	MAPE (%)		RMSE (R\$)		Coef. de U		Intervalo Confiança (%) **
		Tre	Val	Tre	Val	Tre	Val	
Simulação 1	1	1,67	2,70	0,50	1,37	1,15	1,17	65
	3	2,64	4,16	0,79	2,13	1,06	1,01	67
	5	3,38	5,35	1,02	2,66	1,01	1,00	72
Simulação 2	1	2,63	4,19	0,78	2,15	1,81	1,83	47
	3	3,37	5,37	1,02	2,65	1,36	1,26	55
	5	4,08	6,26	1,21	3,05	1,19	1,15	59
Simulação 3	1	1,85	9,20	0,60	4,46	1,38	3,80	09
	3	1,94	7,74	0,64	3,91	0,86	1,85	30
	5	1,91	6,81	0,65	3,50	0,64	1,32	51
Simulação 4	1	1,90	10,57	0,62	5,05	1,41	4,30	08
	3	1,94	8,71	0,64	4,34	0,86	2,06	26
	5	1,91	7,46	0,65	3,83	0,64	1,44	45
Simulação 5	1	1,86	2,90	0,55	1,49	1,26	1,27	63
	3	2,67	4,94	0,80	2,48	1,08	1,18	61
	5	3,30	6,77	0,98	3,36	0,97	1,27	60
Simulação 6	1	2,13	3,38	0,63	1,75	1,45	1,49	55
	3	2,85	5,35	0,85	2,69	1,15	1,28	57
	5	3,44	6,89	1,02	3,54	1,01	1,33	61

* passos à frente.

** percentual de previsões da fase de validação que estão dentro do intervalo de confiança estabelecido pela volatilidade histórica da ação.

Conforme observado na Tabela 7.5, a Simulação 1, envolvendo a análise de multirresolução *wavelet* (utilizando como *wavelet*-mãe a função Haar) para obtenção dos sinais de entrada (sinal de aproximação a_2 , e sinais de detalhes d nos níveis de resolução 1, 2 e 3), foi o experimento que apresentou os melhores resultados,

especialmente, na previsão para 3 e 5 passos à frente. A Simulação 2, que utilizou os mesmos sinais, porém com a utilização da função DAUB4, teve resultados inferiores em comparação ao primeiro experimento.

Já os modelos construídos a partir das simulações 3 e 4, ao contrário dos anteriores, não atingiram resultados satisfatórios, pois na fase de validação da rede todos os critérios de avaliação mostraram, claramente, que eles não tiveram a capacidade de generalização que era esperada. Além de que a Simulação 4, que utilizou as TWDs (com a função Haar no sinal PETR4 e DAUB4 nos demais), apresentou um resultado inferior ao experimento 3.

Apesar de serem variáveis influentes na formação dos preços das ações da PETR4 não foi possível, com os experimentos realizados, mensurar percentualmente esta influência. Também, é preciso destacar que os horizontes de previsão contribuíram para estes resultados; horizontes maiores beneficiariam a análise feita com esta metodologia, visto que, as variáveis SELIC e IPCA são, normalmente, apuradas (ou modificadas no caso da SELIC) em períodos iguais ou superiores a 30 dias.

Em relação às simulações 5 e 6, modelos de previsão compostos por uma única variável de entrada, as principais constatações que se pode fazer são de que os filtros de remoção de ruído não melhoraram as previsões e de que em comparação ao modelo multivariado os seus resultados foram levemente inferiores, especialmente, quando se compara os resultados da Simulação 5 com a 1.

Nos gráficos das figuras 7.4 e 7.5, são visualizados os resultados obtidos na fase de teste da rede neural na previsão das cotações da ação PETR4, para o horizonte de 1 passo à frente. Em azul, tem-se a resposta desejada (cotação de fechamento na BOVESPA) e em vermelho a resposta calculada pelo modelo construída a partir da Simulação 1. Na figura 7.5 as linhas pontilhadas delimitam o intervalo de confiança. A linha superior é a soma da resposta desejada com o percentual referente à volatilidade histórica diária da ação (valor calculado para o conjunto de teste da rede), já a inferior é a resposta desejada subtraída deste mesmo percentual.

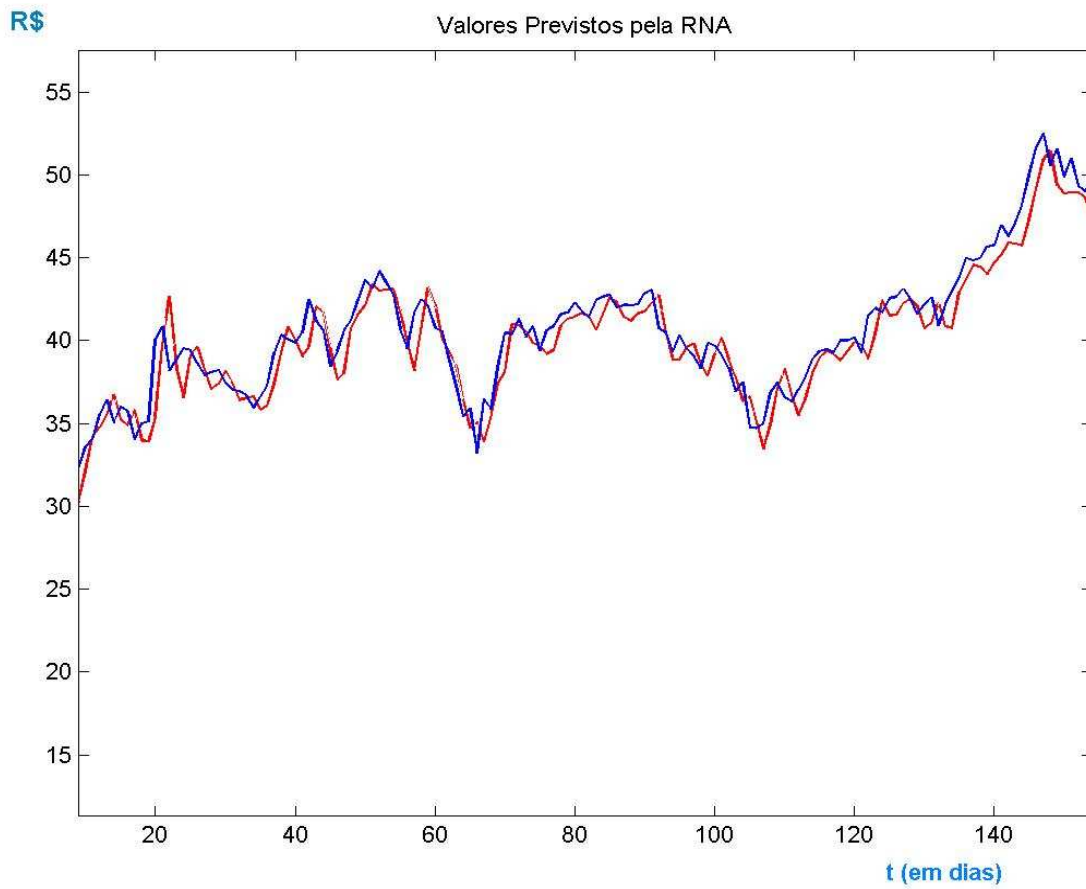


Figura 7.4 – Valores previstos pelo modelo – Simulação 1

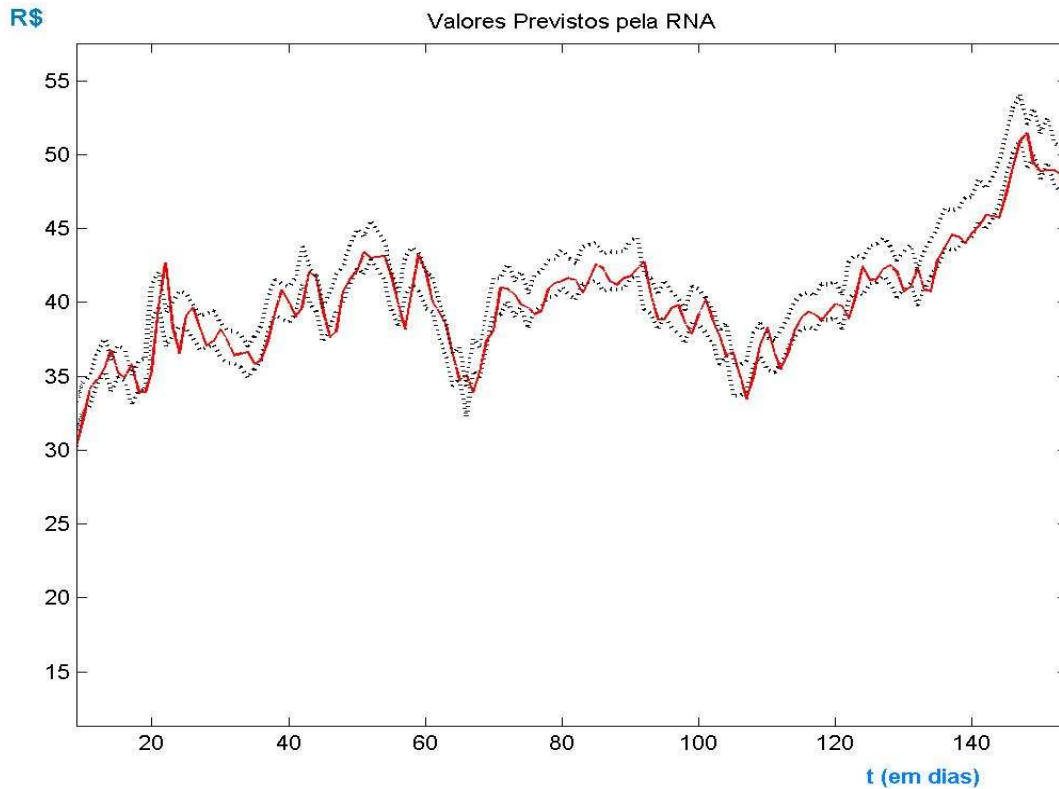


Figura 7.5 – Valores Previstos (Simulação 1) X Intervalo de Confiança

Para mensurar o percentual de influência de cada variável de entrada nas simulações 1 a 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 predição (GALETY, 1996). Na Tabela 7.6 são apresentados os percentuais para cada uma das variáveis.

Tabela 7.6 - Percentual de contribuição ao modelo de predição de cada variável de entrada – Série PETR4

Experimento	Variável	Valor (%)		
		1	3	5
Simulação 1	PETR4_a2	79,15	78,55	80,24
	PETR4_d1	12,02	12,35	11,28
	PETR4_d2	6,42	7,39	6,23
	PETR4_d3	2,40	1,71	2,25
Simulação 2	PETR4_a1	84,36	85,17	86,59
	PETR4_d1	12,97	11,44	12,28
	PETR4_d2	0,99	1,40	0,55
	PETR4_d3	1,68	1,99	0,57

Simulação 3	PETR4	27,98	26,36	26,01
	Petroleo	29,78	29,93	29,97
	SELIC	34,87	36,75	37,66
	IPCA	7,37	6,96	6,36
Simulação 4	PETR4_a2_haar	25,34	24,87	25,06
	Petróleo_a3_db4	28,62	28,82	28,85
	SELIC_a3_db4	36,32	37,05	37,44
	IPCA_a3_db4	9,72	9,25	8,65

Em virtude dos resultados considerados não satisfatórios, obtidos pelos modelos desenvolvidos a partir das simulações 3 e 4, não é possível afirmar que os percentuais calculados para cada variável expressem a verdadeira influência que cada uma dessas variáveis tem nos preços de cotação da ação PETR4.

7.4.2 Previsão das Cotações da Ação TNLP3

Para realizar a previsão das cotações da ação TNPL3 foi utilizada uma série histórica de suas cotações, contendo os preços de fechamento do mercado nos dias que efetivamente houve negociação, do período compreendido entre junho de 2005 até junho de 2007, num total de 512 pontos. Esta série foi dividida em duas partes, uma com 358 pontos (70% do total) e outra com 154. A primeira foi usada para o treinamento da rede e a outra para a fase de teste da rede neural. A figura 7.6 apresenta o gráfico desta série financeira.

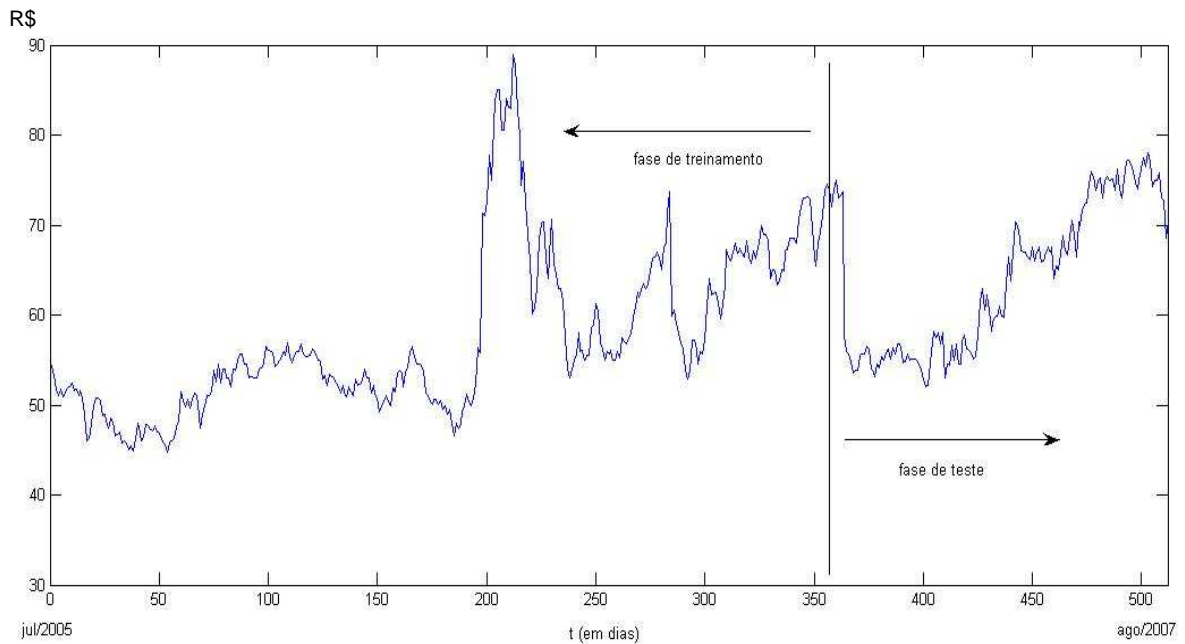


Figura 7.6 – Gráfico com as cotações da ação TNLP3

Ao contrário da série PETR4, as cotações das ações ordinárias da Telemar apresentam uma volatilidade praticamente igual para os conjuntos de treinamento e teste. Na parte usada para o treinamento da rede, a volatilidade dos preços é igual a 3,11% e na fase de teste fica em 3,15%. Se for analisada a série completa (512 pontos), observa-se que a variância da TNLP3 é maior se comparada à PETR4. Os principais motivos para este resultado são: as notícias e/ou os fatos ligados à empresa Telemar e ao mercado de telefonia nacional. Visto que, esta foi uma época de estabilidade no mercado de capitais, marcado pela valorização da maioria dos ativos negociados na BOVESPA.

A seguir, na tabela 7.7 são apresentadas as principais simulações realizadas para a avaliação da metodologia proposta para a predição das cotações da série TNLP3.

Tabela 7.7 - Principais parâmetros dos experimentos – Série TNLP3***

	Variáveis de Entrada	Função Wavelet	Taxa de Aprendizagem	Épocas de Treinamento	No. Neurônios da Camada Oculta
Simulação 1	TNLP3_a2 TNLP3_d1 TNLP3_d2 TNLP3_d3	Haar	0,03	17-11*	07
Simulação 2	TNLP3_a1 TNLP3_d1 TNLP3_d2 TNLP3_d3	DAUB4	0,03	11	07
Simulação 3	TNLP3 MME_12 MME_26 MACD	-	0,02	20-02**	07
Simulação 4	TNLP3	-	0,1	05	25
Simulação 5	TNLP3_max4_s	Haar	0,1	05	25

* Para a previsão 1 passo à frente foram 17 épocas de treinamento, nas demais, foi utilizado 11.

** Para a previsão 5 passos à frente foram 02 épocas de treinamento.

*** Além destas simulações foram realizadas outras, porém estas foram omitidas, visto que, os seus resultados foram pouco significativos.

Já na Tabela 7.8 é apresentada a descrição de cada variável de entrada.

Tabela 7.8 - Descrição das variáveis de entrada – Série TNLP3

Variável	Descrição
TNLP3	Preços de fechamento da ação na BOVESPA
TNLP3_a1	Sinal de aproximação a_1 de TNLP3
TNLP3_a2	Sinal de aproximação a_2 de TNLP3
TNLP3_d1	Sinal de detalhes d_1 de TNLP3
TNLP3_d2	Sinal de detalhes d_2 de TNLP3
TNLP3_d3	Sinal de detalhes d_3 de TNLP3
MME_12	Média móvel exponencial de 12 dias da série temporal TNLP3
MME_26	Média móvel exponencial de 26 dias da série temporal TNLP3
MACD	Convergência e Divergência de Médias Exponenciais (<i>Moving Average Convergence Divergence</i>), construído a partir das séries MME_12 e MME_26
TNLP3_max4_s	Sinal reconstruído após a aplicação do processo de limiarização suave, sendo utilizado como valor para o parâmetro T o <i>minimax</i> dividido por 4.

Na Tabela 7.9 são apresentados os valores dos parâmetros de limiarização, calculados para a série temporal TNLP3 e testados na fase de pré-teste da rede.

Tabela 7.9 – Valor dos parâmetros de limiarização para a série TNLP3

Parâmetro	Valor
Universal	3,3349
Minimax	1,8609
SURE	1,2728
SURE-híbrido	3,3349
Minimax (dividido por 4)	0,4652

Na Tabela 7.10 são apresentados os principais resultados alcançados com cada uma das simulações realizadas para as previsões das cotações da ação TNLP3, um, três e cinco 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 7.10 - Principais resultados obtidos para a série TNLP3

Experimento	Horizonte Previsão *	MAPE (%)		RMSE (R\$)		Coef. de U		Intervalo Confiança (%) **
		Tre	Val	Tre	Val	Tre	Val	
Simulação 1	1	2,44	2,63	2,23	2,40	1,13	1,18	71
	3	4,14	3,95	3,74	3,57	1,04	1,04	77
	5	5,12	4,83	4,68	4,32	0,99	1,06	80
Simulação 2	1	4,22	4,15	3,78	3,79	1,92	1,87	53
	3	5,39	5,32	4,74	5,10	1,31	1,50	67
	5	6,19	5,93	5,63	5,63	1,19	1,39	77
Simulação 3	1	4,37	6,59	3,87	5,09	1,96	2,52	41
	3	5,61	9,77	4,85	7,38	1,35	2,16	49
	5	5,82	8,74	5,34	6,87	1,13	1,69	59
Simulação 4	1	2,79	2,63	2,56	2,54	1,29	1,25	71
	3	4,08	4,17	3,65	3,97	1,01	1,16	78
	5	4,84	4,77	4,16	4,51	0,88	1,11	83
Simulação 5	1	3,48	3,55	3,18	3,35	1,61	1,66	59
	3	4,24	4,70	3,75	4,33	1,04	1,27	72
	5	5,71	7,24	5,33	6,77	1,13	1,66	71

* passos à frente.

** percentual de previsões da fase de validação que estão dentro do intervalo de confiança estabelecido pela volatilidade histórica da ação.

A Simulação 1, que envolveu a análise de multirresolução *wavelet* (utilizando como *wavelet*-mãe a função Haar) para obtenção dos sinais de entrada (sinal de aproximação a_2 , e sinais de detalhes d nos níveis de resolução 1, 2 e 3), foi, novamente, o experimento que apresentou os melhores resultados, especialmente, na previsão para 3 e 5 passos à frente. A Simulação 2, que utilizou a *wavelet*-mãe DAUB4 para o pré-processamento dos sinais de entrada, apresentou resultados pouco satisfatórios, assim como tinha ocorrido com a série PETR4.

Quanto às simulações 4 e 5, modelos de previsão compostos por uma única variável de entrada, os resultados foram semelhantes aos obtidos para a série PETR4. Também se constatou que os filtros de remoção de ruído não melhoraram as previsões e que, em comparação ao modelo multivariado, os seus resultados foram levemente inferiores, especialmente, quando se compara os resultados da Simulação 4 com a 1 na fase de teste (resultados da coluna “*Val*” da Tabela 7.10) da rede.

A Simulação 3, realizada com as médias móveis exponenciais e com o sinal MACD, foi o modelo que apresentou os piores resultados. Porém, deve-se salientar que o período (12 e 26 dias) escolhido para calcular as médias móveis pode ter influenciado neste resultado, uma vez que, foram realizadas apenas previsões de curto prazo (1, 3 e 5 dias).

Os resultados referentes ao percentual de influência de cada variável de entrada nas simulações de 1 a 3 são apresentados na Tabela 7.11.

Tabela 7.11 - Percentual de contribuição ao modelo de predição de cada variável de entrada – série TNLP3

Experimento	Variável	Valor (%)		
		1	3	5
Simulação 1	TNLP3_a2	60,62	71,09	69,39
	TNLP3_d1	19,66	15,68	17,13
	TNLP3_d2	11,86	8,88	10,28
	TNLP3_d3	7,86	4,35	3,20
Simulação 2	TNLP3_a1	76,11	78,50	75,40
	TNLP3_d1	19,22	17,63	16,74
	TNLP3_d2	3,03	0,30	4,10
	TNLP3_d3	1,64	3,57	3,76
Simulação 3	TNLP3	43,08	39,51	29,00
	MME_12	25,06	23,84	33,35
	MME_26	26,31	25,90	34,79
	MACD	5,55	10,75	2,86

Comparando os tipos de sinais e os percentuais constatados para as simulações de número 1 da série PETR4 e da TNLP3 observa-se uma semelhança, indicando que estas variáveis podem ser as mais adequadas para o treinamento da rede, visto que, o modelo construído a partir delas possui uma capacidade de generalização maior. Também, onde foram encontrados os melhores resultados para a fase de teste em ambas as séries temporais.

7.4.3 Previsão das Cotações do FII Almirante Barroso

Para realizar a previsão das cotações do FII Almirante Barroso (ativo negociado no Mercado de Balcão Organizado da BOVESPA com o código FAMB11B) foi utilizada uma série histórica, contendo os preços de fechamento do mercado nos dias que efetivamente houve negociação, do período compreendido entre janeiro de 2006 até junho de 2008, num total de 512 pontos. Esta série foi dividida em duas partes, uma com 358 pontos (70% do total) e outra com 154. A

primeira foi usada para o treinamento da rede e a outra para a fase de teste da rede neural. A figura 7.7 apresenta o gráfico desta série financeira.

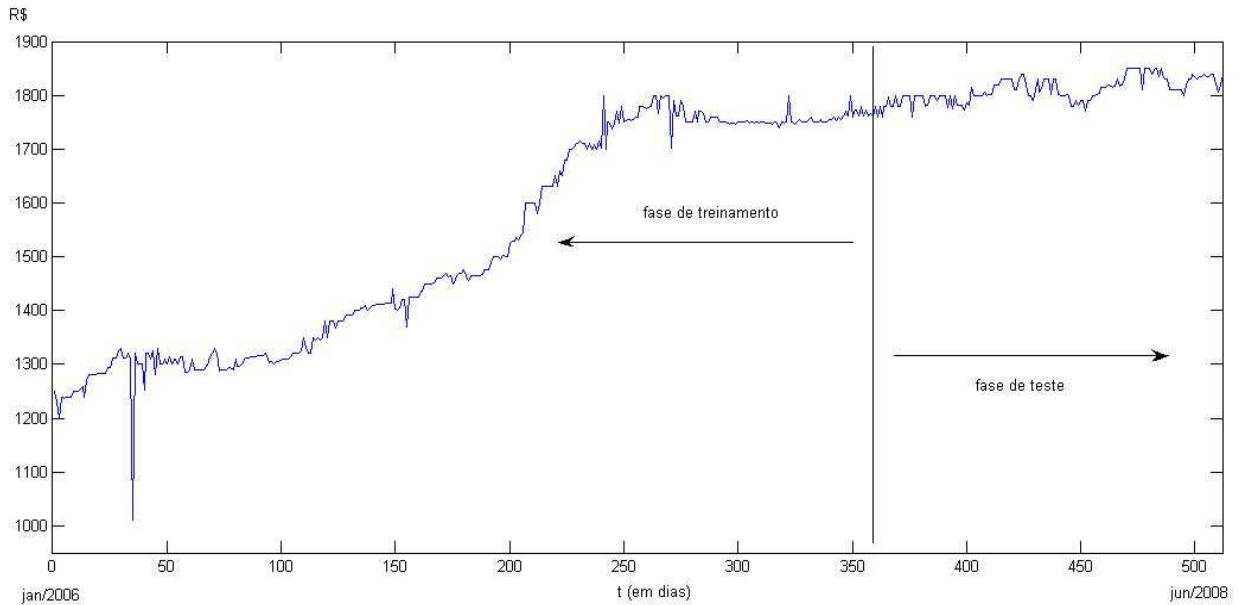


Figura 7.7 – Gráfico com as cotações do FII Almirante Barroso

Analisando o gráfico da figura 7.7 é facilmente observável que existe uma grande diferença entre este e os gráficos das séries PETR4 e TNLP3, no que diz respeito à volatilidade. Enquanto a volatilidade histórica diária das outras séries temporais é de aproximadamente 3%, a do FII é de 2,28% na fase de treinamento e de apenas 0,74% na fase de teste.

A diferença é explicada pelo fato de se tratar de outro ativo financeiro, com características diferentes das ações. O valor das cotas dos FIIs é muito mais dependente da performance do empreendimento imobiliário no qual estão lastreadas do que nas demais variáveis econômicas, apesar de também serem afetados por fatores como inflação, taxa básica de juros e do crescimento da economia.

A seguir na tabela 7.12 são apresentadas as principais simulações realizadas para série FAMB11B:

Tabela 7.12 - Principais parâmetros dos experimentos – Série FAMB11B*

	Variáveis de Entrada	Função Wavelet	Taxa de Aprendizagem	Épocas de Treinamento	No. Neurônios da Camada Oculta
Simulação 1	FAMB_a3 FAMB_d1 FAMB_d2 FAMB_d3	Haar	0,02	18	07
Simulação 2	FAMB_a1 FAMB_d1 FAMB_d2 FAMB_d3	DAUB4	0,02	18	07
Simulação 3	FAMB IND_RENT M_IBOVESPA SELIC	-	0,1	18	07
Simulação 4	FAMB IND_RENT_a3 M_IBOVESPA_a3 SELIC_a3	DAUB4	0,1	18	07
Simulação 5	FAMB	-	0,025	05	25
Simulação 6	FAMB_max4_h	Haar	0,025	05	25

* Além destas simulações foram realizadas outras, porém estas foram omitidas, visto que, os seus resultados foram pouco significativos.

Já na Tabela 7.13 é apresentada a descrição de cada variável de entrada.

Tabela 7.13 - Descrição das variáveis de entrada – Série FAMB11B

Variável	Descrição
FAMB	Preços de fechamento da cota no Mercado de Balcão Organizado da BOVESPA
FAMB_a1	Sinal de aproximação a_1 de FAMB
FAMB_a3	Sinal de aproximação a_3 de FAMB
FAMB_d1	Sinal de detalhes d_1 de FAMB
FAMB_d2	Sinal de detalhes d_2 de FAMB
FAMB_d3	Sinal de detalhes d_3 de FAMB
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
FAMB_max4_h	Sinal reconstruído após a aplicação do processo de limiarização rígido, sendo utilizado como valor para o parâmetro T o <i>minimax</i> dividido por 4.

Na Tabela 7.14 são apresentados os valores dos parâmetros de limiarização calculados para a série temporal FAMB11B e testados na fase de pré-teste da rede.

Tabela 7.14 – Valor dos parâmetros de limiarização para a série FAMB11B

Parâmetro	Valor
Universal	3,3349
Minimax	1,8609
SURE	0,0707
SURE- híbrido	0,0707
Minimax (dividido por 4)	0,4652

Na Tabela 7.15 são apresentados os principais resultados alcançados com cada uma das simulações realizadas para as previsões das cotações do FII Amiranter Barroso, um, 10 e 20 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 7.15 - Principais resultados obtidos – Série FAMB11B

Experimento	Horizonte Previsão *	MAPE (%)		RMSE (R\$)		Coef. de U		Intervalo Confiança (%) **
		Tre	Val	Tre	Val	Tre	Val	
Simulação 1	1	0,77	0,61	24,00	13,28	0,83	0,99	73
	10	1,12	0,95	27,30	20,89	0,72	0,90	97
	20	1,44	1,15	32,30	25,46	0,59	0,93	99
Simulação 2	1	0,88	0,74	24,50	15,89	0,85	1,19	55
	10	1,21	1,01	28,70	21,99	0,76	0,95	96
	20	1,51	1,18	33,60	26,43	0,62	0,96	98

Simulação 3	1	0,89	1,31	29,06	29,27	1,01	2,19	37
	10	0,88	1,13	28,99	24,84	0,77	1,08	88
	20	0,87	1,26	28,18	28,29	0,52	1,03	97
Simulação 4	1	0,87	1,58	29,45	33,22	1,02	2,49	22
	10	0,88	1,08	29,20	23,49	0,78	1,02	95
	20	0,87	1,06	28,40	23,85	0,52	0,87	99
Simulação 5	1	0,74	0,56	23,80	13,18	0,83	0,99	70
	10	1,16	1,00	28,20	21,67	0,75	0,94	95
	20	1,40	1,29	31,30	28,54	0,57	1,04	97
Simulação 6	1	0,82	0,62	24,30	14,66	0,85	1,10	64
	10	1,19	1,01	28,30	22,16	0,75	0,96	95
	20	1,45	1,30	31,90	28,73	0,59	1,05	97

* passos à frente.

** percentual de previsões da fase de validação que estão dentro do intervalo de confiança estabelecido pela volatilidade histórica diária da cota do FII Almirante Barroso.

Assim como nas séries temporais anteriores, a Simulação 1 apresentou os melhores resultados. Contudo, é preciso salientar que na atual série os resultados foram bem superiores aos obtidos para a previsão das ações da Petrobrás e da Telemar.

Porém, observa-se uma diferença no que diz respeito aos modelos construídos a partir das simulações 3 e 4, ao contrário dos estudos realizados com a série PETR4, constatou-se uma melhora na capacidade de generalização da RNA com a utilização das TWDs nas variáveis taxa de juros (SELIC), índice de rentabilidade (IND_RENT) e média de pontos do IBOVSPA (M_IBOVSPA). Sendo que o experimento 4 apresentou bons resultados para a previsão de 10 e 20 passos à frente.

Em relação às simulações 5 e 6 (modelo univariado), as principais conclusões são análogas às realizadas nas séries PETR4 e TNL3. Onde se obteve resultados superiores na Simulação 5, mas inferior ao modelo multivariado do experimento 1.

Para esta série financeira, o percentual de influência de cada variável tem uma importância especial, visto que, trata-se de um ativo financeiro pouco estudado

pelos especialistas da área, não existindo referência na literatura. Na Tabela 7.16 são apresentados os referidos percentuais para cada uma das variáveis.

Tabela 7.16 - Percentual de contribuição ao modelo de predição de cada variável de entrada para a série FAMB11B

Experimento	Variável	Valor (%)		
		1	10	20
Simulação 1	FAMB_a3	86,70	94,98	95,15
	FAMB_d1	4,98	1,73	1,63
	FAMB_d2	5,25	2,24	2,28
	FAMB_d3	3,07	1,05	0,94
Simulação 2	FAMB_a1	96,44	96,09	96,06
	FAMB_d1	2,17	2,17	2,12
	FAMB_d2	0,80	1,46	1,04
	FAMB_d3	0,59	0,28	0,78
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

7.5 Análise dos Resultados

Antes realizar uma análise mais detalhada dos resultados obtidos, é preciso destacar como evoluiu o processo de aprendizagem da rede ao longo das épocas de treinamento. Nos gráficos das figuras 7.8 e 7.9 é mostrada a curva do erro percentual médio absoluto (MAPE) – fase de treinamento da Simulação 1 da série PETR4, curvas semelhantes foram observadas para os demais experimentos.

Observa-se que a inicialização da rede com o valor dos pesos sinápticos iguais a zero e a atualização dos mesmos a cada iteração produz a acentuada queda do erro entre a primeira e a segunda época de treinamento. Nas demais épocas de treinamento o erro evolui lentamente (conforme detalhes da figura 7.9). Salienta-se, ainda, que não foi utilizado nenhum método (algoritmo de otimização) de validação cruzada entre as fases de treinamento e validação com objetivo de selecionar o melhor momento para parar o processo de aprendizagem e, assim, obter o modelo que tenha a maior capacidade de generalização.

O tempo gasto para o treinamento da rede neural foi de aproximadamente 2 minutos para cada época de aprendizagem, utilizando um processador *AMD Turion 64* (velocidade de 2.0 GHz, 512kb de cache L2) e com 1GB de memória *RAM*.

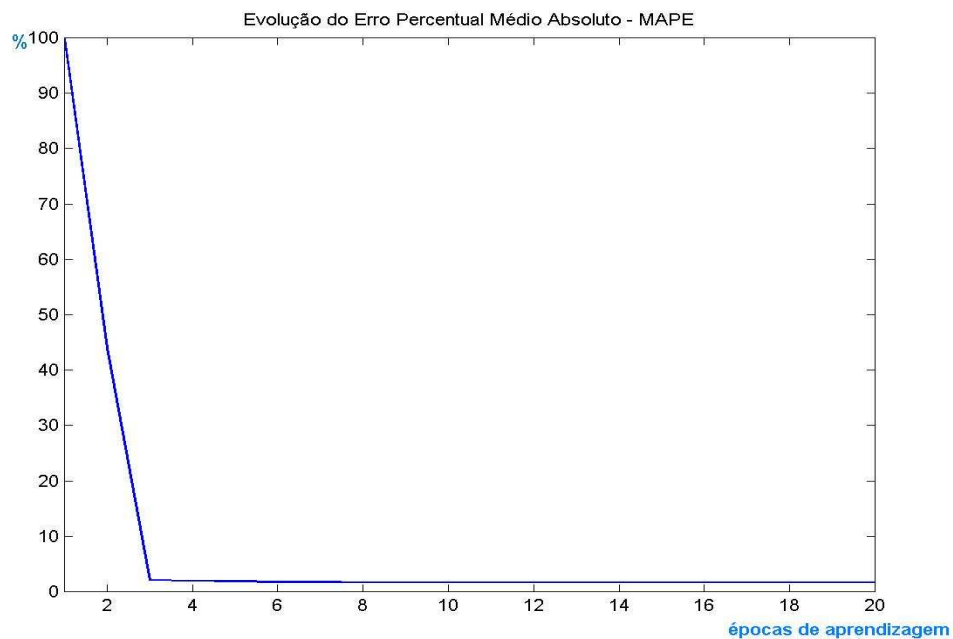


Figura 7.8 – Curva de evolução do MAPE na fase de treinamento da RNA

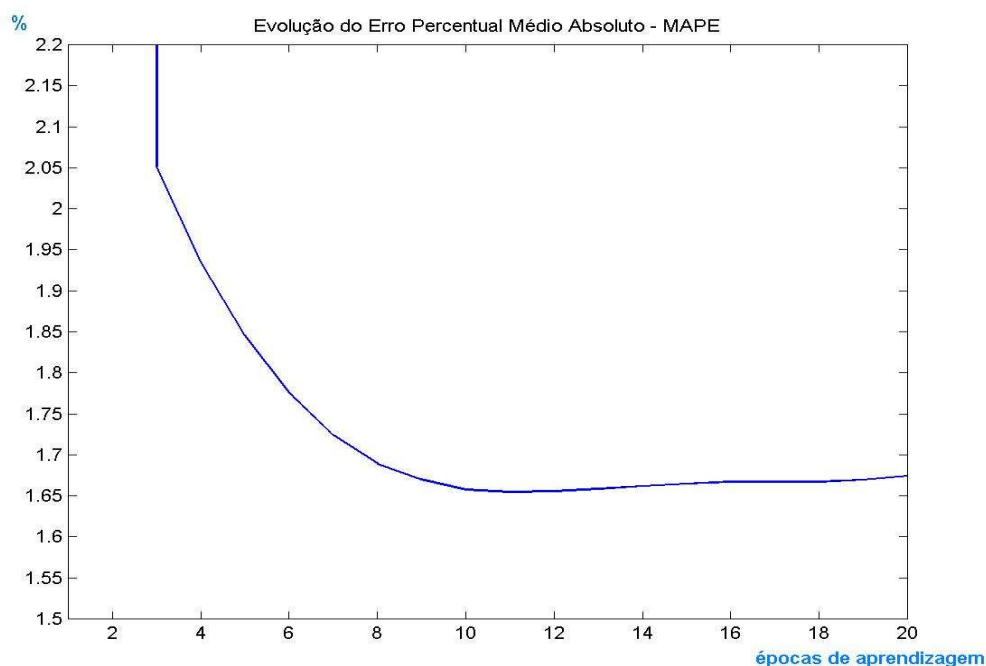


Figura 7.9 – Detalhes da curva de evolução do MAPE

Uma pergunta sempre surge quando se fala em métodos (especialmente aqueles integrantes da escola técnica) aplicados à previsão do mercado de capitais: O método pode ser utilizado na prática?

Para responder esta pergunta é preciso considerar dois aspectos: os resultados obtidos (com as limitações que os estudos de caso possuem) e a complexidade do problema (real).

Em relação aos resultados, além dos critérios apresentados nas tabelas referentes a cada série temporal, salienta-se também, a média dos percentuais de previsões que acompanharam a alta ou queda observada na série original (real). Tais percentuais ficaram em torno de 53% no caso das previsões das ações da Petrobrás e da Telemar (51,10% e 53,07 para as simulações 1 e 5, respectivamente, da série PETR4 e 53,51% e 53,07% para as simulações 1 e 4 da série TNLP3). Já para o FII Almirante Barroso os percentuais ficaram em 67,76% (Simulação 1) e 68,64% (Simulação 5).

Considerando a hipótese que cada operação lucrativa (compra ou venda no momento correto) compensa uma outra errada (onde houve prejuízo), pode-se supor

que ao final de um número grande de operações, o resultado será um lucro de aproximadamente 6% para negociações que envolvam as ações PETR4 e TNLP3. Contudo, sempre cabe ao analista responsável pela operação decidir sobre a conveniência ou não de utilização do método ou sua combinação com outros, o que é, normalmente, recomendado.

Como mencionado, o método exige que haja um grande número de operações, portanto, só é adequado para grandes investidores interessados em operações de curto prazo (especulação) e, mesmo assim, não há garantia de lucro, apenas a expectativa. Para pequenos investidores, os valores das taxas de corretagem e os impostos iriam superar o lucro projetado.

No que diz respeito à complexidade da análise técnica do mercado acionário, Paulos (2004, capítulo 9, pg. 192) a caracteriza da seguinte forma:

“Essa marcha inexorável no sentido da complexidade crescente decorre das iniciativas dos co-investidores, que, ao perceberem (ou inferirem, ou receberem informações) que alguém está obtendo bons resultados por meio de alguma regra simples de análise técnica, tentarão agir da mesma maneira. Para reagir às respostas desses outros investidores, deve-se complicar a própria regra e provavelmente diminuir o excesso de retorno... Não demora muito para que a regra revista-se de complexidade quase aleatória, reduzindo o excesso de retorno a quase zero e voltando a aumentar a dependência em relação ao acaso.”

Neste capítulo que Paulos (2004) dedica ao paradoxo do mercado eficiente e da complexidade dos métodos de análise, o autor reforça a idéia de que a busca por novas metodologias de análise de investimentos resulta no aumento da sua complexidade e na diminuição de sua eficiência. Porém, paradoxalmente, se a maioria dos analistas não almejarem o aprimoramento dos métodos atuais, os poucos que fizerem terão bons resultados com pouco esforço.

Analisando-se em separado o caso do FII Almirante Barroso, é preciso ter cautela com os ótimos resultados obtidos, pois se trata de um ativo de baixa liquidez, pouco conhecido no mercado de capitais, cujos resultados podem não se confirmarem no futuro.

Ainda, em relação aos FIIs, uma importante constatação que se pode fazer é de que o início da crise do mercado imobiliário americano não afetou o valor de suas cotas no mercado secundário até o momento. Mais do que isso, a crise aliada às novas medidas anunciadas pela CVM para incentivar o desenvolvimento dos FIIs, feitas através da edição da Instrução n. 472/2008, tem aumentado o número de investidores interessados neste tipo de ativo financeiro.

Além de que, a partir dos resultados satisfatórios obtidos, especialmente pelo modelo desenvolvido a partir da Simulação 4 da série FAMB, é possível concluir, com a ressalva de que podem estar restritos a este caso, que a taxa básica de juros da economia tem uma influência maior nas cotações dos FIIs do que o desempenho do mercado de ações (que é expresso pelo IBOVESPA). Tal conclusão é importante ao passo que pode subsidiar a decisão de analistas no momento de diversificar as carteiras de investimentos. Onde os FIIs podem se tornar uma alternativa de investimento nessa época de crise no mercado de ações e de tendência de baixa da taxa SELIC.

Em relação às simulações que utilizaram as técnicas de filtragem (*denoising*) de ruído implementadas com as TWDs (simulação 6 das séries PETR4 e FAMB11 e simulação 5 da série TNLP3) a dificuldade foi estabelecer o valor do parâmetro de limiarização (T) ou, em outras palavras, distinguir ruído aleatório de oscilações relevantes que podem auxiliar na previsão destes tipos de sinais.

Neste contexto, surge a justificativa para a utilização dos sinais de detalhes nos níveis 1, 2 e 3 (sinais d_1 , d_2 e d_3), especialmente, nas simulações 1 de todas as três séries financeiras analisadas, obtendo ótimos resultados.

Destaca-se, também, que a decisão de fornecer estes sinais de detalhes como variáveis de entrada da rede neural foi baseada em experimentos que mostraram a superioridade do modelo multivariado (tendo a RNA quatro sinais de entrada) formado pela combinação de um sinal de aproximação e três sinais de detalhes da série temporal original.

Contudo, salienta-se que o horizonte de previsão também é um parâmetro a ser analisado, uma vez que possui uma grande correlação com as variáveis de entrada. Como todas as três séries continham apenas os preços de fechamento do mercado, não foi analisado o comportamento da ação ao longo do dia de negociação, o que poderia ter trazido informações valiosas, principalmente, para as previsões de curto prazo, 1 passo à frente, por exemplo. Ao passo que, a ampliação do horizonte de predição poderia reduzir a importância dos sinais de detalhes, especialmente, os de níveis 1 e 2, que carregam as oscilações de curto prazo dos ativos.

7.6 Considerações

Neste capítulo foram apresentados os resultados dos estudos de caso realizados com base na metodologia proposta neste trabalho. Além de uma análise pontual destes resultados e das principais contribuições que esta pesquisa trouxe para a área da Análise Técnica do mercado de capitais. No próximo capítulo serão expostas as principais conclusões obtidas, até o momento, e algumas sugestões de trabalhos futuros que poderão dar continuidade ao desenvolvimento deste estudo.

CONCLUSÃO

O atual nível de desenvolvimento do mercado de capitais brasileiro exige a otimização dos métodos de análise, sejam eles integrantes da escola fundamentalista ou técnica. Os ciclos de crise, como o que está em curso, reforça ainda mais esta idéia, pois aperfeiçoar as ferramentas de previsão pode ser uma forma de prevenção para prejuízos maiores.

Conseguir caracterizar e mensurar corretamente a correlação existente entre os diversos tipos de ativos negociados em bolsa de valores é uma maneira de aumentar a proteção das carteiras de investimentos ao risco que estão expostas.

Este estudo apresentou um método de predição de séries temporais financeiras. O objetivo principal foi a sua aplicação na previsão de preços futuros das ações negociadas na BOVESPA e das cotas dos FII's negociados no mercado secundário.

Como principal contribuição deste trabalho pode-se citar a flexibilidade do método desenvolvido, visto que, pode ser adaptado com facilidade para qualquer tipo de ativo financeiro. Além de que, podem ser criados “sub-modelos” para cada tipo de ação, selecionando, por exemplo, as variáveis mais apropriadas para o setor de atuação da empresa.

Também, pode ser aplicada na análise de outras séries temporais como, por exemplo, na previsão de vendas de determinado produto. Especialmente, em sinais não-estacionários, cujas tendências de curto e médio prazo mudam repentinamente, o método pode ser empregado com maiores chances de ter sucesso frente a outras técnicas tradicionais.

Os principais resultados obtidos são satisfatórios, uma vez que, trata-se de um problema de solução complexa. A comparação com outros trabalhos é dificultada pelas características das séries temporais que sofreram profundas alterações nos últimos seis anos, resultado da transformação marcante do nosso mercado de capitais que atraiu um número expressivo de investidores e de empresas interessadas em abrir o seu capital.

No que diz respeito à identificação e análise das variáveis que influenciam os preços de um ativo financeiro, os resultados não foram conclusivos. Necessitam de um número maior de experimentos e testes que envolva outras séries financeiras para que se possa afirmar, com precisão, se é possível ou não com esta metodologia analisar, satisfatoriamente, as variáveis que alteram o desempenho de um ativo financeiro no mercado de capitais. Visto que, é de conhecimento dos analistas financeiros que variáveis como a SELIC e o IPCA influenciam o mercado acionário, mas a grande dificuldade enfrentada, não só neste estudo, mas também em outros constantes na literatura, é mensurar percentualmente esta influência.

Contudo, analisando os resultados apresentados neste trabalho pode-se supor que para as cotações dos FII's seja viável utilizar este modelo para identificar a correlação existente entre as variáveis, já para as cotações das ações da Petrobras a metodologia parece não ser adequada.

Também em relação aos filtros de remoção de ruído (*denoising*) há a necessidade de um estudo mais profundo, especialmente, visando à determinação de um parâmetro de limiarização adequado às características das séries históricas de cotações do mercado de ações brasileiro; o que não existe na literatura atualmente. Contudo, os experimentos realizados com estes filtros mostraram que, provavelmente, não seja possível aplicá-los com sucesso na previsão de séries financeiras, apesar, de serem utilizados, atualmente, com êxito na análise de outros sinais unidimensionais, tais como, sinais sonoros e de eletrocardiogramas.

O presente estudo possibilitou novas perspectivas para o desenvolvimento de trabalhos futuros. Como sugestão pode-se citar:

- Otimizar o modelo desenvolvido nesta pesquisa através da implementação de um programa computacional que permita o aumento do número de variáveis de entrada e a seleção das melhores, além da otimização dos parâmetros de treinamento da rede, tais como: o valor da taxa de aprendizagem e o número de épocas de treinamento.
- Otimizar o modelo, através da seleção das variáveis de entrada e da ordem do filtro FIR, para a previsão em diferentes horizontes, inclusive testando nas negociações *day-trade* (operações de compra e venda de uma ação dentro do mesmo dia).
- Utilizar as TWDs e os métodos de remoção de ruído para otimizar os métodos tradicionais de análise técnica, tais como as médias móveis exponenciais e as outras metodologias que derivam delas.
- Utilizar a metodologia apresentada, testando outros tipos de rede neurais aplicada à previsão de séries temporais, tais como redes recorrentes e RBF, testando juntamente com outras funções *wavelets*.
- Desenvolver um método gráfico a partir das TWDs que indique o melhor momento para comprar ou vender um ativo financeiro.

Durante o período de desenvolvimento deste trabalho, foi apresentado e publicado um artigo (completo) nos anais do XXVII ENEGEP (Encontro Nacional de Engenharia de Produção) realizado no Rio de Janeiro/RJ, de 13 a 16 de outubro de 2008. O título do artigo é: “Predição de Séries Temporais Financeiras utilizando *Wavelets* e a Rede Neural TLFN Distribuída”. Também foi apresentado um pôster no 31º CNMAC (Congresso Nacional de Matemática Aplicada e Computacional) realizado em Belém, Pará, de 8 a 11 de setembro de 2008. O título do artigo publicado na forma de resumo nos anais do evento é: “Aplicação da Análise de Multirresolução *Wavelet* e de Redes Neurais Artificiais no Mercado de Ações”. Além da existência de um convite para publicar na Revista Cadernos do IME - Série Estatística 2008, da Universidade do Estado do Rio de Janeiro (UERJ)

Os estudos de caso realizados mostraram a validade da metodologia utilizada, uma vez que, trouxeram contribuições importantes para o desenvolvimento

de outras pesquisas que envolvem a predição de séries temporais. Também, pode ser a base para a criação de *softwares* “híbridos” que sejam capazes de somar as virtudes das Redes Neurais Artificiais com as das Transformadas *Wavelets* Discretas.

Sendo que ao final deste estudo foi possível concluir que o método de predição implementado atingiu resultados satisfatórios. Podendo já ser utilizado como ferramenta de auxílio na decisão dos profissionais que operam no mercado de capitais.

REFERÊNCIAS

ABECIP. Desenvolvido pela Associação Brasileira das Entidades de Crédito Imobiliário e Poupança. 2001-2008. Na seção de Estatísticas apresenta séries temporais de índices de preços. Disponível em: <<http://www.abecip.org.br>>. Acesso em: 01 ago. 2008.

ABELÉM, Antônio Jorge Gomes. *Redes Neurais Artificiais na Previsão de Séries Temporais*. 1994, 100 f. Dissertação (Mestrado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1994.

AMATO, Fernando Bontorim. 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 do VIII SEMEAD. Disponível em: <http://www.ead.fea.usp.br/Semead/8semead/resultado/an_resumo.asp?cod_trabalho=136> Acesso em: 07 jun 2007.

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, George; NARICI, Lawrence; BECKESTEIN, Edward. *Fourier and Wavelet Analysis*. Ed. Springer. New York, 2000.

BOVESPA. Desenvolvido pela Bolsa de Valores de São Paulo. Apresenta diversas informações sobre o mercado de capitais brasileiro. Disponível em: <<http://www.bovespa.com.br/Principal.asp>>. Acesso em: 01 mar. 2009.

BRAGA, Antônio de Pádua; CARVALHO, André Carlos Ponce de Leon Ferreira; LUDERMIR, Teresa Bernarda. *Redes Neurais Artificiais: Teoria e Aplicações*. Ed. LTC. Rio de Janeiro, 2000.

BRASIL. Comissão de Valores Mobiliários. Instrução CVM n. 472, de 31 de outubro de 2008. Diário Oficial [da] *República Federativa do Brasil*, Brasília, DF, 03 nov. 2008.

CAVALCANTE, Francisco; MISUMI, Jorge Yoshio; RUDGE, Luiz Fernando. *Mercado de Capitais: O que é, como funciona*. 6ª edição. Ed. Elsevier. Rio de Janeiro, 2005.

DE CASTRO, Maria Cristina Felippetto. *Predição não-linear de series temporais usando redes neurais RBF por decomposição em componentes principais*. 2001. 186 f. Tese (Programa de Pós-Graduação em Engenharia Elétrica – Doutorado) Universidade Estadual de Campinas - UNICAMP, Campinas/SP, 2001.

DE OLIVEIRA, Hélio Magalhães. *Análise de Sinais para Engenheiros: Uma Abordagem via Wavelets*. Ed. Brasport. Rio de Janeiro, 2007.

DE SOUZA, Eniuce Menezes. *Efeito de Multicaminho de Alta Frequência no Posicionamento Relativo GPS Estático: Detecção e Atenuação Utilizando Wavelets*. 2004. 140 f. Dissertação (Pós-Graduação em Ciências Cartográficas - Mestrado) – Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista, Presidente Prudente, 2004.

DESCHATRE, Gil Ari. *Ganhe nas Bolsas com o seu Micro*. 2ª edição. Ed. Ciência Moderna LTDA. Rio de Janeiro, 1997.

FACHHOCHSCHULE JENA. *Wavelets – Transformationskodierung. Diskrete Wavelet-Transformation (DWT)*. Disponível em: <<http://www.fhjena.de/contrib/fb/et/personal/ansorg/ftp/wavelet/wavelet.htm>>. Acesso em: 15 set. 2007.

FREIMAN, José Paulo; PAMPLONA, Edson de O. *Redes Neurais Artificiais na Previsão do valor de Commodity do Agronegócio*. In: V Encontro Internacional de Finanças. 2005, Santiago, Chile. Disponível em: <<http://www.iepg.unifei.edu.br/edson/download/ArtFreimanChile05.pdf>>. Acesso em: 12 jan. 2007.

GALETY, E. *Neural Networks for Financial Forecasting*. New York: John Wiley & Sons, 1996.

GALVÃO, Roberto K. H. et al. *Estudo Comparativo sobre Filtragem de Sinais Instrumentais usando Transformadas de Fourier e Wavelet*. Química Nova, n. 6, Vol. 24, p. 874-884. Sociedade Brasileira de Química: São Paulo, 2001. Disponível em:

<<http://quimicanova.s bq.org.br/qn/qnol/2001/vol24n6/index.html>>. Acesso em: 15 jun. 2007.

GENÇAY, Ramazan; SELÇUK, Faruk; WHITCHER, Brandon. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Ed. Academic Press / Elsevier. San Diego, California, USA, 2002.

GUTIERREZ, Carlos Enrique Carrasco. *Eliminação do Ruído por Encolhimento de Wavelets: Uma Aplicação à Série de Preço Spot de Energia Elétrica do Brasil*. 2002. Dissertação (Mestrado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2002.

HAYKIN, Simon. *Redes Neurais: Princípios e Práticas*. 2ª edição. Tradução Paulo Martins Engel. Ed. Bookman. Porto Alegre, 2001.

JENSEN, Mark J. *Making Wavelets in Finance*. Financial Engineering News. n. 01, Vol. 01. Ago. 1997. Disponível em: <<http://www.fenews-digital.com/fenews/199708/?pg=1>>. Acesso em: 25 jun. 2007.

MATSURA, Eduardo. *Comprar ou Vender? Como Investir na Bolsa Utilizando Análise Gráfica*. 5ª edição. Ed. Saraiva. São Paulo, 2007.

MORSELEY, Norman. *Modeling Economic Time Series Using a Focused Time Lagged FeedForward Neural Network*. Proceeding of Student Research Day, CSIS, Pace University, 2003. Disponível em: <<http://csis.pace.edu/~etappert/srd2003/paper14.pdf>>. Acesso em: 25 set. 2007.

MUELLER, Alessandro. *Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário*. 1996. Dissertação (Programa de Pós-graduação em Engenharia de Produção - Mestrado) – Universidade Federal de Santa Catarina, Florianópolis, 1996. Disponível em: <<http://www.eps.ufsc.br/disserta96/mueller/index/index.htm>>. Acesso em: 20 maio 2007.

OSÓRIO, Fernando Santos; BITTENCOURT, João Ricardo. *Sistemas Inteligentes baseados em Redes Neurais Artificiais Aplicados ao Processamento de Imagens*. In: I Workshop de Inteligência Artificial. Departamento de Informática - UNISC, Santa Cruz do Sul, 2000. Disponível em: <<http://www.inf.unisinos.br/~osorio/wia-unisc/wia2000-mini.pdf>>. Acesso em: 15 set. 2007.

PARRA FILHO, Domingos; SANTOS, João Almeida. *Metodologia Científica*. 5ª edição. Ed. Futura. São Paulo, 2002.

PAULOS, John Allen. *A lógica do Mercado de Ações – Uma análise prática do Mercado de Ações*. 7ª edição. Ed. Elsevier. Rio de Janeiro, 2004.

PAVELKA, Ales; PROCHÁLKA, Ales. *Algorithms for Initialization of Neural Network Weights*. Institute of Chemical Technology, Department of Computing and Control Engineering. Inproceedings of Matlab Conference. Praga, República Tcheca, 2004. Disponível em: <http://dsp.vscht.cz/konference_matlab/matlab04/pavelka.pdf>. Acesso em: 15 fev. 2008.

PETROBRAS. Desenvolvido pela Petrobras S.A. Na seção *Relações com o Investidor* apresenta séries históricas de cotações dos preços do petróleo no mercado externo. Disponível em: <http://www2.petrobras.com.br/portal/frame_ri.asp?pagina=/ri/port/index.asp&lang=pt&area=ri> Acesso em: 01 jul. 2008.

REIS, Agnaldo J. Rocha; DA SILVA, Alexandre P. Alves. *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. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-17592004000100013>. Acesso em: 01 jun. 2007.

SANTOS, Antônio Raimundo dos. *Metodologia Científica: a construção do conhecimento*. 3 edição. Ed. DP&A. Rio de Janeiro, 2000.

SOARES, Fabrício; FROZZA, Rejane; PAZOS, Ruben Edgardo Panta. *Predição de Séries Temporais Financeiras Utilizando Wavelets e a Rede Neural TLFN Distribuída*. In: XXVIII ENEGEP – Encontro Nacional de Engenharia de Produção: Rio de Janeiro/RJ, 2008. Anais do XXVIII ENEGEP. Disponível em: <http://www.abepro.org.br/enegep2008/resumo_pdf/enegep/TN_STO_074_527_11078.pdf>. Acesso em: 10 mar. 2009.

_____. *Aplicação da Análise de Multirresolução Wavelet e de Redes Neurais Artificiais no Mercado de Ações*. In: XXXI CNMAC – Congresso Nacional de Matemática Aplicada e Computacional: Belém/PA, 2008. Anais do XXXI CNMAC. Disponível em: <http://www.sbmac.org.br/eventos/cnmac/xxxi_cnmac/PDF/121.pdf>. Acesso em: 10 mar. 2009.

TOGNETTI, Marilza Aparecida Rodrigues. *Metodologia da Pesquisa Científica*. Serviço de Biblioteca e Informação do Instituto de Física de São Carlos – Universidade de São Paulo – USP, São Carlos, 2006. Disponível em:

<http://sbi_web.if.sc.usp.br/metodologia_pesquisa_cientifica.pdf>. Acesso em: 10 dez. 2008.

THOMAZ, Carlos E.; VELLASCO, Marley M. B. R. *Análise de Tendências de Mercado por Redes Neurais Artificiais*. Anais do 7º Congresso Brasileiro de Redes Neurais, Natal/RN, 2005. Disponível em: <<http://www.fei.edu.br/~cet/publications.html>> Acesso em: 22 jun. 2007.

THOMÉ, Antônio Carlos Gay. *Redes Neurais – Uma Ferramenta Para KDD e Data Mining*. Núcleo de Computação Eletrônica – NCE, UFRJ, 2007. Disponível em: <http://equipe.nce.ufrj.br/thome/grad/nn/mat_didatico/apostila_kdd_mbi.pdf>. Acesso em: 10 jan. 2007.

WALKER, James S. *A Primer on Wavelets and their Scientific Applications*. Ed. Chapman & Hall/CRC. 1999.

WOLFRAM MathWorld. *Dirichlet Fourier Series Conditions*. Disponível em: <<http://mathworld.wolfram.com/DirichletFourierSeriesConditions.html>>. Acesso em: 10 mar. 2009.

ZANETI JR., Luiz Antônio; ALMEIDA, Fernando C. de. *Exploração do uso de Redes Neurais na Previsão do Comportamento de Ativos Financeiros*. Terceiro SEMEAD – Faculdade de Economia Administração e Contabilidade – USP, 1998. Disponível em: <<http://www.ead.fea.usp.br/Semead/3semead/Resumos/Art061.htm>>. Acesso em: 01 jul. 2007.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)