



UNIVERSIDADE FEDERAL DE PELOTAS  
INSTITUTO DE FÍSICA E MATEMÁTICA  
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

MARÍLIA TERRA DE MELLO

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NO PROCESSO  
DE PRECIFICAÇÃO DE AÇÕES**

Trabalho de Conclusão apresentado como  
requisito parcial para a obtenção do grau  
de Bacharel em Ciência da Computação

Prof<sup>a</sup>. Dr<sup>a</sup>. Gertrudes A. Dandolini  
Orientadora

Pelotas, Outubro de 2004



# **APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NO PROCESSO DE PRECIFICAÇÃO DE AÇÕES**

Monografia defendida e aprovada em 06 de Outubro de 2004 pela banca  
examinadora composta pelos professores:

---

Prof<sup>a</sup>. Gertrudes Aparecida Dandolini, Dr<sup>a</sup>.

---

Prof. João Artur de Souza, Dr.

---

Prof. Edar da Silva Añaña

**Aos meus pais, Celso e Aglaé**

# SUMÁRIO

<b>LISTA DE FIGURAS .....</b>	<b>7</b>
<b>LISTA DE GRÁFICOS .....</b>	<b>8</b>
<b>LISTA DE TABELAS .....</b>	<b>9</b>
<b>RESUMO.....</b>	<b>10</b>
<b>ABSTRACT.....</b>	<b>11</b>
<b>1 INTRODUÇÃO .....</b>	<b>12</b>
<b>1.1 Motivação.....</b>	<b>12</b>
<b>1.2 Objetivos do trabalho.....</b>	<b>13</b>
1.2.1 Objetivo Geral .....	13
1.2.2 Objetivos Específicos .....	13
<b>1.3 Justificativa.....</b>	<b>13</b>
<b>1.4 Organização do trabalho.....</b>	<b>14</b>
<b>2 SELEÇÃO DE CARTEIRAS .....</b>	<b>15</b>
<b>2.1 Introdução.....</b>	<b>15</b>
<b>2.2 Modelo de Markowitz.....</b>	<b>16</b>
<b>2.3 Precificação de Ações .....</b>	<b>17</b>
<b>3 REDES NEURAS ARTIFICIAIS.....</b>	<b>19</b>
<b>3.1 Introdução.....</b>	<b>19</b>
<b>3.2 Aprendizagem.....</b>	<b>22</b>
<b>3.3 Rede Neural Backpropagation.....</b>	<b>25</b>
3.3.1 Introdução .....	25
3.3.2 Ajuste dos Pesos .....	27
3.3.3 Observações .....	29
<b>3.4 Projeto de uma Rede Neural .....</b>	<b>30</b>
<b>3.5 Vantagens.....</b>	<b>30</b>
<b>3.6 Aplicações .....</b>	<b>31</b>
<b>4 APLICAÇÃO E RESULTADOS.....</b>	<b>32</b>
<b>4.1 Introdução.....</b>	<b>32</b>
<b>4.2 Dados.....</b>	<b>32</b>
<b>4.3 Variáveis de entrada da rede .....</b>	<b>33</b>
<b>4.4 Rede Neural .....</b>	<b>34</b>
4.4.1 Introdução .....	34
4.4.2 Inicialização da Rede.....	35
4.4.3 Treinamento .....	36
<b>4.5 Resultados .....</b>	<b>37</b>
<b>5 CONCLUSÕES.....</b>	<b>41</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>43</b>
<b>APÊNDICE A .....</b>	<b>46</b>

**APÊNDICE B ..... 47**  
**APÊNDICE C ..... 48**

## LISTA DE FIGURAS

Figura 3.1: Neurônio Artificial .....	20
Figura 3.2: Estrutura de uma Rede Neural .....	21
Figura 3.3: Rede Neural Feedforward.....	26
Figura 3.4: Correção dos pesos por backpropagation .....	27
Figura 4.1: Tela do Economática com as cotações da ação Petrobrás PN .....	33
Figura 4.2: Fluxograma do processo de treinamento da rede.....	37

## LISTA DE GRÁFICOS

Gráfico 3.1: Função erro para um único peso.....	28
Gráfico 4.1: Variáveis de entrada (Petrobrás PN) normalizadas entre $[-1,1]$ .....	34
Gráfico 4.2: Resultado com dados de teste .....	39
Gráfico 4.3: Resultado com dados de treinamento .....	39
Gráfico 4.4: Resultado com dados de teste .....	39
Gráfico 4.5: Resultado com dados de treinamento .....	39
Gráfico 4.6: Resultado com dados de teste .....	40
Gráfico 4.7: Resultado com dados de treinamento .....	40



## LISTA DE TABELAS

Tabela 4.1: Resultados da rede neural para a ação Eletrobrás ON .....	38
Tabela 4.2: Resultados da rede neural para a ação Petrobrás PN .....	38
Tabela 4.3: Resultados da rede neural para a ação Vale Rio Doce PNA .....	38
Tabela A.1: Backpropagation aplicada à ação Eletrobrás ON .....	46
Tabela A.2: Backpropagation aplicada à ação Petrobrás PN .....	47
Tabela A.3: Backpropagation aplicada à ação Vale Rio Doce PNA .....	48

## RESUMO

O crescimento da utilização das Redes Neurais Artificiais no mercado financeiro e o bom desempenho demonstrado nesta área têm as tornado uma ferramenta atrativa para ser aplicada no processo de seleção de carteiras. A colaboração das redes neurais neste processo está na tarefa de realizar a seleção das ações que farão parte da carteira, a partir da previsão de seus preços futuros. O presente trabalho propõe o desenvolvimento de uma rede neural feedforward, multicamada, com algoritmo de aprendizagem backpropagation, com o objetivo de realizar a previsão dos retornos de ações no horizonte de um mês. O modelo desenvolvido apresentou, em seus resultados, estimativas de retornos aproximadas dos valores reais, mostrando-se eficiente para ser utilizado como uma ferramenta que auxilia o investidor a tomar decisões no processo de formação da sua carteira. Dispondo das estimativas fornecidas pela rede, o investidor passa a ter maior segurança na escolha das ações que lhe oferecem a melhor relação risco-retorno, de acordo com seu perfil, visando a minimização do risco e a maximização do retorno esperado.

**PALAVRAS-CHAVE:** Redes Neurais Artificiais, Precificação de Ações, Seleção de Carteiras

## **THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN THE PORTFOLIO SELECTION PROCESS**

### **ABSTRACT**

The increase of the Artificial Neural Networks use in the financial market and the good performance shown in this area have made them an attractive tool to be used in the portfolio selection process. The neural network contribution to this process is in the task of making the selection of the stocks that will be in the portfolio, by predicting their future prices. This work proposes the development of a multiplayer, feedforward neural network, trained by the backpropagation algorithm, to perform the prediction of the stocks returns to the next month. The developed model presented, in its results, estimates of the returns that approximate to the real values, showing that it is efficient to be used as a tool that helps the investor to make decisions in the process of forming his portfolio. Having the estimates provided by the net, the investor becomes more secure when choosing the stocks that offer the best risk-return relation according to his profile, trying to minimize the risk and maximize the expected return.

**KEYWORDS:** Artificial Neural Networks, Stock Prediction, Portfolio Selection

# 1 INTRODUÇÃO

## 1.1 Motivação

Com o processo de globalização, que resultou em um intenso intercâmbio entre os países, cada vez mais o mercado acionário vem adquirindo uma crescente importância no cenário financeiro internacional. Hoje, o mercado de ações, além de ser uma importante fonte de financiamento empresarial, é também um meio importante de captação de finanças individuais.

O crescimento do interesse pelo mercado de ações provoca o crescimento da concorrência entre os acionistas; logo, a tentativa de realizar a previsão do comportamento deste mercado é de extrema importância para os investidores. No decorrer dos anos, tem ocorrido um avanço considerável na utilização de técnicas matemáticas, estatísticas e computacionais para a previsão de movimentos futuros no mercado financeiro a partir de bases de dados. Dentre as técnicas computacionais utilizadas, destacam-se as Redes Neurais Artificiais, uma das técnicas mais difundidas da Inteligência Artificial.

A rede neural artificial é um sistema de processamento de informação que possui certas características de performance em comum com as redes neurais biológicas [FAU 94]. Ela é a técnica mais indicada para este tipo de aplicação por trabalhar de uma forma diferente dos tradicionais programas computacionais. Para um programa de computador realizar uma tarefa, o programador deve antecipar todas as condições de entrada de dados para que o programa possa chegar a uma solução. Mas, em certos casos, não é possível prever exatamente todas as situações possíveis, como no caso do mercado financeiro. Já a rede neural tenta simular o funcionamento do cérebro humano, adquirindo conhecimento para a solução de um determinado problema através de um processo de aprendizagem. Além disso, sua habilidade de lidar com dados ruidosos, incompletos ou imprecisos, e de prever sistemas não-lineares torna a aplicação de redes neurais no mercado financeiro bastante eficiente, pelo fato deste mercado ser um sistema não-linear visto que ele sofre a influência de fatores políticos, econômicos, entre outros.

Existem hoje, na literatura científica, vários trabalhos utilizando redes neurais artificiais no mercado financeiro obtendo ótimos resultados. Dentre as principais aplicações nesta área, destaca-se a utilização de redes neurais no processo de previsão dos retornos de ativos financeiros (ações, índices de ações, opções...) [LAW 97, LAZ 00, FRE 01, CAR 01, entre outros]. Estes trabalhos são apenas alguns dos exemplos que demonstram a superioridade das redes quando comparadas com modelos tradicionalmente utilizados neste campo de aplicação. O bom desempenho das redes neurais nestes trabalhos as tem tornado uma ferramenta atrativa para aplicação em

vários processos relacionados ao comportamento do mercado financeiro tais como o processo de seleção de carteiras.

Ao realizar um investimento em uma carteira, o investidor espera que o seu retorno seja grande o suficiente para compensar os riscos que está correndo. Ou seja, seu objetivo é tentar minimizar o risco e maximizar o retorno esperado.

A utilização de redes neurais artificiais pode vir a auxiliar o investidor a formar uma carteira ótima que, segundo Bernstein (2000), é a combinação de ativos que oferecem o retorno esperado mais alto em um determinado nível de risco. As redes neurais ajudam no processo de seleção das ações que farão parte da carteira, a partir da previsão de seus preços futuros. Através da estimativa do valor futuro da ação, o investidor adquire maior segurança na tomada de decisão.

A área das finanças caracteriza-se como uma das principais áreas de aplicação de redes neurais artificiais em negócios, por tratar-se de um ramo onde as decisões são baseadas em informações com um alto grau de variabilidade e incerteza [SCH 03]. As redes neurais auxiliam esse processo pela sua capacidade de análise de padrões. Embora métodos estatísticos convencionais possam ser utilizados para prever as tendências do mercado financeiro, as redes neurais são bem mais apropriadas para lidar com oscilações e perturbações.

As redes neurais podem facilitar bastante o trabalho de um investidor de ações e, ao mesmo tempo, aumentar a possibilidade de ganhos, que é hoje um dos maiores objetivos não só do investidor, mas também de todo ser humano.

## **1.2 Objetivos do trabalho**

### **1.2.1 Objetivo Geral**

Desenvolver uma rede neural artificial capaz de realizar a previsão dos preços futuros de ações. Os resultados fornecidos pela rede servirão de base para um processo de seleção das ações mais indicadas para fazer parte de uma determinada carteira.

### **1.2.2 Objetivos Específicos**

- a) Realizar um estudo sobre o processo de seleção de carteiras;
- b) Realizar um estudo sobre redes neurais artificiais, analisando a aplicabilidade desta técnica no processo de seleção de carteiras;
- c) Analisar e coletar os dados que serão utilizados na alimentação da rede neural;
- d) Selecionar a rede neural que melhor se adapta ao problema proposto;
- e) Implementar a rede neural;
- f) Validar o modelo implementado.

## **1.3 Justificativa**

Ao longo dos anos vem se tentando, de várias maneiras, realizar a previsão dos preços futuros das ações (previsão das tendências de alta ou baixa). A utilização das redes neurais artificiais vem entrando neste mercado com muita força visando facilitar o trabalho do investidor, auxiliando-o na tomada de decisão [FRE 99].

O processo de seleção de carteiras é um problema não linear que envolve um espaço de soluções grande demais para se pesquisar todas as soluções possíveis. Além disso, os dados relacionados a este problema apresentam ruídos e podem, algumas vezes, ser incompletos.

Antes do desenvolvimento tecnológico, as pessoas realizavam operações no mercado de ações baseadas na intuição. À medida que o nível de investimento e negociação aumentava, crescia também a procura por ferramentas e métodos que pudessem aumentar os ganhos, diminuindo os riscos.

Os modelos atualmente utilizados na avaliação de ações e na seleção de carteiras têm tido seus resultados questionados nos últimos anos [CAR 01]. Vários estudos mostram que tais modelos nem sempre geram carteiras consideradas eficientes. O principal modelo utilizado é o de Markowitz (1959) o qual, apesar de ser bastante difundido, apresenta algumas limitações.

O avanço da tecnologia computacional possibilitou o desenvolvimento de redes neurais que conseguiram superar as dificuldades encontradas por outros métodos já utilizados na tentativa de realizar previsões no mercado financeiro. Os modelos neurais procuram aproximar o processamento dos computadores dos processos cerebrais, apresentado um grau de interconexão similar à estrutura do cérebro, tornando-os capazes de realizar abstrações tais como aprender [MAR 02]. Esta característica das redes possibilita a extração de regras sem que estas estejam formalizadas explicitamente, além da descoberta de relações não-lineares nos dados de entrada.

Considerando as informações aqui apresentadas e os diversos trabalhos já realizados, acredita-se que a rede neural artificial é a ferramenta ideal para modelar sistemas dinâmicos não-lineares tais como o mercado de ações.

## **1.4 Organização do trabalho**

O presente trabalho está estruturado em cinco capítulos.

No segundo capítulo é apresentada uma visão geral sobre o processo de seleção de carteiras, destacando o modelo desenvolvido por Markowitz e o problema da precificação de ações.

No terceiro capítulo apresenta-se uma introdução às redes neurais. Neste capítulo faz-se uma descrição dos principais tipos de redes e métodos de aprendizagem, apresentando também as principais etapas do projeto de uma rede neural, suas principais vantagens e áreas de aplicação.

A descrição do modelo implementado pode ser vista no quarto capítulo, juntamente com a metodologia empregada e os resultados obtidos.

No quinto capítulo encontram-se as conclusões e a contribuição do trabalho desenvolvido, além de sugestões para o desenvolvimento de trabalhos futuros.

## 2 SELEÇÃO DE CARTEIRAS

### 2.1 Introdução

Ao realizar um investimento no mercado de ações, o investidor pode simplesmente investir em certas ações isoladamente, ou pode investir em carteiras diversificadas, isto é, em uma combinação de ativos, de acordo com seus interesses. A principal vantagem da escolha pela formação de carteiras diversificadas está na possibilidade de ter acesso a alternativas de investimento com níveis de risco relativamente reduzidos, se comparados às ações isoladamente.

Para realizar um investimento em uma carteira de ações deve-se primeiro realizar um processo de seleção de carteiras. “O processo de seleção de carteiras é o processo pelo qual se escolhem, dentre várias alternativas de investimento, os ativos que comporão a carteira com melhor relação risco-retorno, dado o perfil do investidor” [CAR 01]. O objetivo principal de um investimento em uma carteira de ações é obter o máximo de rendimento possível, dadas as restrições de tal carteira [BER 00].

No processo de seleção de carteiras as perspectivas de risco e retorno de vários títulos são analisadas com o objetivo de formar um conjunto de forma que se consiga, através da diversificação, reduzir o risco de uma carteira a níveis algumas vezes menores que o risco do investimento mais seguro que participa da carteira. Pode-se através da diversificação do investimento, ou seja, combinação de ativos de risco com aqueles livres de risco, aumentar o retorno esperado mantendo o risco a níveis iguais ou menores que o risco individual de cada ativo.

O retorno esperado sobre a carteira é a média ponderada dos retornos esperados de cada componente da carteira [BER 00]. Considerando o retorno esperado de uma ação como o retorno que um indivíduo espera que esta ação possa proporcionar no próximo período [ROS 95]. Já o risco de um investimento é a probabilidade de se ganhar menos que o esperado [GON 02], é a possibilidade de algo dar errado. Este risco existe devido a incertezas da atividade comercial, inflação, supervalorização, crises do crédito e crises no câmbio exterior. Na teoria da carteira o risco é geralmente medido como o desvio entre o retorno real e o esperado [BER 00].

Existem dois tipos de risco no mercado financeiro:

a) O risco diversificável (também conhecido como risco não-sistemático) o qual afeta apenas uma empresa ou, na pior das hipóteses, um pequeno grupo de empresas, ou seja, afeta um único ativo ou um pequeno grupo de ativos. Ele deve ser considerado para cada empresa ou ativo separadamente. Este tipo de risco pode ser eliminado por diversificação em uma carteira ampla [ROS 95].

b) O risco não-diversificável (também conhecido como risco sistemático, ou risco de mercado) o qual refere-se a acontecimentos que afetam o mercado como um todo [GON 02]. Este risco afeta um grande número de ativos, cada um deles com maior ou menor intensidade.

A diversificação proporcionada pela carteira pode reduzir e até mesmo eliminar somente o primeiro risco, o diversificável. O mesmo não acontece com o risco sistemático.

A composição de uma carteira pode variar muito devido ao fato de existirem diferentes tipos de investidores que desejam assumir diferentes riscos e retornos. Segundo Bernstein (2000), a decisão fundamental sobre alocação de ativos (quais ações incluir na carteira) é uma das mais importantes em todo o processo de investimento. Esta escolha tem mais influência sobre os retornos de uma carteira composta do que qualquer outra decisão particular [BER 00].

## 2.2 Modelo de Markowitz

De acordo com Cartacho (2001), os trabalhos de Markowitz são considerados o marco inicial das pesquisas nesta área e contêm os princípios básicos da formação de carteiras. Segundo Castro e Gonçalves (2002), o processo de seleção de uma carteira, na visão de Markowitz, pode ser dividido em dois estágios: o primeiro que começa com a observação e a experiência do investidor e termina com suposições sobre a performance futura dos negócios avaliados; e o segundo estágio que começa com as suposições relevantes para os desempenhos futuros e termina com a escolha da carteira de ações. O presente trabalho concentra-se no segundo estágio deste modelo.

A idéia básica do modelo de Markowitz é que os únicos fatores que devem ser levados em consideração na seleção de uma carteira são: o retorno esperado da carteira e o risco calculado pela variância deste retorno [CAS 02, JUD 03], sendo o retorno o fator desejável pelo investidor e o risco o fator indesejável. A contribuição fundamental deste modelo foi a distinção entre a variabilidade do retorno de um ativo financeiro e o seu impacto no risco de uma carteira de investimentos.

A variância da carteira de ações é uma função de correlação nos retornos entre os pares de ativos [DAM 97]. Quanto maior a correlação em retornos entre cada dois ativos, menor será o benefício decorrente da diversificação. Por outro lado, se a correlação for menor do que 1, haverá algum benefício proporcionado pela diversificação. Ou seja, quando dois ou mais ativos pouco relacionados compõem uma carteira, consegue-se um risco menor que a média ponderada dos riscos individuais [GON 02]. À medida que o coeficiente de correlação decresce, também decresce o desvio padrão da carteira [DAM 97], isto é, decresce o risco da carteira.

A chamada Moderna Teoria de Carteiras, baseada no método de Markowitz, fornece a base para a construção de carteiras eficientes (ou ótimas). Segundo Damodaran (1997), a otimização da carteira é a maximização do nível de retorno sujeito ao nível de risco máximo que o investidor está disposto a aceitar. Se o investidor especificar seu nível desejável de retorno, a carteira ótima será a que minimizar a variância sujeita a este nível de retorno. As carteiras que emergem deste processo são denominadas Markowitz.

Esta teoria analisa como o risco coletivo de uma carteira pode ser reduzido. Ela baseia-se no fato de que os riscos específicos de cada ação (os riscos diversificáveis) se compensam mutuamente e acabam desaparecendo por completo a partir de um certo



número de títulos [OIC 02]. O único risco que persiste numa carteira apropriadamente diversificada é a combinação dos riscos sistemáticos (não-diversificáveis) de diferentes ações. A Moderna Teoria de Carteiras tem sido utilizada para quantificar as vantagens da diversificação [OIC 02]. A diversificação do investimento é um meio de reduzir os riscos da carteira sem ocasionar perda de rendimentos.

Apesar de ser a principal referência quando se trata de seleção de carteiras, o modelo de Markowitz apresenta algumas falhas. Seu método para seleção de carteiras baseia-se em premissas que nem sempre são verificadas na realidade [CAR 01]. Para ser aplicado, este modelo necessita que alguns dados sejam fornecidos como entrada, considerando que os valores destes dados refletem bem a realidade. Logo, se os dados de entrada estiverem fora da realidade ou não estiverem disponíveis, o modelo não será capaz de gerar bons resultados.

Gonçalves (2002) alcançou bons resultados com a aplicação do modelo de Markowitz através da programação linear, com o uso de planilhas eletrônicas. Entretanto, o exemplo apresentado em seu trabalho não reflete a complexidade do mercado financeiro real. Já o resultado obtido pelo método de Markowitz no trabalho desenvolvido por Cartacho (2001) foi inferior ao resultado atingido pela rede neural desenvolvida neste mesmo trabalho; porém, quando comparado com o índice de mercado, seu resultado foi considerado melhor.

## **2.3 Precificação de Ações**

As ações são elementos de grande importância para o processo de seleção de uma carteira. Elas podem ser definidas como: “títulos nominativos negociáveis que representam, para quem as possui, uma fração do capital social de uma empresa” [UCP 03]. As ações são conversíveis em dinheiro, a qualquer tempo, pela negociação em bolsas de valores ou no mercado de ações. Elas são ativos de renda variável, ou seja, não oferecem ao investidor uma rentabilidade garantida, previamente conhecida [UCP 03].

Como já foi mencionado anteriormente, a decisão sobre quais ações incluir em uma carteira é uma das mais importantes do processo de investimento. O investidor que deseja compor uma carteira de ações deve escolher a melhor combinação de ações que minimize o risco de acordo com o retorno desejado.

As decisões do investidor são então baseadas no retorno que ele espera que sua carteira possa proporcionar no próximo período. Esta informação depende dos retornos individuais que se espera que cada ação venha a ter. Portanto, para que o investidor possa ter maior segurança no seu investimento, recomenda-se a utilização de um modelo de precificação de ações.

O modelo mais conhecido e utilizado é o CAPM (Capital Asset Pricing Model) o qual é um modelo para precificação de ativos financeiros. Ele é considerado “o tronco da teoria de precificação moderna para mercados financeiros” [CAB 02]. No entanto o CAPM, assim como o modelo de Markowitz para seleção de carteiras, baseia-se em premissas que não se aplicam na realidade [MIL 01, CAS 02], sendo esta a grande limitação deste modelo.

Uma das premissas básicas do modelo CAPM é a do mercado eficiente. A Hipótese de Mercado Eficiente (HME) afirma que os preços dos títulos não variam de forma aleatória, mas refletem todas as informações disponíveis a seu respeito [CAS 02]. Desta forma não é possível a obtenção de lucros superiores àquele lucro normal ajustado ao

risco inerente do negócio, onde o valor do título representa o valor justo ou valor intrínseco da empresa a qual o título representa [MIL 01]. Esta hipótese considera que a informação é amplamente acessível e barata para os investidores, e que todos os dados relevantes e suscetíveis de averiguação são refletidos nos preços dos valores mobiliários.

Existem três níveis de eficiência do mercado. No primeiro nível, denominado forma fraca de eficiência, os preços refletem toda a informação contida no registro dos preços passados. No segundo nível, denominado forma semiforte de eficiência, os preços, além de refletir seu comportamento passado, refletem toda a informação pública disponível. Finalmente no terceiro nível, conhecido como forma forte de eficiência, os preços refletem não só a informação publicada, mas também toda informação que pode ser obtida de forma privilegiada.

A teoria de um mercado eficiente não é verificada na realidade, pois, por exemplo, não há como garantir que as transações não terão custo algum para os investidores, ou que toda informação existente estará disponível também sem custo algum para todos no mercado. Além disso, ela assume que os mercados financeiros não seriam previsíveis, e que seria impossível para um investidor “bater” o mercado. Entretanto, recursos computacionais cada vez mais poderosos têm possibilitado o desenvolvimento e a aplicação de novas e sofisticadas técnicas estatísticas, matemáticas e de Inteligência Artificial na análise dos mercados financeiros na busca da previsibilidade dos preços [BRU 98].

Milani (2001) realiza um teste empírico do modelo CAPM na BOVESPA (Bolsa de Valores de São Paulo) e demonstra que este modelo não reflete adequadamente o comportamento do mercado, pois os resultados obtidos não são compatíveis com os resultados encontrados na prática. Isto ocorre devido à limitação do modelo CAPM citada anteriormente.

O modelo proposto neste trabalho como uma alternativa para a realização da precificação de ações é uma rede neural artificial. Esta escolha foi feita baseada no fato de que a precificação de ações é um problema de detecção da inter-relação entre diferentes variáveis. A rede neural é capaz de realizar a previsão do preço de qualquer ação a partir de informações históricas relacionadas ao mercado financeiro, as quais são passadas para a rede através de variáveis de entrada. A rede consegue aprender os tipos de relações entre essas variáveis e, com base nisto, fornecer uma previsão do preço da ação para o período seguinte.

Dispondo de estimativas dos retornos esperados de cada ação individualmente, os investidores podem definir qual a melhor combinação possível dos ativos que lhes oferece a melhor relação risco-retorno, de acordo com seu perfil. Quanto melhores forem as previsões, mais acertadas serão as alocações das ações dentro de carteiras escolhidas e melhores serão os investimentos realizados.

## 3 REDES NEURAIS ARTIFICIAIS

### 3.1 Introdução

A rede neural artificial é uma técnica de Inteligência Artificial (IA) que tenta simular em máquinas (computadores) o funcionamento do cérebro humano, de uma maneira simplificada. Ela é capaz de reconhecer padrões, extrair regularidades e detectar relações subjacentes em um conjunto de dados aparentemente desconexos. Além disso, ela apresenta habilidade de lidar com dados ruidosos, incompletos ou imprecisos, e de prever sistemas não lineares, o que torna a sua aplicação no mercado financeiro bastante eficiente.

“A Inteligência Artificial é um termo que abrange muitas definições” [TURBAN apud DAN 97]. Ela pode ser definida como a área das ciências da computação que visa o projeto de sistemas inteligentes, ou seja, sistemas que tentam emular algum tipo de inteligência, semelhante à de um ser humano, em termos de processos computacionais. Os sistemas inteligentes são assim denominados por exibirem características que associamos ao comportamento inteligente de um ser humano, como por exemplo: percepção, aprendizagem, raciocínio, comunicação e atuação em ambientes complexos [NIL 98].

Uma rede neural, segundo Haykin (2001), pode ser definida como,

um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela assemelha-se ao cérebro em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurônios (os pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

As redes neurais são formadas por neurônios e conexões entre eles. O neurônio (Figura 3.1) representa uma região onde informações são processadas. Seus três elementos básicos são: os pesos sinápticos, a função de soma e a função de transferência.

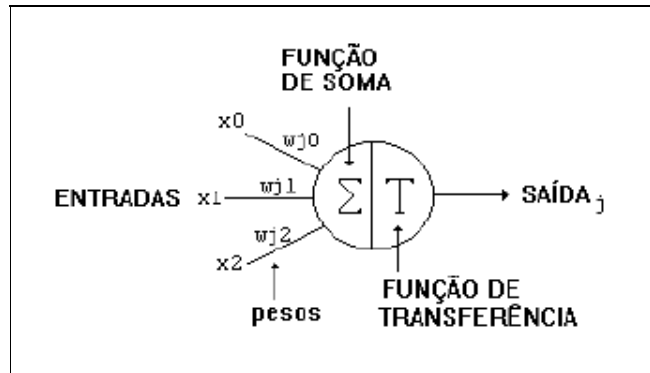


Figura 3.1: Neurônio Artificial  
FONTE: Tafner (1998)

As conexões entre os neurônios, denominadas pesos sinápticos, são responsáveis pelo armazenamento das informações. Além disso, elas definem o efeito que a saída de um neurônio exerce sobre a entrada do neurônio seguinte. Os pesos sinápticos são de grande importância para uma rede neural, pois determinam toda a manipulação de valores da rede [ALV 01].

A função de soma processa os estímulos ponderados pelos respectivos pesos, ou seja:

$$x_j = \sum_i w_{ij} y_i \quad (3.1)$$

onde  $y_i$  é a saída gerada por cada neurônio da camada anterior.

Já a função de transferência, também chamada de função de ativação, limita a amplitude do intervalo do sinal de saída do neurônio para algum valor finito, geralmente no intervalo normalizado  $[0,1]$  ou  $[-1,1]$ .

$$y_j = f(x_j) \quad (3.2)$$

Dentre as principais funções de ativação utilizadas, os tipos básicos são:

a) Função Degrau. É o tipo mais simples de função de ativação. Sua resposta pode assumir dois valores: 0 ou 1, como demonstrado a seguir:

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases} \quad (3.3)$$

b) Função Linear. Um exemplo de função linear pode ser descrito como:

$$f(x) = \begin{cases} 1, & x \geq \frac{1}{2} \\ x, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & x \leq -\frac{1}{2} \end{cases} \quad (3.4)$$

onde  $(-\frac{1}{2}$  e  $\frac{1}{2})$  é o intervalo que define a saída linear e 0 e 1 são os limites mínimo e máximo da função.

c) Função Sigmóide. Esta função assume valores em um intervalo contínuo entre 0 e 1. Sua fórmula é dada por:

$$f(x) = \frac{1}{1 + \exp(-\alpha x)} \quad (3.5)$$

na qual  $\alpha$  determina a inclinação da função.

Além dos três elementos básicos já citados, o neurônio pode ainda apresentar um bias que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação [HAY 01]. O termo bias age como um peso extra nas conexões das unidades cuja entrada é sempre um [FAU 94].

Usualmente as redes neurais apresentam três níveis de camadas de neurônios (Figura 3.2):

- uma camada de entrada: onde os padrões são apresentados à rede;
- uma camada de saída: onde o resultado é apresentado;
- camadas intermediárias ou ocultas: onde é feita a maior parte do processamento, através das conexões ponderadas. Elas situam-se entre a camada de entrada e a camada de saída e podem ser consideradas como extratoras de características.

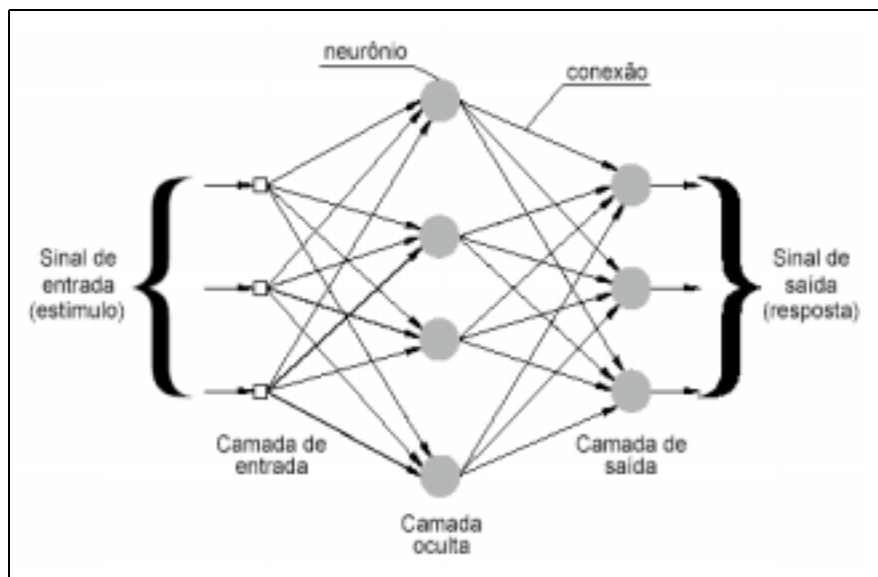


Figura 3.2: Estrutura de uma Rede Neural

FONTE: Alves (2001)

A camada de entrada, na verdade, não é formada por neurônios reais, pois eles não realizam nenhum processamento. Eles simplesmente distribuem os valores das entradas da rede para os neurônios da primeira camada oculta. Já a camada intermediária tem a função de processar a informação provida da camada de entrada. Ela também pode ser denominada camada oculta, pois sua saída não é conhecida pelo usuário.

O tipo de conexão, número de camadas de neurônios e o tipo de treinamento são os aspectos que diferem os tipos de redes neurais. Cada um é mais adequado para determinado tipo de tarefa. Os diferentes tipos de conexões entre os neurônios de uma rede determinam a topologia (ou arquitetura) desta rede. As principais topologias de rede são descritas a seguir.

a) Redes alimentadas adiante (feedforward) – os neurônios estão dispostos em camadas conectadas por pesos unidirecionais na direção entrada → saída, ou seja, as conexões ocorrem apenas entre camadas diferentes e subsequentes. Esta estrutura é totalmente conectada uma vez que todas as saídas dos neurônios de uma camada são conectadas com as entradas de todos os neurônios da camada seguinte.

b) Redes recorrentes – um neurônio pode receber entradas de qualquer outra camada da rede. Destas fazem parte as redes com realimentação nas quais os neurônios da entrada recebem sinais vindos diretamente dos neurônios da saída.

A rede neural deve ter a capacidade de generalização, ou seja, ela deve ser capaz não apenas de classificar as entradas para as quais ela recebe treinamento, mas também de generalizar e classificar entradas que não tenham sido apresentadas. Isto é possível graças a um processo de aprendizagem ao qual a rede é submetida. Esta propriedade permite que a rede encontre respostas corretas mesmo quando os dados disponíveis para as entradas estão incompletos ou danificados.

O desenvolvimento de uma rede neural ainda é um processo de tentativa e erro. A seleção da rede envolve a escolha da topologia da rede (ou arquitetura), da função de transferência e do algoritmo de aprendizagem.

### **3.2 Aprendizagem**

A propriedade mais importante das redes neurais é a habilidade de aprender de seu ambiente e com isso melhorar seu desempenho. Isto é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

O processo de aprendizagem nas redes neurais acontece internamente por meio do ajuste dos pesos sinápticos das conexões durante a exposição dos exemplos, em resposta à quantidade de erros gerados pela rede. Ou seja, a rede neural é capaz de modificar-se em função da necessidade de aprender a informação que lhe foi apresentada [TAF 95].

As redes neurais são treinadas para aprender a partir dos dados de entrada. Assim como o cérebro humano, elas aprendem a partir de experiências e não através de programação. Por este motivo, deve-se tomar bastante cuidado com a formação do conjunto de treinamento. Este conjunto deve ser gerado a partir de dados históricos, ou seja, a partir de experiências e fatos ocorridos no passado.

Como já mencionado anteriormente, a rede deve ser capaz de generalizar. Mas, ao mesmo tempo, deve-se tomar cuidado para que não aconteça um “supertreinamento” e memorização dos dados. Se uma rede neural é submetida a um supertreinamento, ela perde a capacidade de reconhecer padrões fora do conjunto de treinamento. Para evitar esta situação deve-se ter um conjunto de teste com dados diferentes do conjunto de treinamento, e a rede deve ser capaz de classificá-los corretamente, provando assim sua flexibilidade e capacidade de generalização.

Os três principais paradigmas de aprendizagem são apresentados a seguir:

- a) Aprendizagem supervisionada (ou aprendizagem com professor), quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada. O ajuste dos pesos ocorre quando o sistema compara a saída da rede com a resposta desejada previamente conhecida.
- b) Aprendizagem não-supervisionada (ou aprendizagem sem professor), quando não existe um agente externo indicando a resposta desejada para os padrões de entrada. A rede neural utiliza os neurônios como classificadores, e os dados de entrada como elementos de classificação. Esse tipo de rede trabalha essas entradas e se organiza de modo a classificá-las mediante algum critério de semelhança.
- c) Aprendizagem híbrida, que mescla os conceitos apresentados acima. Parte dos pesos é determinada através da aprendizagem supervisionada, enquanto outros são obtidos através da aprendizagem não-supervisionada.

Em um processo de aprendizagem, os pesos dos neurônios são ajustados através de um algoritmo de aprendizagem. O algoritmo de aprendizagem é um conjunto pré-estabelecido de regras bem-definidas para resolução de um problema de aprendizagem [HAY 01]. Ele tem como objetivo encontrar pesos para a rede que permitam que esta gere saídas compatíveis com as desejadas. Os algoritmos existentes diferem entre si pela forma como ocorre o ajuste dos pesos sinápticos dos neurônios, ou seja, pela regra de aprendizagem adotada.

De acordo com Haykin (2001), existem cinco regras básicas de aprendizagem através das quais os pesos sinápticos de uma rede podem ser ajustados: aprendizagem por correção de erro, baseada em memória, hebbiana, competitiva e aprendizagem de Boltzmann.

#### *Aprendizagem por correção de erro*

A aprendizagem por correção de erro é baseada no paradigma de aprendizagem supervisionada no qual a saída desejada para cada padrão de entrada é fornecida para a rede. O sinal de saída gerado pela rede, representado por  $y_k(n)$ , é comparado com a resposta desejada, representada por  $d_k(n)$ , produzindo um sinal de erro  $e_k(n)$ . Este erro é dado por:

$$e_k(n) = d_k(n) - y_k(n) \quad (3.6)$$

O sinal de erro é então utilizado para ajustar os pesos das conexões com o objetivo de aproximar o sinal de saída  $y_k(n)$  da resposta desejada  $d_k(n)$ , reduzindo o erro  $e_k(n)$ .

#### *Aprendizagem baseada em memória*

Nesta regra de aprendizagem, todas as experiências passadas são armazenadas em uma grande memória de exemplos de entrada-saída classificados corretamente:  $(x_i, d_i)$ , onde  $i = 1..n$ ,  $x_i$  representa um padrão de entrada e  $d_i$  representa a resposta desejada correspondente [HAY 01]. Quando se deseja classificar um vetor de teste  $X$  (não visto antes), o algoritmo responde buscando e analisando os dados de treinamento em uma “vizinhança local” de  $X$ .

Esta regra envolve dois ingredientes essenciais, e a forma como eles são definidos é que vai diferenciar os algoritmos de aprendizagem baseada em memória entre si. Os ingredientes são: o critério utilizado para definir a vizinhança local do vetor de teste  $X$  e a regra de aprendizagem aplicada aos exemplos de treinamento em uma “vizinhança local” de  $X$  [HAY 01].

### *Aprendizagem hebbiana*

O postulado de aprendizado de Hebb é a mais antiga regra de aprendizagem existente [HEBB apud DAN 97]. O princípio básico desta regra é: se dois neurônios em ambos os lados de uma conexão são ativados simultaneamente, então a força desta conexão é seletivamente aumentada. A forma mais simples de aprendizagem hebbiana é descrita por [HAY 01]:

$$w_{jk}(n+1) = w_{jk}(n) + \eta y_k(n)x_j(n) \quad (3.7)$$

onde  $w_{jk}$  é o peso sináptico do neurônio  $k$ ,  $x_j$  e  $y_k$  são os sinais pré-sináptico e pós-sináptico, respectivamente, deste peso, e  $\eta$  é uma constante positiva que define a taxa de aprendizagem.

Segundo Dandolini (1997), uma vantagem desta regra é que a aprendizagem é feita localmente, ou seja, a mudança nos pesos depende somente da ativação dos dois neurônios conectados pelo peso, o que simplifica bastante a complexidade da aprendizagem.

### *Aprendizagem competitiva*

Na regra de aprendizagem competitiva, os neurônios da camada de saída competem entre si para se tornarem ativos, considerando que somente um neurônio pode estar ativo em um determinado instante. Este fenômeno é conhecido como ‘winner-take-all’, isto é, ‘o vencedor leva tudo’.

Para um neurônio ser o vencedor, isto é, estar ativo, seu campo local induzido  $v_k$  para um padrão de entrada  $x$  deve ser o maior dentre todos os neurônios da rede [HAY 01]. Quando isto acontece, o sinal de saída  $y_k$  deste neurônio é igual a um. Caso contrário o sinal de saída é colocado em zero, como demonstrado a seguir:

$$y_k = \begin{cases} 1 & \text{se } v_k > v_j \text{ para todos } j, j \neq k \\ 0 & \text{caso contrário} \end{cases} \quad (3.8)$$

Nesta regra, a variação  $\Delta w_{jk}$  que é aplicada ao peso  $w_{jk}$  é definida por:

$$\Delta w_{jk} = \begin{cases} \eta(x_j - w_{jk}) & \text{se o neurônio } k \text{ vencer a competição} \\ 0 & \text{se o neurônio } k \text{ perder a competição} \end{cases} \quad (3.9)$$

onde  $\eta$  é a taxa de aprendizagem. Como efeito desta regra de aprendizagem, o vetor de peso  $w_k$  é movido na direção do padrão de entrada  $x$ , a partir do neurônio vencedor  $k$ .

Este tipo de aprendizagem é adequado para descobrir características nos dados de entrada que podem ser utilizadas para agrupar padrões similares.

### *Aprendizagem de Boltzmann*

A regra de aprendizagem de Boltzmann é um algoritmo de aprendizagem estocástico que realiza o ajuste dos pesos baseando-se na probabilidade e na mecânica estatística. A rede neural que utiliza esta regra é denominada máquina de Boltzmann [HAY 01].

Os neurônios nesta máquina formam uma estrutura recorrente e podem assumir dois estados: ligado (+1) ou desligado (-1). Os estados de cada neurônio na máquina determinam o valor de uma função de energia que caracteriza esta rede:



$$E = -\frac{1}{2} \sum_j \sum_k w_{jk} x_j x_k \quad j \neq k \quad (3.10)$$

onde  $x_j$  é o estado do neurônio  $j$  e  $w_{jk}$  é o peso sináptico entre os neurônios  $j$  e  $k$ . A restrição  $j \neq k$  é para garantir que nenhum dos neurônios tenha auto-realimentação. O equilíbrio é alcançado quando esta função de energia alcança um mínimo.

A máquina escolhe um neurônio ao acaso e, em um determinado passo do processo de aprendizagem, troca seu estado de  $+x_k$  para  $-x_k$ , a uma pseudotemperatura  $T$ , com probabilidade:

$$P(x_k \rightarrow -x_k) = \frac{1}{1 + \exp(-\Delta E_k / T)} \quad (3.11)$$

onde  $\Delta E_k$  é a variação de energia resultante da troca.

Os neurônios desta rede podem ser visíveis ou ocultos e o sistema é operado em dois modos: condição presa, no qual os neurônios visíveis estão presos em estados específicos determinados pelo meio; e condição de operação livre, onde os neurônios visíveis e ocultos podem operar livremente [HAY 01]. Permite-se que a máquina opere em ambos os modos até que o equilíbrio seja alcançado. Quando isto acontece, os pesos  $w_{jk}$  nas conexões entre os nós  $j$  e  $k$  são ajustados com base na diferença entre as probabilidades  $p_{jk}$  (P) e  $p_{jk}$  (L).  $p_{jk}$  (P) é a probabilidade dos elementos  $j$  e  $k$  estarem na condição presa, e  $p_{jk}$  (L) é a probabilidade de ambos estarem na condição de parada livre [DAN 97].

### 3.3 Rede Neural Backpropagation

#### 3.3.1 Introdução

A rede neural comumente denominada backpropagation é na verdade uma rede neural feedforward (Figura 3.3), multicamada, treinada pelo algoritmo backpropagation. Segundo Lawrence (1997), esta rede é a mais utilizada no desenvolvimento de redes neurais para o mercado financeiro. O desenvolvimento deste algoritmo de aprendizagem foi um dos marcos mais importantes das pesquisas em redes neurais artificiais [DAN 97]. Este trabalho foi o primeiro que possibilitou o ajuste dos pesos em redes multicamadas feedforward, abrindo caminho para a elaboração de redes neurais mais genéricas. Devido a grande popularidade do método backpropagation, seu nome é utilizado para denominar as redes que o utilizam no seu treinamento.

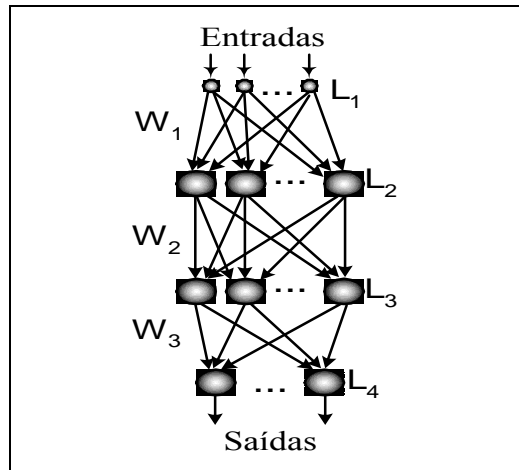


Figura 3.3: Rede Neural Feedforward

A atratividade do método backpropagation vem do conjunto de equações bem definidas e explícitas para correção dos pesos da rede [DAN 97]. Este algoritmo consiste em realizar a retropropagação do erro gerado na comparação entre a saída da rede e a saída desejada com o objetivo de minimizar o erro total da saída gerada pela rede. O treinamento de uma rede através deste algoritmo envolve três etapas: a propagação dos dados da camada de entrada para a camada de saída da rede, o cálculo e a retropropagação do erro gerado pela rede, e o ajuste dos pesos [FAU 94].

Na primeira etapa, estímulos de entrada são apresentados à rede e as ativações fluem até chegarem à camada de saída, gerando um resultado. Já na segunda e terceira etapas, o resultado obtido pela rede é comparado com a saída desejada e o erro gerado é computado para as unidades de saída. Os pesos conectados às unidades de saída são então ajustados para reduzir este erro. Em seguida, o erro da camada de saída é utilizado para derivar estimativas de erro para as unidades da(s) camada(s) oculta(s), para que o erro seja então propagado para trás até a conexão da camada e entrada. O fluxo de informação deste processo é ilustrado de uma forma resumida na Figura 3.4.

O método backpropagation atualiza os pesos incrementalmente, depois de analisar cada par entrada-saída. Depois da apresentação de todos os pares entrada-saída diz-se que uma época foi concluída [RIC 94]. Este treinamento, em geral, requer muitas épocas.

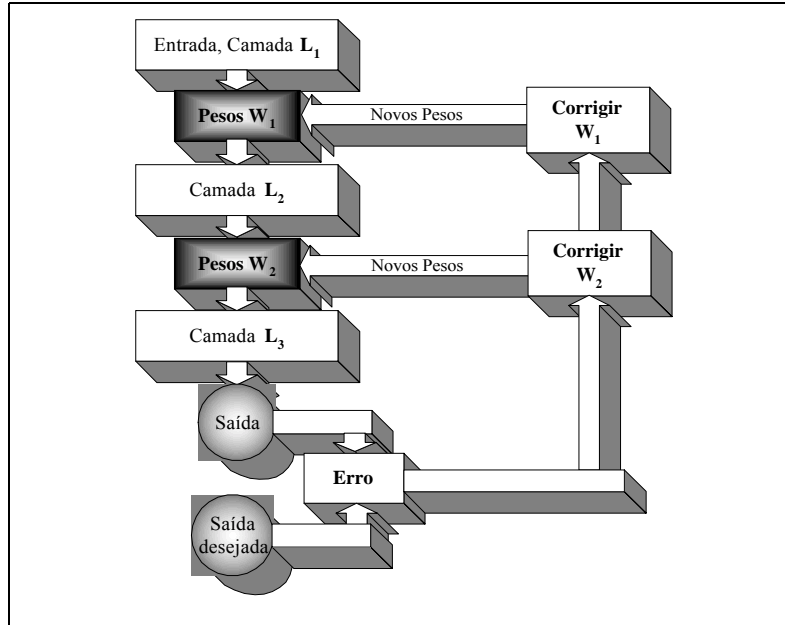


Figura 3.4: Correção dos pesos por backpropagation

### 3.3.2 Ajuste dos Pesos

Uma regra de aprendizagem por correção de erro, a regra delta, é onde está baseado o ajuste dos pesos realizado pelo algoritmo backpropagation. Este ajuste é feito baseado na retropropagação do erro através da qual o erro gerado pelos neurônios na camada de saída é distribuído para os demais neurônios da rede.

Mesmo conhecendo o erro global da rede, não é possível determinar os pesos exatos para poder corrigi-lo. Entretanto, com base nesta informação, pode-se estabelecer a direção na qual os pesos devem ser ajustados para minimizar o erro quadrado total da saída da rede. Conhecida esta direção, é possível ajustar os pesos até que o menor erro global seja atingido. O ajuste de um peso  $w_{ij}$  que define seu valor para a próxima iteração é definido por:

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (3.12)$$

A variação  $\Delta w_{ij}$  que é aplicada ao peso  $w_{ij}$  deve ser proporcional ao sinal de entrada  $x_j$ , que é definido pelos sinais de saída da camada anterior ponderados pelos pesos, e ao erro gerado na saída. Ela é dada por:

$$\Delta w_{ij}(n) = \eta \delta_j(n) x_j(n) \quad (3.13)$$

onde  $\eta$  é a taxa de aprendizagem e  $\delta_j$  é o gradiente local do erro para o neurônio  $j$ .

A taxa de aprendizagem é um valor positivo, geralmente menor do que 1, que regula a intensidade com que as atualizações dos parâmetros (pesos) serão efetuadas. Taxas muito baixas, próximas de zero, tendem a fazer com que o aprendizado seja bastante lento, porém taxas muito altas, próximas de 1, podem fazer com que a rede oscile, como se estivesse aprendendo e desaprendendo, e as vezes nem consiga chegar a um patamar aceitável de aprendizado. O valor da taxa de aprendizado não precisa permanecer fixo

durante todo o treinamento. Em algumas implementações ela pode ser adaptativa e controlada pela própria rede.

O gradiente local do erro é determinado através do método gradiente descendente. Ele é o termo responsável pela distribuição do erro da camada de saída para as camadas anteriores. O ajuste dos pesos ( $\Delta w$ ) deve ser realizado na direção contrária ao gradiente, conforme mostra o Gráfico 3.1. Se o peso  $w(n)$  (valor do peso na iteração  $n$ ) está a esquerda do erro mínimo, o ajuste  $\Delta w$  deve ser positivo para que  $w(n+1)$  (valor do peso da próxima iteração) esteja mais próximo do valor de  $w$  que minimiza o erro. Por outro lado, se o peso  $w(n)$  está a direita do erro mínimo, o ajuste  $\Delta w$  deve ser negativo.

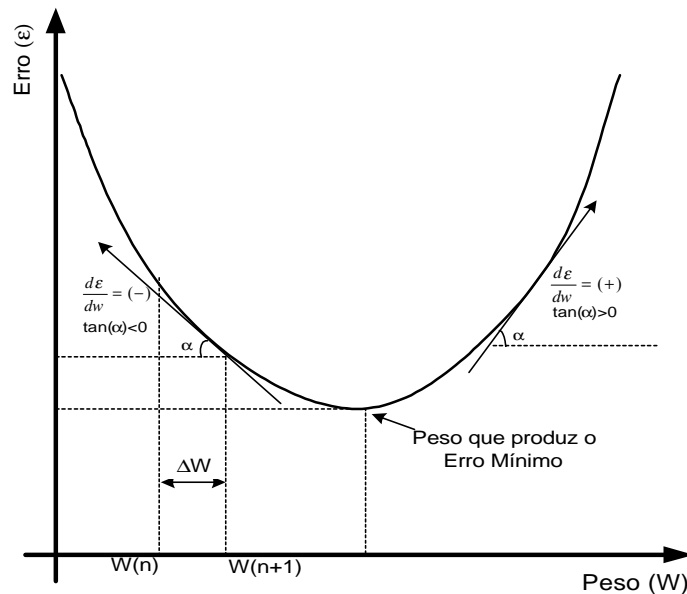


Gráfico 3.1: Função erro para um único peso

O ajuste dos pesos para os neurônios da camada de saída em um algoritmo backpropagation é diferenciado do ajuste dos pesos para os neurônios da camada oculta. A seguir será demonstrado o processo de ajuste dos pesos para cada caso separadamente.

#### *Ajuste dos pesos na camada de saída*

Devido ao fato da aprendizagem ser do tipo supervisionada, o resultado desejado para a camada de saída  $l$  é fornecido para a rede. Com isto, pode-se fazer uma comparação deste com o resultado obtido pela rede nesta camada, gerando um sinal de erro que é utilizado para realizar o ajuste dos pesos dos neurônios desta camada:

$$\varepsilon_{l,j}(n) = d_{l,j}(n) - y_{l,j}(n) \quad (3.14)$$

Tendo o valor do erro, seu gradiente local é definido como:

$$\delta_{l,j}(n) = \varepsilon_{l,j}(n) \frac{\partial y_{l,j}(n)}{\partial x_{l,j}(n)} \quad (3.15)$$

A expressão para o ajuste dos pesos entre a camada de saída  $l$  e a camada  $l-1$  é então determinada como:

$$\Delta w_{l-1,i,j}(n) = \eta \delta_{l,j}(n) y_{l-1,i}(n) \quad (3.16)$$

#### *Ajuste dos pesos na camada oculta*

Da mesma forma que os neurônios na camada de saída devem apresentar saídas próximas aos alvos ( $d$ ), os neurônios na(s) camada(s) oculta(s) também devem exibir saídas determinadas, contudo desconhecidas. A saída desejada para este tipo de camada não é informada para a rede. Neste ponto o algoritmo backpropagation justifica seu nome, retropropagando o erro gerado pelos neurônios na camada de saída para as camadas internas, distribuindo o erro para cada um dos neurônios nas camadas ocultas.

O gradiente local do erro para a(s) camada(s) oculta(s) é então definido como:

$$\delta_{l,j}(n) = \sum_k [\delta_{l+1,k}(n) w_{l,j,k}(n)] \frac{\partial y_{l,j}(n)}{\partial x_{l,j}(n)} \quad (3.17)$$

E o ajuste dos pesos entre os neurônios da camada  $l$  e da camada  $l-1$  é determinado pela Equação (3.16) a qual também determina este ajuste para a camada de saída.

### **3.3.3 Observações**

a) A escolha da função de transferência em uma rede neural backpropagation deve obedecer aos requisitos de continuidade, diferenciabilidade e monotonicidade. Estes requisitos são exigidos pelo algoritmo backpropagation para permitir que uma expressão analítica, para o ajuste dos pesos da rede, seja obtida.

b) Um termo extra que pode ser adicionado ao ajuste dos pesos na tentativa de melhorar performance da rede é a taxa de momentum. Esta taxa é um parâmetro de uso opcional, de valor positivo menor do que 1, cuja utilização visa imprimir uma dinâmica no treinamento tal que, eventualmente, possibilite que o algoritmo livre-se de mínimos locais durante o processo de busca pelo mínimo global. Para utilizá-la, os pesos de um ou mais padrões de treinamento anteriores devem ser salvos. Sua utilização é recomendada quando alguns dados de entrada são muito diferentes da maioria [FAU 94]. Normalmente resulta em uma aprendizagem mais rápida, mas pode causar instabilidade em alguns casos se for muito grande. Quando a taxa de momentum  $\mu$  é utilizada, a expressão para o ajuste dos pesos entre a camada de saída  $l$  e a camada  $l-1$  passa a ser definida como:

$$\Delta w_{l-1,i,j}(n) = \eta \delta_{l,j}(n) y_{l-1,i}(n) + \mu [w_{l-1,i,j}(n) - w_{l-1,i,j}(n-1)] \quad (3.18)$$

### 3.4 Projeto de uma Rede Neural

Diante de um projeto de uma rede neural, não pensamos mais em procedimentos, regras ou fórmulas algorítmicas de processamento de dados, mas sim em tipos de dados de entrada, dados de saída e tratamento de dados [TAF 95]. “O projeto de uma rede neural é baseado diretamente nos dados do mundo real, permitindo-se que o conjunto de dados fale por si mesmo” [HAY 01].

O projeto de um sistema neural consiste de diversas etapas que devem ser executadas em seqüência, de forma interativa e até mesmo com diversos ciclos de repetição. A construção do sistema começa pela identificação e coleta dos dados históricos relevantes para o problema. O passo seguinte é a preparação e adequação dos dados ao formato requerido pela rede neural, ou seja, a formatação dos dados. Neste processo cria-se uma escala, estabelecendo um novo intervalo válido dentro do qual todos os dados são colocados. Os intervalos mais utilizados são  $[-1,1]$  ou  $[0,1]$ . Este processo também é conhecido como normalização dos dados.

Após a escolha de uma representação para os dados do problema, deve-se separar os dados em dois conjuntos: o conjunto de treinamento o qual é gerado a partir de dados históricos, ou seja, a partir de experiências e fatos ocorridos no passado; e o conjunto de teste com o qual o funcionamento da rede é testado. Cada dado do conjunto de teste é apresentado uma única vez ao sistema.

O próximo passo é realizar a escolha do modelo neural a ser adotado e definir a topologia da rede. Em seguida realiza-se o desenvolvimento, treinamento e otimização do modelo, seguidos pela validação do mesmo. Na etapa de validação, faz-se uma comparação do resultado obtido pela rede com o resultado desejado.

Por fim, após a validação do modelo, realiza-se a aplicação do mesmo.

### 3.5 Vantagens

Uma das vantagens da rede neural artificial que mais se destaca, principalmente quando aplicada no mercado financeiro, é a sua capacidade de modelar e prever sistemas não-lineares. Este é o grande diferencial das redes neurais quando comparadas com outros métodos como, por exemplo, modelos estatísticos. Além desta, existem inúmeras outras vantagens atribuídas às redes neurais, dentre as quais destacam-se:

- capacidade de encontrar soluções eficientes para problemas do mundo real;
- habilidade de lidar com dados ruidosos, incompletos ou imprecisos;
- capacidade de análise e reconhecimento de padrões;
- capacidade de resolver problemas práticos sem a necessidade da definição de regras ou de modelos precisos;
- capacidade de buscar a solução através de um método próprio de treinamento e auto-aprendizado;
- execução em paralelo;
- alta capacidade de adaptação e generalização;

A principal limitação das redes neurais é não ter capacidade de explicação [DAN 07]. Os dados entram na rede e uma previsão sai, mas o tipo de relacionamento entre as

variáveis não é revelado [FRA 01], assim como detalhes de como as redes raciocinam com o dados para chegar às conclusões também não são fornecidos. Por este motivo, as redes neurais são recomendadas para serem aplicadas em áreas de conhecimento cujas teorias ainda não conseguem explicar adequadamente o comportamento dos fenômenos observados ou em áreas de grande complexidade que não necessitam de modelos precisos da realidade física do problema. Ou seja, a maior limitação das redes neurais não representa um problema para a sua aplicação no mercado financeiro.

### **3.6 Aplicações**

Mesmo com algumas restrições, a área de redes neurais tem demonstrado sua potencialidade em diversas aplicações, superando expectativas e gerando resultados até então não alcançados com qualquer outra técnica, seja computacional ou convencional.

As redes neurais artificiais podem ser treinadas para encontrar soluções, interpretar e classificar dados, reconhecer padrões, aproximar funções e prever eventos futuros. Atualmente, são inúmeras as áreas nas quais as redes neurais têm sido aplicadas, e os bons desempenhos alcançados têm incentivado pesquisadores a fazer das redes neurais uma alternativa de solução para problemas nas mais diversas áreas de atuação. Os principais domínios de aplicação das redes neurais são: processamento de imagens, setor militar, robótica, biologia e medicina, telecomunicações e o setor financeiro.

Dentre as principais aplicações das redes neurais na área de finanças destacam-se: aprovação de crédito, detecção de fraude, previsão de falência, previsão no mercado de ações (previsão do preço de ações e de índices econômicos), seleção de carteiras, análise de riscos, diagnóstico de empresas [DAN 97, FRA 01, SCH 03, entre outros].

## 4 APLICAÇÃO E RESULTADOS

### 4.1 Introdução

Neste capítulo apresenta-se o desenvolvimento de uma rede neural feedforward com algoritmo de treinamento backpropagation para realizar a previsão dos preços de três ações: Eletrobrás ON, Petrobrás PN e Vale Rio Doce PNA, no horizonte de um mês. A implementação deste modelo visa fornecer uma ferramenta de auxílio ao investidor, proporcionando-lhe maior segurança na tomada de decisão.

Depois da descrição do modelo desenvolvido, são apresentados os resultados obtidos nos treinamentos e nos testes da rede neural aplicada às ações dadas.

### 4.2 Dados

Os dados utilizados neste trabalho são reais e foram calculados a partir das cotações mensais obtidas no programa Economática\* (sendo ajustados aos dividendos e deflacionados) e organizados no Microsoft Excel para o posterior uso na rede neural. O Economática é um software que fornece um grande banco de dados a respeito dos títulos listados em diversas bolsas do mundo, sendo utilizado por analistas de investimentos como uma das ferramentas de apoio às suas decisões de alocação de recursos, pois ele disponibiliza uma vasta gama de informações de interesse destes [MIL 01]. Um exemplo de tela do Economática, apresentando as cotações da ação Petrobrás PN, pode ser visualizado na Figura 4.1.

O conjunto de dados foi dividido em duas partes: a primeira parte (de janeiro/1995 a dezembro/1998) foi utilizada no processo de treinamento da rede, a outra parte (de janeiro/1999 a dezembro/2000) foi utilizada no teste da mesma.

Antes de serem apresentados à rede, os dados ainda passaram por um pré-processamento, sendo normalizados dentro de dois intervalos. Os dados utilizados na entrada da rede foram normalizados para o intervalo  $[-1,1]$ . Já os valores das respostas desejadas para a rede foram normalizados no intervalo  $[0,1]$ . O objetivo da normalização é diminuir a influência causada por valores que se destacam excessivamente em relação aos demais, ou seja, diminuir a distância entre os valores de variáveis muito espaçadas. O tratamento dos dados incompletos também faz parte do pré-processamento. Atribuiu-se o valor zero para cada dado não disponível.

---

\* [www.economatica.com.br](http://www.economatica.com.br)



Data	Q Negs	Q Tds	Volume\$	Fechamento	Abertura	Mínimo	Máximo	Médio
08/98	22.778	73.152k	1.797M	19,274	19,43	14,217	24,643	-
10/98	16.328	43.137k	1.167M	23,707	19,47	17,38	27,176	-
11/98	11.629	34.134k	1.209M	27,074	24,181	24,181	30,662	-
12/98	8.704,00	35.645k	1.016M	21,481	27,074	19,95	28,499	-
01/99	16.720	68.832k	1.540M	21,081	22,108	13,387	32,42	-
02/99	10.894	36.848k	1.104M	23,912	21,546	21,236	27,437	-
03/99	13.886	48.083k	1.798M	35,759	24,193	24,193	38,927	-
04/99	12.858	45.020k	2.108M	40,119	35,761	35,308	42,971	-
05/99	12.627	49.057k	2.287M	37,903	40,573	33,796	41,756	-
06/99	8.004,00	27.014k	1.250M	41,028	37,772	35,117	41,899	-
07/99	8.318,00	30.815k	1.510M	37,759	41,028	37,598	42,988	-
08/99	8.830,00	32.042k	1.387M	37,203	37,598	33,621	39,064	-
09/99	8.291,00	26.847k	1.338M	40,843	37,203	37,203	45,309	-
10/99	8.244,00	34.095k	1.218M	43,474	40,843	39,445	45,388	-
11/99	8.850,00	25.732k	1.586M	53,052	43,404	43,056	57,126	-
12/99	7.888,00	24.992k	1.655M	62,054	53,258	52,575	62,728	-
01/00	8.376,00	20.203k	1.418M	54,483	62,189	54,063	63,807	-
02/00	8.012,00	20.654k	1.474M	61,311	54,483	54,25	63,03	-
03/00	7.225,00	16.980k	1.285M	63,924	61,311	60,844	68,775	-
04/00	8.621,00	24.386k	1.736M	58,174	63,853	54,171	67,714	-
05/00	8.597,00	20.433k	1.403M	58,295	58,701	55,996	60,862	-
06/00	8.514,00	21.710k	1.710M	72,549	58,968	58,834	74,164	-
07/00	15.884	25.640k	2.039M	60,848	74,413	60,822	77,075	-
08/00	28.973	55.318k	4.158M	70,445	61,183	57,928	74,33	-
09/00	14.684	21.603k	1.768M	68,915	71,595	65,075	74,536	-
10/00	15.282	26.227k	1.974M	64,024	67,298	58,757	69,20	-
11/00	11.835	20.814k	1.570M	59,856	63,758	59,478	66,921	-
12/00	15.247	23.588k	1.809M	57,254	60,474	55,448	61,388	-
01/01	16.177	25.988k	2.000M	67.828	57.279	57.279	70.41	-

Figura 4.1: Tela do Economática com as cotações da ação Petrobrás PN

### 4.3 Variáveis de entrada da rede

Uma das decisões mais importantes na construção de uma rede neural é a escolha das variáveis nas quais a rede irá se basear para poder aprender. Quando as redes neurais são utilizadas para realizar previsão financeira, dados sobre diversos indicadores econômicos são utilizados como valores de entrada da rede para que na saída obtenha-se alguma informação, como neste caso, o preço de fechamento da ação para o próximo período.

Tomando por base o trabalho desenvolvido por Cartacho (2001), para cada ação foram coletados os seguintes dados para serem utilizados como variáveis de entrada: preço de abertura, mínimo, máximo e de fechamento; volume de negociação; índice Ibovespa; taxa do dólar; taxa da poupança; taxa de juros (Selic 30 dias) e índice Dow Jones. Além destes, também foram utilizados os preços de fechamento nos períodos (t-1), (t-2) e (t-3), e as médias móveis de cinco (5), dez (10) e quinze (15) meses das cotações do preço de fechamento, formando um conjunto de dezesseis variáveis de entrada. O Gráfico 4.1 apresenta as variáveis de entrada da rede, da ação Petrobrás PN, com valores normalizados no intervalo [-1,1].

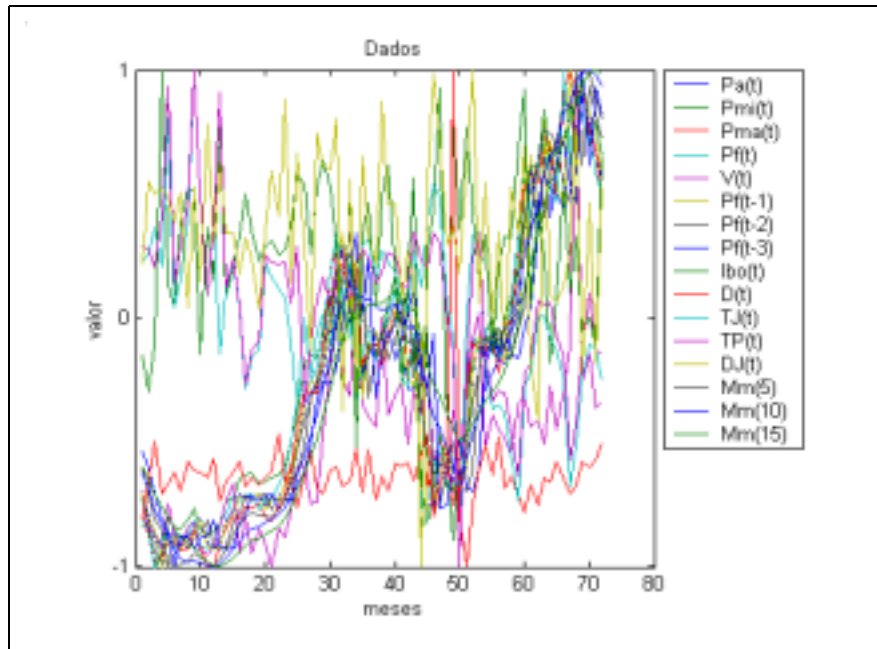


Gráfico 4.1: Variáveis de entrada (Petrobrás PN) normalizadas entre [-1,1]

A média móvel é uma ferramenta bastante utilizada em previsões no mercado financeiro. Nela os preços dos últimos  $n$  períodos são utilizados para filtrar um pouco eventuais variações excessivas de um período para outro e visualizar mais claramente a possível tendência do mercado. Utiliza-se neste trabalho a média móvel ponderada, na qual os últimos preços coletados adquirem maior importância, de acordo com a seguinte fórmula:

$$M(n) = (P[t]*n + P[t-1]*(n-1) + P[t-2]*(n-2) + \dots + P[t-n+1]*1) / (n + (n-1) + (n-2) + \dots + 1) \quad (4.1)$$

Os preços de fechamento dos períodos  $t$ ,  $(t-1)$ ,  $(t-2)$  e  $(t-3)$  foram utilizados com o objetivo de fornecer à rede uma memória de curto e médio prazo.

“O índice BOVESPA (Ibovespa) é o índice da Bolsa de Valores de São Paulo, que mede a lucratividade de uma carteira teórica de ações” [BOV 99]. Ele “retrata o comportamento dos principais papéis negociados na bolsa de valores, constituindo o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro” [BOVESPA apud VIE 00].

A taxa Selic é atualmente a taxa de referência dos juros no mercado financeiro [UCP 01].

## 4.4 Rede Neural

### 4.4.1 Introdução

A utilização de uma rede neural artificial exige que uma série de escolhas, não triviais, sejam feitas na busca de um modelo que apresente um resultado considerado satisfatório. Dentre as principais escolhas estão: a topologia da rede, o algoritmo de

aprendizagem, a função de ativação, a taxa de aprendizagem, a taxa de momentum e o número ideal de épocas.

Optou-se pelo modelo de rede neural feedforward, freqüentemente utilizado em aplicações financeiras, com algoritmo de treinamento backpropagation, no qual os pesos são ajustados baseados na regra de aprendizagem por correção de erro. A rede, modelada e executada no software Matlab, foi estruturada em três camadas, ou seja, apenas uma camada oculta foi utilizada, pois, de acordo com Fausett (1994), duas ou mais camadas ocultas podem beneficiar determinadas aplicações, mas uma única camada oculta é considerada suficiente.

A camada de entrada é formada por dezesseis neurônios, um para cada variável de entrada da rede. O número de neurônios na segunda camada foi sendo modificado ao longo dos testes. Na terceira camada há apenas um neurônio que representa a variável de saída da rede: o preço de fechamento da ação para o mês seguinte

Além do número de neurônios da camada oculta, a taxa de aprendizagem, a taxa de momentum e o número de épocas também foram sendo modificados durante a realização de inúmeros testes na busca da combinação que produzisse o melhor resultado.

A função tangente hiperbólica (tanh) foi definida como a função de ativação para os neurônios da camada oculta. Ela assume valores em um intervalo contínuo entre  $-1$  e  $1$ . Sua fórmula é dada por:

$$\tanh(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)} \quad (4.2)$$

A função de ativação aplicada ao neurônio da camada de saída é a seguinte função linear:

$$f(x) = \begin{cases} 1 & \text{se } x > m \\ \frac{x}{m} & \text{se } 0 < x \leq m \\ 0 & \text{se } x \leq 0 \end{cases} \quad (4.3)$$

onde  $m$  assume valores diferentes dependendo da combinação dos parâmetros e da ação para a qual a rede está sendo treinada.

É importante ressaltar que a combinação que resulta na melhor previsão para uma ação não é a mesma combinação que gera o melhor resultado para outra. Portanto, é necessário configurar uma rede neural com parâmetros diferentes para cada ação.

#### 4.4.2 Inicialização da Rede

Como já mencionado anteriormente, os dados utilizados na rede neural foram obtidos a partir do software Economática e posteriormente organizados no Excel. Os dois conjuntos de dados, de treinamento e de teste, são armazenados em um mesmo arquivo xls da seguinte maneira: cada linha representa um padrão de entrada (referente a um dado mês) e cada coluna representa uma variável de entrada, excluindo a última que contém a resposta desejada.

Após a leitura dos dados a partir do arquivo xls, a primeira tarefa a ser realizada é a separação dos dados de entrada da rede das respostas desejadas. Em seguida os dados de entrada são normalizados no intervalo  $[-1,1]$  e as respostas desejadas são normalizadas no intervalo  $[0,1]$ .

Antes de serem apresentados à rede, os dados são mais uma vez separados. Nesta etapa faz-se a divisão dos grupos de dados em dois conjuntos: de treinamento e de teste. O conjunto de treinamento é formado por 48 vetores de entrada, referentes aos 48 meses de janeiro de 1995 a dezembro de 1998. Já o conjunto de teste é formado por 24 vetores, referentes aos 24 meses de janeiro de 1999 a dezembro de 2000.

Antes de iniciar o processamento da rede, também é feita a escolha dos valores iniciais para os pesos, que foram definidos de forma randômica dentro do intervalo  $[-1,1]$ .

#### **4.4.3 Treinamento**

Considerando que o objetivo da rede neural neste trabalho é realizar a previsão do preço das ações para o mês seguinte, seu treinamento foi realizado da maneira descrita a seguir.

Tendo definido o número de épocas  $p$  para o treinamento, realiza-se a apresentação dos 48 pares entrada-saída de treinamento  $p$  vezes para que a rede seja treinada. Após o treinamento, é feito o teste com o primeiro padrão do conjunto de teste. Guarda-se o resultado obtido (a resposta da rede e o erro gerado) em um vetor, adiciona-se o padrão testado ao conjunto de treinamento e treina-se novamente a rede  $p$  vezes, agora com um padrão a mais. Depois deste treinamento realiza-se o teste com o próximo padrão do conjunto de teste. Adiciona-se o novo resultado ao vetor de resultados, adiciona-se o padrão ao conjunto de treinamento voltando a treinar a rede  $p$  vezes, com o novo conjunto de treinamento. Este processo repete-se até que todos os padrões de teste tenham sido testados.

Os passos para o treinamento da rede, listados anteriormente, são demonstrados no fluxograma ilustrado na Figura 4.2, de acordo com a seguinte nomenclatura:

$i = 16$  (número de variáveis de entrada);

$t = 1$  (padrão de entrada atual para teste);

$TR = 48$  (número de padrões de entrada para treinamento);

$TT = 24$  (número de padrões de entrada para teste);

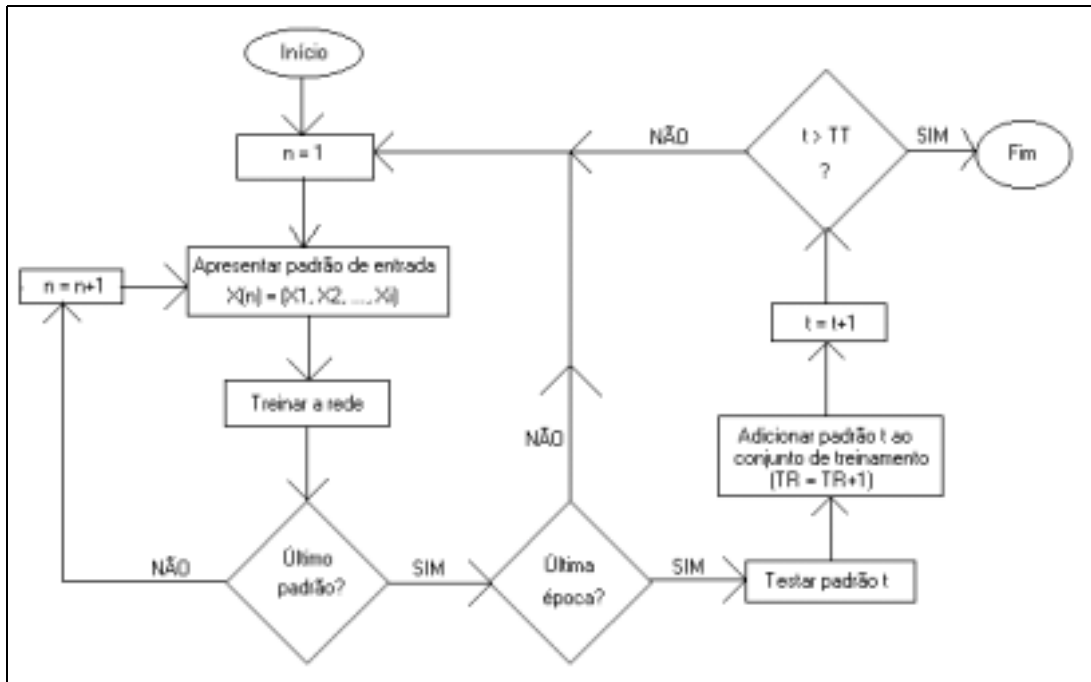


Figura 4.2: Fluxograma do processo de treinamento da rede

## 4.5 Resultados

Nesta seção são apresentados alguns resultados obtidos com a rede backpropagation aplicada às ações Eletrobrás ON, Petrobrás PN e Vale Rio Doce PNA. Além dos testes propriamente ditos (realizados com os dados de teste), foram realizados testes com os dados de treinamento para verificar se a rede realmente estava aprendendo.

Os testes foram realizados para cada ação separadamente com o objetivo de encontrar valores para os parâmetros que combinados realizassem a melhor previsão. Os parâmetros que tiveram seus valores alterados durante os testes são: o número de neurônios da camada oculta, a taxa de aprendizagem, a taxa de momentum, o número de épocas de treinamento e o valor  $m$  da função linear.

Uma boa medida de desempenho para uma rede com aprendizado supervisionado é o cálculo do erro quadrático médio, dado pela seguinte fórmula:

$$erro = \frac{1}{n} \sum_j^n (erro_j)^2 \quad (4.3)$$

Esta foi umas das medidas utilizadas para realizar a avaliação da rede desenvolvida neste trabalho. As Tabelas 4.1, 4.2 e 4.3 demonstram alguns resultados, destacando as combinações de parâmetros que geraram os menores erros para cada uma das ações. Mais resultados podem ser encontrados nos apêndices A, B e C.

Tabela 4.1: Resultados da rede neural para a ação Eletrobrás ON

Neurônios	Épocas	Momentum	Taxa de Aprendizagem	$m$	Erro Treinamento	Erro Teste
16	75	0.6	0.1	6000	0.0424	0.0072
15	100	0.1	0.8	6000	0.0937	0.0200
13	100	0.5	0.01	1000	0.0378	0.0069
19	75	0.6	0.01	5000	0.0368	0.0068

Tabela 4.2: Resultados da rede neural para a ação Petrobrás PN

Neurônios	Épocas	Momentum	Taxa de Aprendizagem	$m$	Erro Treinamento	Erro Teste
6	30	0.1	0.8	2000	0.2870	0.0128
6	30	0.8	0.8	6000	0.0895	0.0141
10	35	0.01	0.8	1500	0.3091	0.0103
8	30	0.8	0.8	6000	0.0915	0.0144

Tabela 4.3: Resultados da rede neural para a ação Vale Rio Doce PNA

Neurônios	Épocas	Momentum	Taxa de Aprendizagem	$m$	Erro Treinamento	Erro Teste
10	20	0.3	0.7	2000	0.2447	0.0187
10	20	0.5	0.5	2000	0.1152	0.0227
7	20	0.5	0.8	2500	0.2364	0.0140
10	20	0.9	0.5	2000	0.1076	0.0583

Ao avaliar a performance de uma rede neural, é aconselhável a utilização de uma segunda medida de desempenho, além do cálculo do erro. Neste trabalho, a performance da rede desenvolvida foi avaliada a partir de sua capacidade de prever os preços das ações. Além da avaliação do erro médio, foi feita a análise de gráficos com o resultado obtido pela rede e o resultado desejado (com a utilização de dados de teste e dados de treinamento).

Os Gráficos 4.2 e 4.3 apresentam o resultado obtido pela rede em comparação com o resultado desejado para a ação Eletrobrás ON.

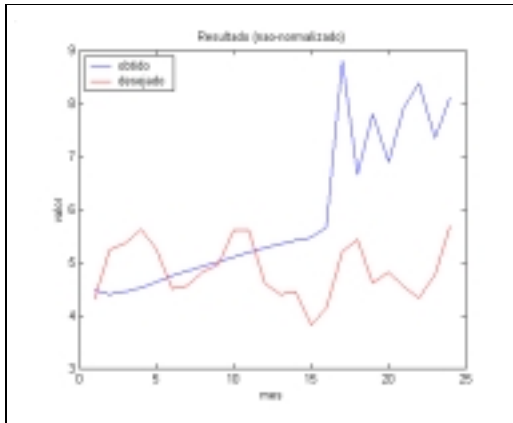


Gráfico 4.2: Resultado com dados de teste

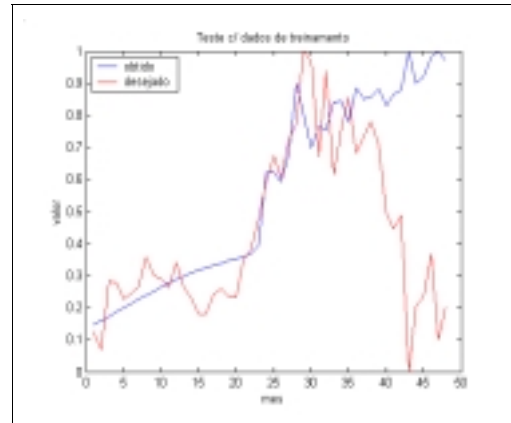


Gráfico 4.3: Resultado com dados de treinamento

Os Gráficos 4.4 e 4.5 apresentam o resultado obtido pela rede em comparação com o resultado desejado para a ação Petrobrás PN.

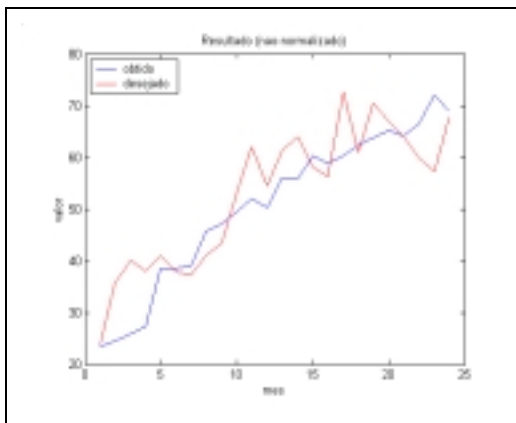


Gráfico 4.4: Resultado com dados de teste

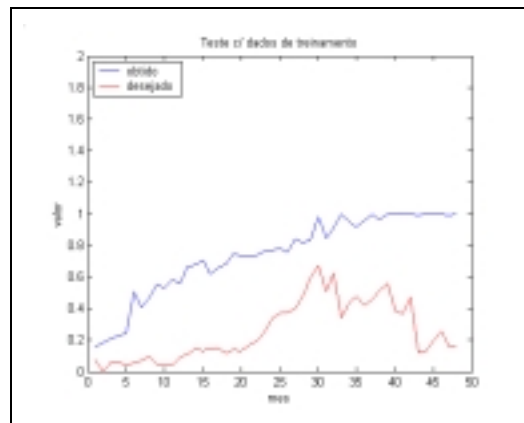


Gráfico 4.5: Resultado com dados de treinamento

Os Gráficos 4.6 e 4.7 apresentam o resultado obtido pela rede em comparação com o resultado desejado para a ação Vale Rio Doce PNA.

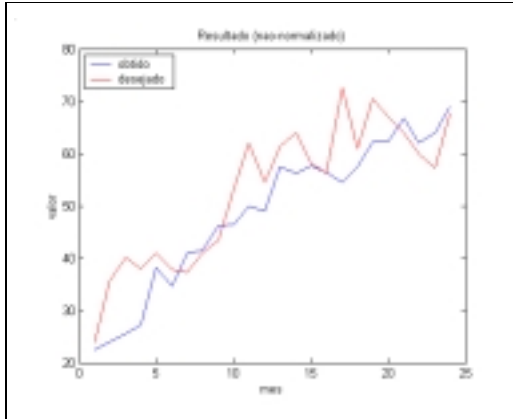


Gráfico 4.6: Resultado com dados de teste

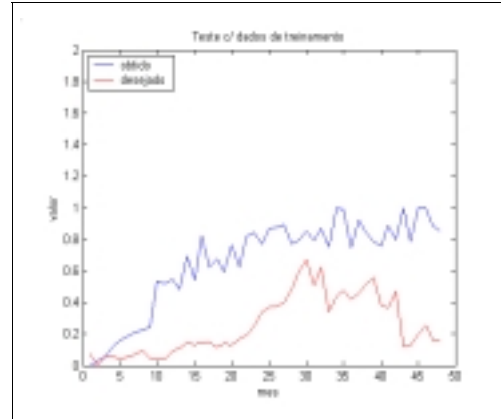


Gráfico 4.7: Resultado com dados de treinamento



## 5 CONCLUSÕES

O problema da seleção de carteiras de investimentos tem atraído o interesse tanto de simples investidores quanto de profissionais que atuam no dia a dia das instituições financeiras. A disponibilidade de recursos computacionais tem possibilitado o desenvolvimento de métodos capazes de alcançar resultados importantes, como por exemplo a capacidade de realizar previsões no comportamento do mercado de ações, estabelecendo novos rumos para pesquisas no mercado financeiro.

O presente trabalho foi motivado pela possibilidade de desenvolvimento de uma ferramenta de auxílio ao investidor de ações, que proporcionasse maior segurança na sua tomada de decisão. A escolha das redes neurais artificiais justifica-se, principalmente, pela sua capacidade de captar não-linearidades no sistema sem qualquer intervenção humana e pelo fato delas superarem as dificuldades encontradas pelos métodos tradicionalmente utilizados neste campo de aplicação.

A aplicação de redes neurais no processo de seleção de carteiras está na parte da previsão dos preços das ações (primeiro passo do processo) para que a seleção das ações ideais que farão parte da carteira possa ser feita. Mesmo não sendo perfeitas nas suas previsões, as redes superam todos os outros métodos (superando suas dificuldades) permitindo que se realizem previsões sobre as estimativas de retornos com certa confiabilidade nos resultados. Estes resultados são de extrema importância para o processo de seleção de carteiras, pois quanto melhores forem as previsões, mais acertadas serão as alocações das ações dentro de carteiras escolhidas e melhores serão os investimentos realizados.

Os resultados obtidos com a rede desenvolvida neste trabalho demonstram que ela consegue captar as tendências dos preços, conseguindo algumas vezes aproximar a estimativa de retorno do valor real. Isto confirma a idéia de que as redes neurais podem facilitar o trabalho dos investidores de ações, proporcionando grandes possibilidades de obtenção de ganhos, que é hoje um dos maiores objetivos não só do investidor, mas também de todo ser humano.

A grande contribuição deste trabalho foi mostrar que a união de duas áreas de conhecimento, Administração Financeira e Ciência da Computação, pode ser de grande benefício tanto para estudantes quanto para profissionais de ambas as áreas. O desenvolvimento do presente trabalho confirma o crescente destaque da administração financeira como uma área de aplicação para a informática e, ao mesmo tempo, demonstra a importância dos resultados atingidos com técnicas computacionais (em especial as redes neurais) no desenvolvimento de pesquisas em áreas financeiras,

alimentando até mesmo a esperança de que um dia será possível entender melhor os sistemas dinâmicos e caóticos tais como o mercado financeiro.

A utilização de redes neurais artificiais no mercado financeiro continuará sendo uma área de pesquisa de grande interesse enquanto pesquisadores e investidores esforçarem-se para realizar previsões, com o objetivo de melhorar seus retornos. Portanto, são inúmeras as possibilidades de trabalhos futuros dentre as quais destacam-se:

a) A implementação de outros tipos de redes neurais. A realização de testes com diferentes topologias de rede e/ou diferentes algoritmos de aprendizagem pode vir a melhorar o desempenho do modelo, aproximando ainda mais a saída obtida do resultado desejado.

b) A construção de um sistema híbrido formado por uma rede neural artificial e um algoritmo genético. A inclusão do algoritmo genético no modelo tem como objetivo definir os pesos que as ações devem ter na carteira visando otimizá-la. As redes neurais realizam a precificação de ações para servir de base para a escolha daquelas que farão parte da carteira. A função do algoritmo genético seria otimizar esta escolha, definindo quanto o investidor deve investir em cada ação dentro da carteira.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [ALV 01] ALVES, C. A. M. **Uma Ferramenta de Extração de Regras de Redes Neurais**. 2001. Dissertação (Mestrado em Ciências em Engenharia Civil) – COPPE, UFRJ, Rio de Janeiro.
- [BER 00] BERNSTEIN, P.; DAMODARAN, A. **Administração de Investimentos**. Traduzido por Cyro C. Patarra e José Carlos B. Santos. Porto Alegre: Bookman, 2000.
- [BOV 99] BOVESPA. **Mercado de Capitais – Introdução**. 1999. Disponível em: <<http://www.bovespa.com.br>>. Acesso em: mar. 2004.
- [BRU 98] BRUNI, A. L.; FAMÁ, R. Eficiência, Previsibilidade dos Preços e Anomalias em Mercados de Capitais: Teoria e Evidências. **Caderno de Pesquisas em Administração**, São Paulo, v.1, n.7, 1998.
- [CAB 02] CABRAL, **Mercados Financeiros: Uma Metodologia de Ensino de Estratégias de Investimento**. 2002. Tese (Doutorado em Engenharia de Produção) – UFSC, Florianópolis.
- [CAR 01] CARTACHO, M. S. **A utilização de um modelo híbrido Algoritmo Genético/Redes Neurais no processo de seleção de carteiras**. 2001. Dissertação (Mestrado em Administração) – Faculdade de Ciências Econômicas, UFMG, Belo Horizonte.
- [CAS 02] CASTRO Jr, F. H. F. C.; FAMÁ, R. As Novas Finanças e a Teoria Comportamental no Contexto da Tomada de Decisão sobre Investimentos. **Caderno de Pesquisas em Administração**, São Paulo, v.9, n. 2, p. 25-35, abr/jun. 2002.
- [DAM 97] DAMODARAN, A. **Avaliação de Investimentos**. Traduzido por Bazán Tecnologia e Linguística (Carlos Trieschmann e Ronaldo Rego). Rio de Janeiro: Qualitymark Ed., 1997.
- [DAN 97] DANDOLINI, G. A. **Um procedimento para avaliação da saúde financeira de pequenas empresas: estudo de um caso usando redes**

- neurais artificiais**. 1997. Dissertação (Mestrado em Engenharia de Produção) – UFSC, Florianópolis.
- [FAU 94] FAUSETT, L. **Fundamentals of Neural Networks – Architecture, Algorithms, and Applications**. [S.l.]: Prentice Hall International, Inc., 1994.
- [FRA 01] FRANCIS, L. The Basics of Neural Networks Demystified. **Contingencies**, p. 56-61, nov/dez. 2001.
- [FRE 99] FREITAS, A. A. C.; SILVA, I. N. Análise Técnica de Títulos Financeiros Através de Redes Neurais Artificiais. In: IV CONGRESSO BRASILEIRO DE REDES NEURAIAS, 1999, São José dos Campos. **Anais...** São José dos Campos, IV SBRN, 1999. p. 067-071.
- [FRE 01] FREITAS, S. O. **Utilização de um modelo baseado em Redes Neurais para a precificação de opções**. 2001. Dissertação (Mestrado em Administração) – Faculdade de Ciências Econômicas, UFMG, Belo Horizonte.
- [GON 02] GONÇALVES Jr, C.; PAMPLONA, E. O.; MONTEVECHI, J. A. B. Seleção de Carteiras Através do Modelo de Markowitz para Pequenos Investidores (Com o Uso de Planilhas Eletrônicas). In: IX SIMPEP, 2002, Bauru.
- [HAY 01] HAYKIN, S. **Redes Neurais, Princípios e Prática**. 2.ed. Porto Alegre: Bookman, 2001.
- [JUD 03] JÚDICE, J. J.; RIBEIRO, C. O.; SANTOS, J. P. J. **Análise Comparativa dos Modelos de Seleção de Carteiras de Acções de Markowitz e Konno**. 2003. Disponível em: <<http://dragao.co.it.pt/~judice/Articles/JudSantosRibeiro2003.pdf>>. Acesso em: set. 2004.
- [LAW 97] LAWRENCE, R. **Using Neural Networks to Forecast Stock Market Prices**. Dez. 1997. Disponível em: <<http://www.cs.uiowa.edu/~rlawrenc/research/Papers/nn.pdf>>. Acesso em: nov. 2003.
- [LAZ 00] LAZO, J. G.; VELLASCO, M. M.; PACHECO, M. A. A Hybrid Genetic-Neural System for Portfolio Selection and Management. In: SIXTH INTERNATIONAL CONFERENCE ON ENGINEERING APPLICATIONS OF NEURAL NETWORKS – EANN, 2000. **Proceedings...** Kingston Upon Thames, UK 17-19, jul. 2000.
- [MAR 02] MARTINS, R. X. Redes Neurais, O Aprendizado Reproduzido em Máquinas. **Tópicos Avançados em Ciência da Computação – Inteligência Artificial**, Varginha, v.5, n.5, mai. 2002. Disponível em: <<http://www.rxmartins.pro.br/interest.htm>>. Acesso em: nov. 2003.

- [MIL 01] MILANI, L. H. P. **Uma Contribuição para o Entendimento do Mercado de Ações, Através de um Teste Empírico do CAPM na BOVESPA no Período de 1996 a 2000**. 2001. Dissertação (Mestrado em Engenharia de Produção) – UFSC, Florianópolis.
- [NIL 98] NILSSON, N. J. **Artificial Intelligence: A New Synthesis**. San Francisco: Morgan Kaufmann Publishers, Inc. 1998.
- [OIC 02] ORGANIZAÇÃO INTERNACIONAL DO CAFÉ. **Diversificação nos Países Exportadores de Café**. Londres, 2002.
- [RIC 94] RICH, E.; KNIGHT, K. **Inteligência Artificial**. 2. ed. São Paulo: Makron Books do Brasil, 1994.
- [ROS 95] ROSS, S. A.; WESTERFIELD, R. W.; JAFFE, J. F. **Administração Financeira**. [S.l.]: Atlas, 1995.
- [SCH 03] SCHAEGLER, C. A. et al. **Redes Neurais**. São Leopoldo: UNISINOS, 2003.
- [TAF 95] TAFNER, M. A.; XEREZ, M.; FILHO, I. R. **Redes Neurais Artificiais: Introdução e Princípios de Neurocomputação**. Blumenau: EKO, 1995.
- [TAF 98] TAFNER, M. A. Redes Neurais Artificiais: Aprendizado e Plasticidade. **Revista Cérebro & Mente**, n.5, mai. 1998. Disponível em: <<http://www.cerebromente.org.br/n05/tecnologia/rna.htm>>. Acesso em: set. 2004.
- [UCP 03] UNIVERSIDADE CATÓLICA DE PELOTAS. Escola de Ciências Econômico-Empresariais. **Administração Financeira: O Ambiente Financeiro – Apostila 01**. Pelotas: UCPel, 2003.
- [VIE 00] VIEIRA, R. S.; THOMÉ, A. C. G. **Avaliação de Redes Neurais Aplicadas à Previsão de Índices de Mercados de Ações**. Rio de Janeiro: UFRJ, 2000.

## APÊNDICE A

### RESULTADOS DA REDE NEURAL PARA A AÇÃO ELETROBRÁS ON

Tabela A.1: Backpropagation aplicada à ação Eletrobrás ON

Neurônios	Épocas	Momentum	Taxa de Aprendizagem Inicial	$m$	Erro Treinamento	Erro Teste
20	100	0.1	0.8	6000	0.0947	0.0243
15	100	0.1	0.8	6000	0.0937	0.0200
15	100	0.3	0.5	5000	0.0727	0.0410
10	100	0.3	0.8	5000	0.0741	0.0396
10	150	0.3	0.6	5500	0.0918	0.0639
12	150	0.9	0.6	5000	0.1887	0.3252
12	125	0.9	0.6	10000	0.1345	0.2786
12	50	0.9	0.1	10000	0.0790	0.0273
17	50	0.9	0.1	4500	0.0729	0.0333
18	35	0.9	0.1	3000	0.0668	0.0265
18	35	0.5	0.1	3000	0.0350	0.0092
18	50	0.5	0.1	2000	0.0316	0.0202
18	100	0.5	0.5	2000	0.1616	0.1971
8	50	0.5	0.5	2000	0.1611	0.1987
8	30	0.5	0.5	2500	0.1234	0.1549
13	30	0.5	0.8	2500	0.1674	0.2853
13	50	0.5	0.8	6000	0.1283	0.1072
10	50	0.9	0.1	8000	0.0595	0.0368
10	50	0.9	0.7	8000	0.1693	0.2901
11	80	0.1	0.9	6000	0.1039	0.0353
9	60	0.3	0.8	5000	0.1159	0.0880
20	30	0.9	0.1	10000	0.1177	0.0222
17	30	0.3	0.8	6000	0.1134	0.0243
20	25	0.5	0.5	5500	0.1243	0.0329
20	35	0.3	0.7	6000	0.0978	0.0192
19	50	0.3	0.7	6000	0.0981	0.0202
19	75	0.6	0.01	5000	0.0368	0.0068
19	75	0.8	0.01	3000	0.0431	0.0063
15	80	0.9	0.1	10000	0.0520	0.0371
15	60	0.5	0.5	5000	0.1023	0.0658
15	50	0.7	0.01	4500	0.0359	0.0072
10	35	0.5	0.01	3000	0.0398	0.0074
10	100	0.6	0.01	1500	0.0426	0.0075
13	100	0.5	0.01	1000	0.0378	0.0069
14	60	0.5	0.01	800	0.0371	0.0118
14	40	0.3	0.8	1200	0.2781	0.3629
16	35	0.4	0.9	3500	0.1574	0.1956
16	75	0.6	0.1	6000	0.0424	0.0072

## APÊNDICE B

### RESULTADOS DA REDE NEURAL PARA A AÇÃO PETROBRÁS PN

Tabela A.2: Backpropagation aplicada à ação Petrobrás PN

Neurônios	Épocas	Momentum	Taxa de Aprendizagem Inicial	$m$	Erro Treinamento	Erro Teste
6	30	0.1	0.8	2000	0.2870	0.0128
6	30	0.8	0.8	6000	0.0895	0.0141
10	35	0.01	0.8	1500	0.3091	0.0103
8	30	0.8	0.8	6000	0.0915	0.0144
8	20	0.6	0.8	2500	0.2073	0.0143
8	40	0.01	0.8	2000	0.2732	0.0177
10	20	0.01	0.8	2000	0.2730	0.0227
10	50	0.01	0.8	2000	0.2999	0.0140
10	30	0.8	0.8	6000	0.0881	0.0171
5	40	0.01	0.8	2000	0.2825	0.0188
15	20	0.01	0.8	2000	0.2673	0.0249
12	30	0.1	0.8	2000	0.2598	0.0156
12	20	0.5	0.1	2500	0.0272	0.1043
12	30	0.9	0.1	1500	0.0255	0.0735
17	25	0.9	0.1	1500	0.0286	0.0731
17	20	0.5	0.5	1000	0.1953	0.0158
18	20	0.5	0.1	1000	0.0232	0.0866
19	15	0.5	0.01	1500	0.0194	0.1346
20	25	0.5	0.01	800	0.0197	0.0808
20	50	0.5	0.05	500	0.0236	0.0676
20	35	0.7	0.05	250	0.0285	0.0496
9	35	0.9	0.1	200	0.1055	0.0771
9	20	0.9	0.1	100	0.1019	0.0604
11	25	0.5	0.5	300	0.2818	0.0586
11	20	0.5	0.5	700	0.2645	0.0284
13	20	0.5	0.9	700	0.3223	0.0718
13	20	0.2	0.9	4000	0.1669	0.0653
15	25	0.7	0.3	4000	0.0505	0.0744
15	35	0.5	0.5	4000	0.0569	0.0574
15	40	0.9	0.1	3000	0.0396	0.1040
11	30	0.9	0.1	1000	0.0269	0.0827
10	15	0.9	0.1	1000	0.0627	0.0679
10	10	0.5	0.01	800	0.0191	0.1109
18	30	0.9	0.1	5000	0.0377	0.1049
18	25	0.3	0.8	1000	0.3595	0.0529
17	15	0.3	0.8	2000	0.2783	0.0127
17	15	0.1	0.8	2500	0.2034	0.0373

## APÊNDICE C

### RESULTADOS DA REDE NEURAL PARA A AÇÃO VALE RIO DOCE PNA

Tabela A.3: Backpropagation aplicada à ação Vale Rio Doce PNA

Neurônios	Épocas	Momentum	Taxa de Aprendizagem Inicial	$m$	Erro Treinamento	Erro Teste
10	20	0.3	0.7	2000	0.2447	0.0187
10	20	0.5	0.5	2000	0.1152	0.0227
10	20	0.9	0.5	2000	0.1076	0.0583
10	50	0.3	0.7	2000	0.2447	0.0187
10	100	0.9	0.1	5000	0.0289	0.1240
7	20	0.5	0.8	2500	0.2364	0.0140
8	30	0.9	0.1	5000	0.0286	0.1237
8	20	0.5	0.5	1500	0.2183	0.0334
9	15	0.1	0.8	1500	0.3385	0.0306
9	15	0.7	0.3	2000	0.0831	0.0761
12	10	0.7	0.3	1800	0.1111	0.0665
12	10	0.5	0.1	2000	0.0096	0.1456
12	30	0.5	0.1	450	0.0236	0.0589
13	25	0.3	0.6	200	0.3454	0.0756
14	20	0.6	0.2	1500	0.0462	0.1110
14	15	0.3	0.01	1000	0.0082	0.2184
14	25	0.3	0.01	180	0.0084	0.0880
15	20	0.9	0.1	5000	0.0247	0.0931
15	20	0.7	0.3	500	0.1277	0.0346
15	25	0.7	0.3	2500	0.0615	0.0816
16	15	0.5	0.01	400	0.0082	0.1048
16	30	0.9	0.01	6000	0.0379	0.2061
17	30	0.5	0.5	1000	0.2453	0.0263
17	20	0.5	0.3	1600	0.0836	0.0678
17	20	0.1	0.7	1200	0.3372	0.0262
18	20	0.5	0.5	1500	0.1948	0.0347
18	10	0.7	0.2	600	0.0775	0.0321
19	10	0.8	0.1	1000	0.0265	0.0561
19	20	0.3	0.01	300	0.0082	0.0989
19	20	0.2	0.8	1500	0.4170	0.0290
20	15	0.1	0.8	1500	0.3514	0.0340
20	10	0.7	0.5	1700	0.1466	0.0277
20	10	0.9	0.8	10000	0.0358	0.0371
20	20	0.6	0.9	8000	0.0861	0.0619
11	25	0.9	0.7	5000	0.0288	0.0472
11	20	0.5	0.8	5000	0.1424	0.0950
11	15	0.1	0.3	500	0.0771	0.0563



