

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
DEPARTAMENTO DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**UMA APLICAÇÃO DE REDES NEURAS ARTIFICIAIS
NA PREVISÃO DO MERCADO ACIONÁRIO**

ALESSANDRO MUELLER

*Dissertação submetida ao Programa de Pós-graduação em Engenharia de Produção
como requisito parcial para a obtenção do título de Mestre em Engenharia.*

FLORIANÓPOLIS, JULHO DE 1996.

UMA APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NA PREVISÃO DO MERCADO ACIONÁRIO

ALESSANDRO MUELLER

Esta dissertação foi julgada adequada como requisito parcial para a obtenção do título de Mestre em Engenharia e aprovada em sua forma final pelo programa de Pós-graduação em Engenharia de Produção.



Prof. RICARDO MIRANDA BARCIA, Ph.D.
Coordenador do Curso

Banca examinadora :



Prof. CARLOS ERNANI FRIES, M.Eng.

Orientador



Prof. EMÍLIO ARAÚJO MENEZES, Dr.



Prof. NORBERTO HOCHHEIM, Dr.



Prof. PLÍNIO STANGE, Dr.

À minha família, com muito amor.

AGRADECIMENTOS

Desenvolver uma dissertação não é uma tarefa fácil. Requer acima de tudo muita força de vontade e espírito de luta. Muitos obstáculos surgem ao longo do caminho e é preciso realmente acreditar no sonho inicial para seguir em frente. Mas seria pretensão afirmar que todo esse caminho é percorrido sozinho. Sem o esforço de muitos, talvez fosse impossível. Dessa forma, não poderia deixar de agradecer às pessoas que considero indispensáveis para que mais este projeto pudesse ser realizado.

À minha família, pelo carinho e confiança. A nossa união me deu forças para tornar realidade alguns sonhos e crescer. Sei que independentemente de qualquer situação sempre poderei contar com o apoio e amor de vocês.

Agradeço ao meu orientador, professor Carlos Ernani Fries, pela compreensão e credibilidade em mim depositada durante o desenvolvimento deste trabalho.

À professora Marley M. B. R. Vellasco, da PUC - RJ, pela dedicação. A sua contribuição foi fundamental para a concretização desta dissertação.

Quero agradecer também ao Altair, pela oportunidade e confiança no meu potencial. O seus ensinamentos e ideais foram essenciais na conquista do meu espaço e na maneira de observar o mundo.

À Fernanda, minha grande companheira, com quem compartilhei momentos inesquecíveis nestes últimos anos.

Ao Gláucio, que embora não tenha acompanhado a minha luta para concluir esta dissertação, sempre me estimulou na busca de conhecimento e, mesmo distante, continua sendo muito importante.

À Joyce, uma pessoa muito especial na minha vida, pelo apoio e amizade.

E finalmente, gostaria de agradecer à minha grande amiga Patrícia, pela inestimável ajuda e incentivo durante toda esta jornada, sempre presente nos momentos realmente significantes para mim. *I don't have words ...*

*If we could first know where we are and whither we are tending,
we could better judge what to do and how to do it.*

A. Lincoln

RESUMO

Métodos de previsão convencionais de séries temporais têm alcançado limitado sucesso na realização de prognósticos de séries econômicas. Este comportamento é devido à dificuldade desses modelos em manipular observações decorrentes de ambientes extremamente dinâmicos, como o mercado de ações. Redes neurais artificiais são, a princípio, capazes de tratar com o problema de instabilidade estrutural entre as observações de uma série temporal. Neste sentido, este trabalho procura investigar a habilidade dos modelos conexionistas em realizar previsões acuradas de séries de preços de ações. É proposta uma forma alternativa de antecipação do comportamento futuro dessas séries, através da identificação de regularidades no movimento da cotação das ações no mercado. Os resultados obtidos pela aplicação de técnicas de redes neurais artificiais são analisados empiricamente e confrontados com aqueles gerados pelos métodos previsão clássicos.

Palavras-chave: redes neurais artificiais, modelos conexionistas, séries temporais, previsão.

ABSTRACT

Classical forecasting methods of time series have achieved limited success in the realization of economic series forecasts. This behavior is caused by the difficulty of these models in manipulate observations from extremely dynamic environments, like the stock market. Artificial neural networks are, in principle, capable to dealing with the problem of structural instability among the time series observations. In this sense, this work investigates the habilities of the connexionist models in realize accurate forecasts of stock prices series. It is proposed an alternative form to predict the future behavior of these series, by identifying the regularities in the movement of the stock market. The results obtained from the application of artificial neural networks techniques are empirically analysed and confronted with the results reached by the classical forecasting methods.

Keywords : artificial neural networks, connecionist models, time series, forecasting.

LISTA DE FIGURAS

Figura 3.1.	Observações de uma série temporal com previsões de origem t e horizonte h	14
Figura 3.2.	Estágios da construção de um modelo ARIMA	26
Figura 4.1.	Neurônio biológico	33
Figura 4.2.	Unidade de processamento	39
Figura 4.3.	Funções de ativação	41
Figura 4.4.	Mínimo local	52
Figura 5.1.	<i>Janelamento</i> de uma série temporal	56
Figura 6.1.	Rede neural de múltiplas camadas <i>fully connected</i>	61
Figura 6.2.	Função tangente hiperbólica ($\tan H$)	62
Figura 6.3.	Gráfico do preço de fechamento diário de uma ação	67
Figura 6.4.	Gráfico dos preços previstos para a ação pelo modelo de redes neurais, considerando a propriedade de convergência	68
Figura 6.5.	Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência	69
Figura 6.6.	Gráfico de dispersão entre o preço real da ação e a previsão deste preço pelo modelo conexionista, considerando a propriedade de convergência	70
Figura 6.7.	Distribuição de frequência dos erros da previsão da rede neural, considerando a propriedade de convergência	70
Figura 6.8.	Gráfico comparativo entre os preços reais e os previstos pela Média Móvel, considerando a propriedade de convergência	71
Figura 6.9.	Gráfico comparativo entre os preços reais e os previstos pelo Alisamento Exponencial Simples, considerando a propriedade de convergência	71
Figura 6.10	Distribuição de frequência dos erros da previsão da Média Móvel, considerando a propriedade de	

	convergência	72
Figura 6.11.	Distribuição de frequência dos erros da previsão do Alisamento Exponencial Simples, considerando a propriedade de convergência	72
Figura 6.12.	Gráfico do preço de fechamento diário de uma ação, referente ao período de tempo imediatamente subsequente ao treinamento	73
Figura 6.13	Gráfico dos preços previstos para a ação pelo modelo de redes neurais, considerando a propriedade de generalização	74
Figura 6.14	Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização	74
Figura 6.15	Gráfico de dispersão entre o preço real da ação e a previsão deste preço pelo modelo conexionista, considerando a propriedade de generalização	75
Figura 6.16	Distribuição de frequência dos erros da previsão da rede neural, considerando a propriedade de generalização	76
Figura 6.17	Gráfico comparativo entre os preços reais e os previstos pela Média Móvel, considerando a propriedade de generalização	77
Figura 6.18	Gráfico comparativo entre os preços reais e os previstos pelo Alisamento Exponencial Simples, considerando a propriedade de generalização	77
Figura 6.19	Distribuição de frequência dos erros da previsão da Média Móvel, considerando a propriedade de generalização	78
Figura 6.20	Distribuição de frequência dos erros da previsão do Alisamento Exponencial Simples, considerando a propriedade de generalização	78
Figura 6.21	Gráfico do erro das previsões em relação ao número de iterações do treinamento	81
Figura 6.22	Gráfico do erro das previsões para diferentes conjuntos	

	de treinamento	82
Figura 6.23	Gráfico evolutivo de recursos aplicados no mercado de ações ou em Caderneta de Poupança	89

LISTA DE TABELAS

Tabela 3.1	Freqüência relativa de uso de técnicas de previsão em relação ao horizonte de previsão (em %)	28
Tabela 3.2	Freqüência relativa dos graus de familiaridade com os diferentes métodos de previsão (em %)	28
Tabela 3.3.	Freqüência relativa dos graus de satisfação com os diferentes métodos de previsão (em %)	29
Tabela 6.1	Erro percentual absoluto médio das previsões de padrões usados no treinamento da rede neural	79
Tabela 6.2	Erro percentual absoluto médio das previsões de padrões não usados no treinamento da rede neural	80
Tabela 6.3	Intervalo de confiança para o erro percentual absoluto médio das previsões de cinco conjuntos de padrões usados no treinamento da rede neural	82
Tabela 6.4	Intervalo de confiança para o erro percentual absoluto médio das previsões de cinco conjuntos de padrões não usados no treinamento da rede neural	83
Tabela 6.5	Erro percentual absoluto médio das previsões de padrões usados no treinamento da rede neural para diferentes valores do termo <i>momentum</i>	84
Tabela 6.6	Erro percentual absoluto médio das previsões de padrões não usados no treinamento da rede neural para diferentes valores do termo <i>momentum</i>	85
Tabela 6.7	Preço das ações para um período de tempo de vinte dias .	87
Tabela 6.8	Valores obtidos com a aplicação dos recursos no mercado acionário e em Caderneta de Poupança	88

SUMÁRIO

RESUMO	vi
ABSTRACT	vii
LISTA DE FIGURAS	viii
LISTA DE TABELAS	xi
1. INTRODUÇÃO	1
2. MERCADO DE AÇÕES	4
2.1. Considerações Iniciais	4
2.2. Definição	5
2.3. Espécies de Ações	6
2.4. Forma das Ações	6
2.5. Direitos conferidos pelas Ações	7
2.6. Análise do Preço das Ações	8
2.6.1. Análise Fundamentalista	9
2.6.2. Análise Técnica	10
3. SÉRIES TEMPORAIS	12
3.1. Definição	12
3.2. Objetivo	13
3.3. Previsão de Séries Temporais	13
3.4. Métodos de Previsão de Séries Temporais	15
3.4.1. Métodos de Decomposição de Séries Temporais	17
3.4.2. Métodos Simples de Previsão de Séries Temporais	18
3.4.2.1. Média Móvel	19
3.4.2.2. Alisamento Exponencial Simples	19
3.4.2.3. Alisamento Exponencial Linear	21
3.4.2.4. Alisamento Exponencial Sazonal e Linear de Winter ...	21
3.4.3. Métodos Avançados de Previsão de Séries Temporais	22

3.4.3.1. Modelos Autoregressivo e Médias Móveis	23
a) Modelo Autoregressivo	23
b) Modelo de Médias Móveis	23
c) Modelo Autoregressivo e de Médias Móveis	24
3.4.3.2. Método de Box e Jenkins	24
3.5. Considerações sobre os Métodos de Previsão de Séries Temporais	27
3.6. Medidas de Acuidade dos Métodos de Previsão de Séries Temporais	29
4. REDES NEURAIS	32
4.1. Redes Neurais Biológicas	32
4.2. Redes Neurais Artificiais	33
4.2.1. Definição	34
4.2.2. Histórico das Redes Neurais Artificiais	34
4.2.3. Topologia das Rede Neurais Artificiais	36
4.2.3.1. Camadas	37
4.2.3.2. Conexões e Pesos	38
4.2.3.3. Unidades Visíveis e Unidades Ocultas	38
4.2.4. Representação do Conhecimento em uma Rede Neural Artificial	39
4.2.5. Características das Redes Neurais Artificiais	42
4.2.5.1. Aprendizado	42
4.2.5.2. Generalização	43
4.2.5.3. Abstração	43
4.2.6. Aprendizado da Rede Neural Artificial	44
4.2.7. Estratégias de Aprendizado	45
4.2.8. Algoritmo de Aprendizado <i>Backpropagation</i>	46
4.2.8.1. Etapas do Algoritmo <i>Backpropagation</i>	47
4.2.9. Considerações sobre o Algoritmo <i>Backprppagation</i>	49
4.2.9.1. Dados para Treinamento	49
4.2.9.2. Tamanho da Rede Neural Artificial	49
4.2.9.3. Pesos e Parâmetros de Aprendizado	50

4.2.9.4. Mínimo Local	51
5. REDES NEURAIS ARTIFICIAIS E A PREVISÃO DO PREÇO DE AÇÕES	53
5.1. Análise de Viabilidade	53
5.2. Modelo de Rede Neural Proposto	55
5.3. Identificação de Regularidades em uma Série Temporal	56
5.4. Critérios de Avaliação das Previsões	57
6. APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NA PREVISÃO DO PREÇO DE AÇÕES	58
6.1. Definição do Problema	58
6.2. Modelagem do Problema	59
6.2.1. Arquitetura do Modelo	59
6.2.2. Parâmetros do Modelo	61
6.2.2.1. Função de Ativação	62
6.2.2.2. Taxa de Aprendizado	63
6.2.2.3. <i>Momentum</i>	63
6.2.2.4. Época	64
6.2.3. Aprendizagem Móvel	64
6.2.4. Dados para Aprendizagem	65
6.2.5. Tempo de Treinamento	66
6.2.6. Horizonte de Previsão	66
6.3. Análise dos Resultados	67
6.3.1. Convergência	67
6.3.2. Generalização	73
6.3.3. Estabilidade	78
6.3.3.1. Quanto à Arquitetura do Modelo Conexionista	79
6.3.3.2. Quanto ao Tempo de Treinamento	80
6.3.3.3. Quanto aos Diferentes Conjuntos de Treinamento	81
6.3.3.4. Quanto aos Termos do <i>Gradiente Decrescente</i>	83
6.3.3.5. Quanto às Condições Iniciais	85
6.4. Aplicação do Modelo de Previsão considerando uma Carteira	

Simplificada de Ações	85
6.5. Considerações Finais	89
7. CONCLUSÕES E RECOMENDAÇÕES	91
7.1. Conclusões	91
7.2. Recomendações	93
REFERÊNCIAS BIBLIOGRÁFICAS	95
ANEXO	100

CAPÍTULO I

1 - INTRODUÇÃO

A observação da realidade é consequência da manifestação de eventos complexos e incertos no decorrer do tempo. Embora estes eventos não sejam exatamente os mesmos, eles também não são completamente diferentes. Há uma linha de continuidade, similaridade e predictibilidade nestes eventos que permite generalizar eventos futuros, freqüentemente de modo correto, a partir de experiências passadas.

A antecipação do comportamento futuro dos acontecimentos sempre despertou interesse nas mais diversas áreas do conhecimento humano. Em finanças, especificamente, a possibilidade de se realizar algum prognóstico permite a tomada de decisão antecipada de um investidor em relação às expectativas do mercado.

Quando o conhecimento exato das leis que regem um determinado fenômeno é expresso através de equações precisas, as quais podem a princípio ser resolvidas, é possível prever o comportamento futuro dos eventos pela aplicação de um modelo matemático. Neste sentido, vários procedimentos de previsão de séries temporais foram desenvolvidos com o objetivo de solucionar o problema da previsão. Entretanto, apesar da diversidade de métodos de previsão de séries temporais, nem sempre a solução encontrada pelo emprego destes métodos é satisfatória. Um ponto possível desta deficiência talvez seja a instabilidade dos parâmetros de tais métodos, mas é mais provável que a causa seja a ausência de não-linearidade nos modelos de previsão [REFENES1993c].

Outra alternativa para a realização de previsões consiste na investigação empírica de uma série temporal em busca da identificação de regularidades presentes nas observações da série de interesse. O desafio está no reconhecimento destas regularidades, que não são sempre evidentes, sendo freqüentemente mascaradas por ruídos [REFENES1993c]. De acordo com [REFENES1993a], há fenômenos que

parecem aleatórios, sem aparente periodicidade, embora recorrentes em um sentido geral.

Devido a sua natureza indutiva, redes neurais artificiais podem sobrepor a formulação teórica e inferir relações não-lineares complexas entre as observações de uma série temporal [REFENES1993b]. Desta forma, arquiteturas de redes neurais têm recebido considerável atenção nos últimos anos em função de suas habilidades em realizar a aprendizagem, sendo aplicadas em um grande número de situações e demonstrando desempenho superior às abordagens convencionais [REFENES1993a].

Este trabalho investiga o potencial preditivo das redes neurais artificiais na manipulação das observações de séries de preços de ações, notadamente, uma situação atípica em relação às séries temporais normalmente utilizadas para avaliar a preditibilidade destes modelos.

No capítulo 2 são apresentados alguns conceitos básicos necessários para a compreensão do mercado acionário. A definição de ação é abordada, assim como a classificação e formas de circulação das ações. Também são colocados os métodos usuais de análise de ações.

O capítulo 3 é dedicado ao estudo de séries temporais e a sua aplicabilidade na previsão de valores futuros destas séries. São introduzidos os conceitos de série temporal e previsão, além da classificação e descrição de alguns métodos de previsão. Ainda, são feitas determinadas considerações a respeito destes métodos e sobre as medidas de acuidade usualmente adotadas na avaliação de tais procedimentos.

No capítulo 4, a estrutura e o funcionamento das redes neurais biológicas são descritos, de modo a estabelecer uma analogia com os modelos de redes neurais artificiais. Na seqüência, uma possível definição de rede neural, sua organização, características e histórico da abordagem baseada nestes modelos são apresentados. Também é descrita a forma de representação do conhecimento em uma rede neural e como o conhecimento contido em sua estrutura é utilizado no processo de aprendizado, e conseqüente realização de previsões.

No capítulo 5 é avaliada a possibilidade de previsão dos preços das ações através de técnicas de redes neurais artificiais. É proposto um modelo de rede neural e critérios para avaliação desse modelo. Ainda, é abordado um método de identificação de regularidades presentes nas séries temporais modeladas.

No capítulo 6, a abordagem de modelos conexionistas é aplicada na manipulação de séries de cotações de ações e na investigação da habilidade desses modelos em prever valores futuros de ações. O modelo de redes neurais artificiais proposto é especificado e é feita a análise dos resultados obtidos com o aprendizado. A estabilidade dos valores previstos também é criticada e uma aplicação dos modelos conexionistas considerando uma carteira simplificada de ações é analisada.

O capítulo 7 apresenta as conclusões inferidas pela aplicação do método de redes neurais artificiais na previsão da cotação de ações e algumas recomendações para trabalhos futuros.

No Anexo são apresentados alguns gráficos comparativos entre os preços reais de determinadas ações e os valores previstos pela aplicação dos modelos conexionistas na manipulação dessas séries temporais.

CAPÍTULO II

2 - MERCADO DE AÇÕES

Este capítulo introduz alguns conceitos básicos relativos ao mercado de ações. É apresentado o conceito de ação, assim como uma classificação e as formas de circulação das ações. Também são apresentados os métodos usuais de análise de ações.

2.1. Considerações Iniciais

Uma empresa em fase de crescimento necessita de recursos para financiar seus projetos de expansão. O ideal seria que a empresa tivesse acumulado superávits no passado para poder financiar suas expansões com recursos próprios. Entretanto, muitas vezes esta condição não pode ser satisfeita.

O financiamento desses projetos pode, então, ser obtido através de empréstimos. Contrair empréstimos junto a uma instituição financeira envolve o pagamento de juros e prazo determinado para saldar o montante financiado. Os pagamentos são compulsórios, independente da empresa operar com lucro ou prejuízo.

Outra possibilidade para a empresa manter seu programa de investimento sem se submeter ao mercado de crédito é a abertura de seu capital pela emissão de ações. Os recursos oriundos de emissão de ações são devolvidos ao investidor através da distribuição de dividendos e da retenção de lucros que, se forem adequadamente reaplicados, valorizarão o preço das ações no mercado. O mercado de ações, segmento do mercado de capitais que opera com títulos de renda variável, objetiva canalizar recursos para as empresas através do capital de risco [OLIVEIRA1986].

Segundo a Revista Investimentos, o mercado acionário encerrou o ano de 1995 apresentando uma rentabilidade real acumulada baixa, mesmo após ter registrado índices

de valorização bastante elevados em determinados meses. O desempenho geral observado no ano anterior foi fundamentalmente um reflexo do período de adaptação pelo qual está passando a economia brasileira sob os efeitos das medidas introduzidas pelo programa de estabilização econômica.

Esse comportamento pode ser interpretado como um momento de ajustamento do mercado às novas diretrizes do governo. Em decorrência da quebra das receitas inflacionárias, a estrutura especulativa baseada nas altas taxas de inflação revelou-se insustentável. A queda anunciada dos juros também é relevante na perspectiva de um investidor de longo prazo, que observa hoje juros muito menores que aqueles praticados no passado. Assim, para tornar-se mais competitivo, técnico e compatível com o novo cenário de estabilidade, o mercado de ações teve que buscar novas estratégias e produtos para se adaptar à nova realidade.

Apesar dos contratempos enfrentados no ano passado, onde a desaceleração do programa de privatização causou forte impacto, o mercado brasileiro se coloca entre os quatro mercados mais rentáveis do mundo neste primeiro trimestre. Desta forma, a retomada do crescimento econômico aliada ao processo de privatização das estatais, tendem a impulsionar o mercado com uma maior dispersão acionária e a consolidação do hábito de poupar através do mercado de ações.

Diante de uma situação onde se verifica a demanda por modelos de previsão que possam auxiliar no gerenciamento do risco envolvido na aplicação de recursos em ações, surgem instituições detentoras de conhecimento específico para avaliar o comportamento do mercado acionário.

2.2. Definição

A ação é um título representativo da fração unitária do capital social de uma sociedade anônima ou companhia, denominação deste tipo de empresa na legislação brasileira, conferindo a seu proprietário o direito de participação nessa sociedade [CASTRO1979].

2.3. Espécies de Ações

Quanto à espécie, as ações normalmente são classificadas em [SANVICENTE1988]:

- **Ações Ordinárias** : as ações ordinárias conferem direito de voto a seu titular nas deliberações das assembléias de acionistas - autoridade máxima da companhia -, além da participação nos lucros da sociedade [OLIVEIRA1986]. Normalmente, a cada ação corresponde um voto [CASTRO1979]. Conseqüentemente, o proprietário de ações ordinárias tem responsabilidades e obrigações correspondentes ao montante de ações possuídas [OLIVEIRA1986].
- **Ações Preferenciais** : as ações preferenciais normalmente não dão direito de voto, característica das ações ordinárias reconhecida pelos diversos sistemas jurídicos [CASTRO1979]. Entretanto, possuem vantagens e preferências em relação às ações ordinárias, protegendo seus titulares contra decisões tomadas nas assembléias de acionistas [SÁ1987]. Essas preferências consistem em prioridade na distribuição de dividendos, no reembolso de capital no caso de dissolução da empresa e na acumulação destas vantagens [CASTRO1979] [OLIVEIRA1986] [SÁ1987] [SANVICENTE1988] [RUDGE1993].

Segundo [RUDGE1993], a legislação brasileira ainda prevê a existência de ações de fruição, ou de gozo, ou ainda partes beneficiárias, apesar de essas formas acionárias não serem objeto de negociação em mercados organizados.

2.4. Forma das Ações

[RUDGE1993] classifica as ações ordinárias e preferenciais quanto à forma de circulação em nominativas e escriturais.

- **Ações Nominativas** : são aquelas cujos certificados identificam seu titular, também inscrito no Livro de Registro de Ações Nominativas [CASTRO1979]. A transferência entre titulares exige o expresse consentimento do acionista vendedor e a inscrição do acionista comprador no Livro de Transferência de Ações Nominativas [OLIVEIRA1986] [SÁ1987].
- **Ações Escriturais** : não são representadas por certificados, sendo títulos mantidos em conta de depósito, em nome de seus titulares, na instituição financeira que o estatuto da empresa designar [CASTRO1979]. Segundo [SÁ1987], a propriedade da ação escritural é presumida pelo registro na conta de depósito de ações. Esta modalidade de ação objetiva difundir a propriedade de ações nominativas, facilitar a circulação - proporcionada pela transferência mediante ordem à instituição financeira e registro dessa transação -, e eliminar os custos com a emissão de certificados [CASTRO1979].

2.5. Direitos conferidos pelas Ações

A Lei N° 6.404/76, que regulamenta o funcionamento das sociedades anônimas, confere aos acionistas o recebimento de dividendos e bonificações, e direitos de subscrição no caso de a empresa pretender aumentar seu capital pela captação de novos recursos [CASTRO1979].

- **Dividendo** : de acordo com [CASTRO1979], os dividendos são decorrentes da distribuição de uma parcela do lucro da empresa aos acionistas, sendo pagos em dinheiro. O dividendo se constitui na renda do investimento em ações, contribuindo decisivamente para o valor da ação [SÁ1987].
- **Bonificação** : em decorrência do aumento do capital da sociedade anônima, realizado com a incorporação de reservas ao capital da empresa, é feita uma distribuição de ações aos acionistas na forma de bonificação [CASTRO1979]. A bonificação distribuída aos acionistas é efetuada na proporção de sua

participação no capital social da companhia e normalmente na mesma espécie das ações possuídas [OLIVEIRA1986] [SÁ1987], representando a atualização da cota de participação do acionista no capital da empresa. A decisão da empresa em efetuar uma bonificação não implica em qualquer variação patrimonial, apesar de que o anúncio de bonificação pode, eventualmente, provocar a valorização das ações da companhia no mercado [SÁ1987].

- **Subscrição** : a subscrição corresponde ao direito garantido aos acionistas ordinários e preferenciais na prioridade de aquisição de ações da empresa, quando esta aumentar o seu capital social pela emissão de novos títulos [RUDGE1993]. Através da subscrição de ações é facultado aos acionistas, adquirirem, pelo preço de emissão, uma parcela das novas ações proporcional àquelas possuídas [OLIVEIRA1986]. Normalmente o aumento de capital pela subscrição de novas ações é realizado quando as condições de mercado se apresentarem favoráveis, permitindo, então, que os acionistas ao subscreverem ações, ganhem a diferença entre o preço de mercado e o valor da emissão do título [OLIVEIRA1986].

2.6. Análise do Preço das Ações

O investimento no mercado acionário advém da expectativa de ganhos decorrentes do rendimento das ações adquiridas e de sua valorização no mercado. *“Quem compra ações, mesmo inconscientemente, especula sobre a possibilidade de rendimentos futuros, o comportamento do mercado e a possibilidade de valorização do papel”* [OLIVEIRA1986].

A avaliação do investimento em ações é explicada pelas correntes fundamentalista e técnica [CASTRO1979] [OLIVEIRA1986] [RUDGE1993] [SÁ1987]. Na realidade, os métodos defendidos por estas correntes de pensamento são complementares e dificilmente utiliza-se exclusivamente um dos dois métodos [SÁ1987]. É usual o emprego simultâneo destas técnicas, onde *“a análise fundamentalista é adequada para a seleção daquelas ações que devem ser compradas, enquanto a análise técnica serve*

para determinar a ocasião em que devem ser compradas” [SÁ1987]. Pela combinação destes métodos, mais precisas serão as possibilidades de maximizar ganhos e minimizar perdas em uma carteira de ações.

2.6.1. Análise Fundamentalista

A hipótese admitida pela análise fundamentalista é a existência de um valor natural ou intrínseco para cada ação [CASTRO1979]. [OLIVEIRA1986] define o valor intrínseco de uma ação como sendo aquele dado pela remuneração efetiva¹ mais a remuneração potencial² do investimento; desta forma, o valor intrínseco de uma ação está correlacionado com o desempenho da companhia emissora.

Em muitas situações, o preço de mercado para a ação não coincide com seu valor intrínseco devido às imperfeições e ineficiência do mercado na avaliação desse título. Segundo [VANHORNE1972], o problema consiste, então, em determinar se o mercado está avaliando corretamente o preço da ação; ou seja, verificar se os preços desses títulos refletem adequadamente o verdadeiro valor dos direitos de propriedade de uma sociedade anônima. Neste sentido, a análise fundamentalista, através da identificação e estudo de fatores que influenciam no desempenho das empresas, procura determinar o valor intrínseco de uma ação, favorecendo a tomada de decisão dos investidores do mercado acionário. De acordo com [CASTRO1979] e [RUDGE1993], normalmente a estimação desse valor é encontrada pelo cálculo do valor atual das ações em função do fluxo dos lucros por ação efetivamente distribuídos em um determinado intervalo de tempo, refletindo a política de distribuição da empresa e sua capacidade de gerar lucro. As demonstrações financeiras e outras informações relevantes que ilustrem o comportamento passado e presente da sociedade, são instrumentos importantes para se realizar a projeção da evolução futura da empresa. A conjuntura econômica e considerações políticas adotadas também devem ser analisadas, pois exercem influência no preço e valor das ações negociadas [SÁ1987]. Assim, para os fundamentalistas, o

¹ A remuneração efetiva é aquela que se pode considerar praticamente certa, avaliando o comportamento histórico do desempenho da empresa [OLIVEIRA1986].

² A remuneração potencial depende do comportamento futuro da sociedade anônima [OLIVEIRA1986].

preço de uma ação depende do desempenho da empresa no seu contexto econômico e que somente pela avaliação deste desempenho é possível a estimação do preço da ação [OLIVEIRA1986].

Como o preço justo - valor intrínseco - de uma ação se fundamenta na expectativa de lucros futuros, [RUDGE1993] coloca que o momento adequado para a compra de uma ação é quando o preço deste título estiver baixo e para a venda, quando estiver alto. Assim, se o investidor achar que a cotação da ação está acima do preço considerado justo, não é opção de investimento; caso contrário, avaliado o risco envolvido na aquisição e na manutenção dessa ação, poderá vir a sê-la.

Os defensores da escola fundamentalista colocam que teoricamente os valores e os preços de mercado deveriam ser iguais [OLIVEIRA1986]. A não observância de tal fato se deve a subestimativas ou superestimativas do preço da ação, que sofre influências de toda ordem, inclusive de caráter psicológico, provocando oscilações rápidas de maior ou menor amplitude, quanto mais ineficiente e desinformado for o mercado [SÁ1987] [VANHORNE1972].

Entretanto, os preços das ações tendem a se movimentar em torno do seu valor intrínseco. Assim, a longo prazo, o mercado, apesar de suas imperfeições, avalia corretamente o valor de uma ação [SÁ1987].

2.6.2. Análise Técnica

A análise técnica assume que os preços das ações apresentam tendência de movimento condicionada a uma dependência significativa dos preços observados no passado [CASTRO1979]. Baseada na hipótese de que "*a história se repete*" [OLIVEIRA1986], essa abordagem procura explicar as oscilações dos preços dos títulos pela própria evolução desses preços no mercado, realizando previsões sobre o comportamento futuro da precificação das ações [LEVY1969] [SÁ1987].

Segundo os defensores da escola técnica, a oscilação nos preços é determinada pela interpretação dada pela massa de investidores a fatores fundamentais que, a princípio, influenciam a variação dos preços das ações e que se modificam constantemente [SÁ1987]. Para interpretar o desempenho das ações e do mercado, a análise técnica utiliza artifícios gráficos de forma a ilustrar a série histórica de preços, permitindo a identificação de tendências e formações na cotação das ações, traduzindo o comportamento do mercado e avaliando a participação de massas de investidores na indução dessas oscilações [RUDGE1993].

Para [CASTRO1979], através da análise das tendências observadas nos gráficos é possível estabelecer oportunidades mais precisas de negociação. As formações, sinais gráficos decorrentes de oscilações nos preços das ações, são utilizadas como instrumentos de previsão das cotações pela constatação empírica de que após o aparecimento dessas, normalmente, os preços apresentam comportamento evolutivo idêntico ao observado.

A análise técnica assume que *“as ações são avaliadas de acordo com os caprichos do mercado e não de acordo com qualquer valor resultante da rentabilidade potencial e de risco da empresa”* [VANHORNE1972]. A sua preocupação está centrada na busca de características consideradas recorrentes no tempo, sobre as quais define probabilidades de ocorrência no futuro, associando níveis de incerteza à realização da previsão [CASTRO1979].

CAPÍTULO III

3 - SÉRIES TEMPORAIS

Este capítulo é dedicado ao estudo de séries temporais e sua aplicabilidade na previsão de valores futuros dessas séries. São introduzidos os conceitos de série temporal e previsão, além da classificação e descrição de alguns métodos de previsão. Também são feitas determinadas considerações a respeito desses métodos e sobre as medidas de acuidade usualmente adotadas na avaliação de tais procedimentos.

3.1. Definição

A classe de fenômenos cujo processo observacional e conseqüente quantificação numérica gera uma seqüência de dados distribuídos no tempo é denominada *série temporal* [SOUZA1989].

A natureza de uma série temporal e a estrutura de seu mecanismo gerador estão relacionadas com o intervalo de ocorrência das observações no tempo [ANDERSON1971]. Caso o levantamento das observações da série possa ser feito a qualquer momento do tempo, a série temporal é dita *contínua*, sendo denotada por $x(t)$ [GRANGER1977]. Entretanto, de acordo com [GRANGER1977] e [NELSON1973], na maioria das séries, as observações são tomadas em intervalos de tempo discretos e equidistantes.

Uma série temporal *discreta* pode ser representada por $X^T = \{x_1, x_2, \dots, x_T\}$, sendo que cada observação discreta x_t está associada a um instante de tempo distinto, existindo uma relação de dependência serial entre essas observações [SOUZA1989].

3.2. Objetivo

[GRANGER1977] coloca como objetivo inicial da análise de séries temporais a realização de inferências sobre as propriedades ou características básicas do mecanismo gerador do processo estocástico¹ das observações da série. Assim, através da abstração de regularidades contidas nos fenômenos observáveis de uma série temporal existe a possibilidade de se construir um modelo matemático como uma representação simplificada da realidade [BARBANCHO1970].

Após a formulação do modelo matemático, obtido pela seleção entre as alternativas de classes de modelos identificadas como apropriadas para essa representação e subsequente estimação de seus parâmetros, é possível utilizá-lo para testar alguma hipótese ou teoria a respeito do mecanismo gerador do processo estocástico e realizar a previsão de valores futuros da série temporal [GRANGER1977] [NELSON1973].

3.3. Previsão de Séries Temporais

Segundo [BARBANCHO1970], uma previsão é uma manifestação relativa a sucessos desconhecidos em um futuro determinado. A previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisão, visando atingir determinados objetivos [MORETTIN1981].

Considerando um conjunto de observações de uma série temporal coletadas até o instante t e de um modelo que represente esses fenômenos, a previsão do valor da série no tempo $t+h$ (figura 3.1) pode ser obtida.

¹ Modelo que descreve a estrutura de probabilidade de uma seqüência de observações [BOX1976].

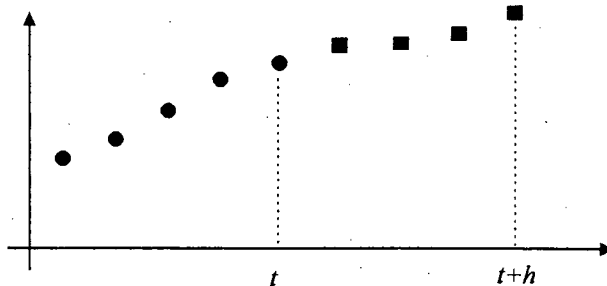


Figura 3.1. Observações de uma série temporal com previsões de origem t e horizonte h .

Dependendo do valor assumido pelo horizonte de previsão, [SOUZA1989] e [WHEELWRIGTH1985] classificam as previsões dos valores futuros de uma série temporal como de curto, médio ou longo prazo. Desse modo, diante da possibilidade de existência de diferentes horizontes de previsão, [REFENES1993a] especifica técnicas distintas para prognosticar os valores futuros de uma série temporal :

- *previsão múltiplos passos* : esta abordagem, adotada para longos horizontes de previsão, procura identificar as tendências gerais e os pontos de inflexão² mais relevantes na série temporal. Na *previsão múltiplos passos*, o conjunto de valores correntes é empregado na realização da previsão para determinado instante; esta previsão é, então, introduzida entre as observações passadas, compondo, desta forma, um novo conjunto de dados, sobre o qual será obtida a previsão do tempo subsequente.
- *previsão simples passo* : nesta técnica não há incorporação de previsões aos dados utilizados para encontrar a próxima previsão, sendo esta, independente dos valores anteriormente previstos. A previsão é feita apenas para o período de tempo imediatamente posterior ao atual, a partir das observações da série temporal.

² Pontos onde se verifica inversão na tendência de movimento da série temporal [REFENES1993a].

Para [SOUZA1989], a garantia da otimalidade das previsões de uma série temporal somente é alcançada adotando como horizonte de previsão o instante de tempo imediatamente subsequente à origem t .

Naturalmente, a investigação do poder preditivo do modelo especificado como o mais adequado para explicar o mecanismo gerador das observações de uma série temporal é um processo empírico de verificação, visto que são feitas comparações entre as previsões e as observações, a fim de confirmar a habilidade do modelo matemático em descrever a estrutura definida pelos dados da série temporal analisada [BARBANCHO1970]. É importante salientar que o caráter previsional de um determinado modelo, segundo [BARBANCHO1970], não pode ser sustentado quando forem identificadas mudanças estruturais nas observações obtidas a partir da origem, em relação àquelas utilizadas para a elaboração do modelo.

3.4. Métodos de Previsão de Séries Temporais

[WHEELWRIGTH1985] define um método de previsão como sendo o conjunto de procedimentos usados no desenvolvimento de uma determinada previsão.

Os métodos de previsão de séries temporais, classificados como métodos quantitativos³, baseiam suas previsões na extrapolação de características de observações passadas e no inter-relacionamento entre essas observações, fornecendo previsões acuradas se o futuro apresentar comportamento similar ao passado [WHEELWRIGTH1985].

De acordo com [WHEELWRIGTH1985], a maioria dos métodos de previsão de séries temporais se baseia na suposição de que observações passadas contém todas as informações sobre o padrão de comportamento da série temporal e esse padrão é recorrente no tempo. O propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse

³ Métodos de previsão que utilizam dados quantitativos históricos para a obtenção de prognósticos futuros [WHEELWRIGTH1985].

padrão para prever os valores futuros da série temporal. Assim, pela identificação desse componente, a previsão para períodos de tempo subsequentes ao observado pode ser desenvolvida.

Apesar de quase a totalidade dos métodos de previsão de séries temporais estar fundamentada apenas na análise das observações da série de interesse para a especificação de algum modelo que descreva essas observações, alguns procedimentos de previsão tentam explicar o comportamento de uma série temporal pela evolução dos fenômenos observacionais de outras séries. Desta forma, dependendo do número de séries temporais envolvidas na modelagem, [SOUZA1989] classifica os métodos de previsão em *univariados*, *funções de transferência* e *multivariados*.

Os *métodos univariados*, que compreendem a maior parte dos métodos de previsão de séries temporais, consideram somente uma única série para a realização dos prognósticos. As previsões decorrentes da aplicação de métodos univariados podem estar relacionadas apenas com as informações contidas na série histórica de interesse (métodos baseados na estatística clássica) ou também, além de incorporarem essas informações, consideram outras supostamente relevantes e que não estão contidas na série analisada (métodos baseados na estatística bayesiana).

Aquelas metodologias nas quais a série de interesse é explicada não só pelo seu passado histórico, como também por outras séries temporais não-correlatas entre si, são conhecidas como *funções de transferência*. Esta classe de métodos de previsão envolve, portanto, mais de uma série temporal, com a ressalva de que a relação de causalidade entre estas séries é perfeitamente conhecida.

Os *métodos multivariados* abrangem os procedimentos de previsão que associam mais de uma série temporal na efetivação de prognósticos sem, no entanto, qualquer imposição com relação à causalidade entre essas séries.

Diante disso, há uma variedade enorme de métodos de previsão de séries temporais, cada qual com suas capacidades e limitações [WHEELWRIGTH1985]. Qualquer que seja a classificação desses métodos, é possível utilizar um número muito

grande de métodos diferentes para descrever o comportamento de uma série particular. A seleção do método de previsão adequado depende de vários fatores, tais como o comportamento do fenômeno observável ou o conhecimento a priori que se tenha sobre a sua natureza e do objetivo da análise.

[SOUZA1989] e [WHEELWRIGTH1985] relacionam os métodos de previsão baseados exclusivamente em uma única série histórica de dados em métodos de *decomposição*, métodos *simples* de previsão e métodos *avançados* de previsão de séries temporais.

3.4.1. Métodos de Decomposição de Séries Temporais

Os métodos de decomposição assumem que uma série temporal é constituída por um conjunto de componentes não-observáveis [SOUZA1989]. Dessa forma, pela identificação das componentes individuais presentes no padrão básico da série histórica de dados (tendência, ciclo, sazonalidade e aleatoriedade), a extrapolação para o futuro pode ser realizada [WHEELWRIGTH1985].

[WHEELWRIGTH1985] expressa o relacionamento entre as componentes não-observáveis da série temporal através da equação (3.1).

$$x_t = f(S_t, T_t, C_t, E_t) \quad (3.1)$$

onde S_t corresponde à componente sazonal para o período t ;

T_t é a componente de tendência no período t ;

C_t é a componente de ciclo no período t e

E_t é a componente aleatória no período t .

Segundo [MORETTIN1981], a componente sazonal representa as flutuações da série de acordo com algum fator de sazonalidade. O ciclo apresenta um comportamento similar à componente sazonal, embora tenha normalmente comprimento maior que

aquela. Justamente pelo fato de não apresentar duração uniforme, a identificação da componente ciclo é mais problemática. A tendência representa o aumento ou declínio gradual nos valores das observações de uma série temporal. Com a remoção das componentes de sazonalidade, ciclo e tendência, a componente aleatória fica determinada [WHEELWRIGTH1985].

[WHEELWRIGTH1985] coloca que vários procedimentos para a decomposição de séries temporais foram desenvolvidos, cada qual tentando isolar as componentes não-observáveis da série o mais acuradamente possível. O objetivo desses procedimentos consiste em remover cada uma das componentes, permitindo que o comportamento da série temporal seja melhor compreendido e, conseqüentemente, prognosticar valores futuros mais apropriados.

3.4.2. Métodos Simples de Previsão de Séries Temporais

Os métodos assim classificados efetuam a previsão do valor futuro da série temporal pelo alisamento das observações passadas da série de interesse [WHEELWRIGTH1985]. Assumindo que os valores extremos da série representam flutuações aleatórias [MORETTIN1981], o propósito desses métodos consiste em identificar o padrão básico presente nos dados históricos e, então, usar esse padrão para prever valores futuros.

[MORETTIN1981] associa a grande popularidade desses métodos à simplicidade, à eficiência computacional e à razoável previsão obtida. Entre os métodos simples de previsão destacam-se o da *Média Móvel*, o *Alisamento Exponencial Simples*, o *Alisamento Exponencial Linear* e o *Alisamento Exponencial Sazonal e Linear de Winter* [WHEELWRIGTH1985], os quais são apresentados sucintamente na seqüência.

3.4.2.1. Média Móvel

Esse método considera como previsão para o período futuro a média das observações passadas recentes [WHEELWRIGHT1985].

Em [MORETTIN1981] a média móvel para o período de tempo t é definida por

$$x_t = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n} \quad (3.2)$$

onde n representa o número de observações incluídas na média x_t .

O termo *média móvel* é utilizado porque à medida que a próxima observação se torna disponível, a média das observações é recalculada, incluindo essa observação no conjunto de observações e desprezando a observação mais antiga [MORETTIN1981].

[WHEELWRIGHT1985] coloca que quanto maior o número de observações incluídas na média móvel, maior o efeito de alisamento na previsão. Assim, caso a série temporal apresente muita aleatoriedade ou pequenas mudanças nos padrões dessa série, um número maior de valores podem ser utilizados no cálculo da média móvel, obtendo-se uma previsão mais alisada. Entretanto, se houver pouca flutuação aleatória nos dados ou mudança significativa no padrão da série, um número menor de observações deve ser incluído no conjunto de valores empregado na determinação da média para que se possa reagir a essas alterações mais rapidamente.

3.4.2.2. Alisamento Exponencial Simples

A princípio, o método conhecido como Alisamento Exponencial Simples se assemelha ao da Média Móvel por extrair das observações da série temporal o comportamento aleatório pelo alisamento dos dados históricos. Entretanto, a inovação introduzida pelo Alisamento Exponencial Simples advém do fato de este método atribuir

pesos diferentes a cada observação da série. Enquanto que na Média Móvel as observações usadas para encontrar a previsão do valor futuro contribuem em igual proporção para o cálculo dessa previsão, no Alisamento Exponencial Simples as informações mais recentes são evidenciadas pela aplicação de um fator que determina essa importância [WHEELWRIGTH1985].

Segundo Wheelwright [WHEELWRIGTH1985], o argumento para o tratamento diferenciado das observações da série temporal é fundamentado na suposição de que as últimas observações contêm mais informações sobre o futuro e, portanto, são mais relevantes para a previsão.

[WHEELWRIGTH1985] especifica o método Alisamento Exponencial Simples através da equação (3.3).

$$F_{t+1} = \alpha x_t + (1 - \alpha) F_t \quad (3.3)$$

onde F_{t+1} representa a previsão no tempo $t + 1$ e

α é o peso atribuído à observação x_t , $0 < \alpha < 1$.

De acordo com [MORETTIN1981] e [GRANGER1977], o valor assumido por α determina o ajuste aplicado aos dados. Quanto menor o valor da constante α , mais estáveis serão as previsões, visto que a utilização de baixo valor de α implica na atribuição de peso maior às observações passadas e, conseqüentemente, qualquer flutuação aleatória no presente contribui com menor importância para a obtenção da previsão. Contudo, não há metodologia que oriente quanto à seleção de um valor apropriado para α , sendo normalmente encontrado por tentativa e erro [WHEELWRIGTH1985]. Um procedimento mais objetivo seria a seleção do valor de α que forneça a *“melhor previsão das observações contidas na série temporal”* [MORETTIN1981].

3.4.2.3. Alisamento Exponencial Linear

Quando o método Alisamento Exponencial Simples é aplicado na previsão de séries temporais que apresentam tendência entre as observações passadas, os valores prognosticados superestimam (ou subestimam) os valores reais [MORETTIN1981]. Desta forma, a acuidade das previsões fica prejudicada.

Para evitar esse erro sistemático, o método Alisamento Exponencial Linear foi desenvolvido procurando reconhecer a presença de tendência na série de dados [WHEELWRIGTH1985]. O valor da previsão obtido através deste método é alcançado pela aplicação da equação (3.4) [WHEELWRIGTH1985].

$$F_{t+m} = S_t + mT_t \quad (3.4)$$

onde S_t corresponde à previsão no tempo t , conforme equação (3.5);

T_t representa a componente de tendência, obtida pela equação (3.6) e

m é o horizonte de previsão.

$$S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (3.5)$$

onde α é o peso atribuído à observação x_t , $0 < \alpha < 1$.

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (3.6)$$

onde β é o coeficiente de alisamento, análogo a α .

3.4.2.4. Alisamento Exponencial Sazonal e Linear de Winter

Este método produz resultados similares ao Alisamento Exponencial Linear, sendo, no entanto, capaz de manipular séries temporais que além de apresentarem tendência nos dados, apresentam também sazonalidade [WHEELWRIGTH1985].

As equações (3.7), (3.8), (3.9) e (3.10) definem o referido método de previsão.

$$\hat{S}_t = \alpha \frac{x_t}{I_{t-l}} + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (3.7)$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (3.8)$$

$$I_t = \gamma \frac{x_t}{S_t} + (1 - \gamma)I_{t-l} \quad (3.9)$$

$$F_{t+m} = (S_t + mT_t)I_{t-l+m} \quad (3.10)$$

onde I_t corresponde ao alisamento do fator de sazonalidade $\frac{x_t}{S_t}$;

l é o intervalo da sazonalidade e

γ corresponde ao peso atribuído ao fator de sazonalidade.

3.4.3. Métodos Avançados de Previsão de Séries Temporais

No universo dos métodos de previsão de séries temporais mais complexos encontram-se os modelos Autoregressivo e Médias Móveis (AR, MA e ARMA), modelos Autoregressivo Integrado de Médias Móveis (ARIMA), Filtros de Kalman e AEP, modelos ARARMA de Parzen, modelos ARMA Multivariáveis (MARMA), entre outros [WHEELWRIGTH1985]. Os métodos assim classificados obtêm a previsão de algum valor futuro da série temporal pela combinação dos valores reais passados e/ou dos erros ocorridos.

Os modelos AR, MA, ARMA e ARIMA são descritos a seguir.

3.4.3.1. Modelos Autoregressivo e de Médias Móveis

[WHEELWRIGTH1985] descreve três procedimentos capazes de representar as observações de uma série temporal estacionária⁴: modelo *Autoregressivo* (AR), modelo de *Médias Móveis* (MA) e o modelo *Autoregressivo e de Médias Móveis* (ARMA).

a) Modelo Autoregressivo

A especificação de um modelo Autoregressivo (AR) é dada pela equação (3.11) [WHEELWRIGTH1985].

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + e_t \quad (3.11)$$

onde x_t corresponde à observação da série temporal no tempo t ,

ϕ_p corresponde ao parâmetro do modelo AR de ordem p e

e_t representa o erro de eventos aleatórios que não podem ser explicados pelo modelo.

Caso as observações da série temporal possam ser representadas pela equação (3.11), a ordem do modelo puder ser determinada e os parâmetros estimados⁵, é possível prever o valor futuro da série em análise [WHEELWRIGTH1985].

b) Modelo de Médias Móveis

Um modelo de Médias Móveis (MA) fica definido conforme equação (3.12) [WHEELWRIGTH1985].

⁴ A ausência de tendência na série temporal é referida como estacionariedade [WHEELWRIGTH1985].

⁵ Normalmente, os parâmetros do modelo são estimados a partir da série histórica dos dados através de uma análise de regressão [BOX1976] [JUDGE1988].

$$x_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (3.12)$$

onde e_t representa o erro de eventos aleatórios que não podem ser explicados pelo modelo.

θ_q corresponde ao parâmetro do modelo MA de ordem q .

A equação (3.12) é similar à equação (3.11), exceto pelo fato de que o valor previsto para a observação x_t depende dos valores dos erros observados em cada período passado, ao invés das observações propriamente ditas [WHEELWRIGTH1985].

c) Modelo Autoregressivo e de Médias Móveis

Wheelwright e Makridakis [WHEELWRIGTH1985] especificam o modelo misto Autoregressivo e de Médias Móveis (ARMA) através da equação (3.13), como sendo a combinação dos modelos AR e MA.

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (3.13)$$

Analisando a equação (3.13) é possível verificar que os modelos ARMA relacionam os valores futuros com as observações passadas, assim como também com os erros passados apurados entre os valores reais e os previstos.

3.4.3.2. Modelo de Box e Jenkins

O destaque atribuído ao modelo de George Box e Gwilyn Jenkins [BOX1976], que também pode ser incluído nesta classificação, é devido principalmente a sua fundamentação teórica, sendo a princípio capaz de manipular séries temporais de qualquer natureza.

O método de Box e Jenkins consiste na busca de um modelo ARIMA (AutoRegressive Integrate Moving Average) que represente o processo estocástico gerador da série temporal, a partir de um modelo ARMA aplicável na descrição de séries temporais estacionárias, estendendo esse conceito para séries temporais não-estacionárias [NELSON1973].

Genericamente, um processo ARIMA(p,d,q) pode ser representado pela equação (3.14).

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (3.14)$$

sendo $w_t = x_t - x_{t-d}$,

onde ϕ_p e θ_q são os parâmetros dos processos Autoregressivo e de Média Móvel de ordem p e q (ARMA(p,q));

e_t corresponde ao erro de eventos aleatórios que não podem ser explicados pelo modelo e

d equivale ao grau de homogeneidade não-estacionária⁶.

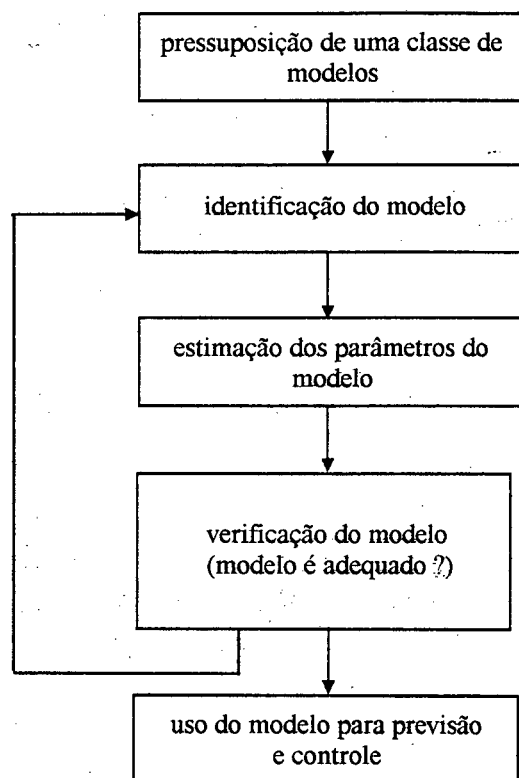
De acordo com [BOX1976], a estratégia para construção de um modelo ARIMA envolve uma abordagem iterativa que pode ser sumarizada conforme ilustra a figura 3.2.

O objetivo da *identificação* é determinar os valores de p , d e q do modelo ARIMA(p,d,q) [WHEELWRIGTH1985]. Inicialmente, a série temporal X^T é diferenciada para se obter uma série estacionária. Com isso, o processo fica reduzido a um modelo ARMA(p,q). Em seguida, a ordem do processo ARMA é identificada pela análise dos coeficientes de autocorrelação e autocorrelação parcial⁷. Ainda nesta etapa são efetuadas estimativas preliminares dos parâmetros do modelo identificado.

⁶ Uma série temporal que se torna estacionária após a diferenciação de d períodos é dita ser homogênea não-estacionária de grau d [JUDGE1988].

⁷ A análise dos coeficientes de autocorrelação e autocorrelação parcial condicionará os valores assumidos por p e q , respectivamente [BOX1976].

Após a identificação do modelo que seja uma representação adequada do mecanismo gerador da série, a *estimação* dos parâmetros desse modelo é efetuada. [JUDGE1988] coloca que os parâmetros do processo AR são estimados através de métodos de regressão; caso o processo MA esteja envolvido, a estimação dos parâmetros deste modelo é obtida pela aplicação de algum algoritmo de otimização não-linear.



Fonte : [BOX1976]

Figura 3.2. Estágios da construção de um modelo ARIMA.

Estimado o modelo, a *verificação* de sua habilidade em representar os fenômenos observáveis da série temporal é confirmada pela análise dos erros do modelo proposto. Caso a inadequação fique evidenciada, o ciclo de identificação, estimação e verificação é novamente aplicado, até que a representação apropriada seja encontrada.

Após a validação do modelo, a *previsão* dos valores futuros da série temporal modelada pode, enfim, ser obtida.

3.5. Considerações sobre os Métodos de Previsão de Séries Temporais

Makridakis e Wheelwright [WHEELWRIGHT1985] investigaram o poder preditivo de vários métodos comumente utilizados na previsão de séries temporais. Pela comparação dos resultados alcançados, identificaram situações onde esses métodos apresentam melhor desempenho e definiram alguns critérios para a seleção dos procedimentos de previsão, pela confrontação dos objetivos a atingir.

O estudo desenvolvido por tais pesquisadores constatou que o incremento da complexidade e da sofisticação estatística dos métodos de previsão de séries temporais não implica, necessariamente, em uma melhora na acuidade da previsão. *“Métodos simples de previsão podem apresentar desempenho extremamente satisfatório sob certas condições”* [WHEELWRIGHT1985]. Além disso, métodos de previsão menos complexos normalmente permitem alcançar total compreensão de suas suposições e limitações, e de interpretação de seus resultados. Assim, antes de se adotar um método de previsão mais complexo, é necessário avaliar os benefícios que um método dessa natureza pode gerar em relação ao custo de sua aplicação.

Segundo [WHEELWRIGHT1985], como a acuidade de uma previsão fica determinada não apenas pelo horizonte de previsão especificado, mas também pelas características das observações da série temporal investigada, a otimização desse critério pode ser alcançada com a aplicação de mais de um método de previsão. A combinação de previsões ou a verificação da consistência desses valores permite aumentar a confiabilidade da previsão e reduzir a possibilidade de grandes desvios.

Mentzer e Cox [MENTZER1984] também analisaram a frequência de aplicação de diferentes métodos de previsão de séries temporais considerando como critérios o horizonte de previsão, familiaridade com os procedimentos preditivos e nível de satisfação decorrente do emprego desses procedimentos. Em relação ao horizonte de previsão (tabela 3.1), o Alisamento Exponencial Simples e a Média Móvel são mais frequentemente usados para períodos curtos, menos para médio prazo e menos ainda

para longo prazo, comprovando evidências empíricas que indicam melhor desempenho destes métodos para horizonte de previsão menores [WHEELWRIGTH1985]. Independente do horizonte de previsão adotado, pouca aplicabilidade do método de Box e Jenkins pode ser observada.

Método	Horizonte de Previsão		
	Curto	Médio	Longo
Alisamento Exponencial Simples	24	17	6
Média Móvel	24	22	5
Decomposição	9	13	5
Box e Jenkins	5	6	2

Fonte : [MENTZER1984].

Tabela 3.1. Frequência relativa de uso de técnicas de previsão em relação ao horizonte de previsão (em %).

A tabela 3.2 sintetiza os estudos de [MENTZER1984] quanto à familiaridade do grupo entrevistado com diferentes métodos de previsão. Analisando os resultados dessa pesquisa, fica evidente que métodos mais simples de previsão são mais conhecidos que métodos avançados de séries temporais.

Método	Familiaridade		
	Muito familiar	Vagamente familiar	Completamente não familiar
Média Móvel	85	7	8
Alisamento Exponencial Simples	73	12	15
Decomposição	42	9	49
Box e Jenkins	26	9	65

Fonte : [MENTZER1984].

Tabela 3.2. Frequência relativa dos graus de familiaridade com os diferentes métodos de previsão (em %).

A satisfação ao se adotar algum desses métodos de previsão também foi estudada e os resultados descritos na tabela 3.3.

Método	Satisfação		
	Satisfatório	Indiferente	Insatisfatório
Alisamento Exponencial Simples	60	19	21
Média Móvel	58	21	21
Decomposição Clássica	55	14	31
Box e Jenkins	30	13	57

Fonte : [MENTZER1984].

Tabela 3.3. Frequência relativa dos graus de satisfação com os diferentes métodos de previsão (em %).

Analisando a tabela 3.3, pode ser observado que o Alisamento Exponencial Simples foi considerado o mais satisfatório entre os métodos de previsão avaliados. Esse resultado é consistente com estudos empíricos [WHEELWRIGTH1985] que constataram ser esse método capaz de alcançar considerável acuidade, além de ser de fácil compreensão e utilização. O método de Médias Móveis também apresentou alto nível de satisfação, enquanto que o método de Box e Jenkins ficou caracterizado como o de menor satisfação e maior insatisfação.

3.6. Medidas de Acuidade dos Métodos de Previsão de Séries Temporais

De acordo com Makridakis [WHEELWRIGTH1985], a suposição básica de qualquer técnica de previsão de séries temporais é que o valor observado na série fica determinado por um padrão que se repete no tempo e por alguma influência aleatória. Isto significa dizer que mesmo quando o padrão exato que caracteriza o comportamento da série temporal tenha sido isolado, algum desvio ainda existirá entre os valores da previsão e os valores realmente observados. Essa aleatoriedade não pode ser prevista;

entretanto, se isolada, sua magnitude pode ser estimada e usada para determinar a variação ou erro entre as observações e previsões realizadas.

A acuidade de um método de previsão pode ser mensurada através de muitas medidas de erro, dentre as quais pode-se citar [WHEELWRIGTH1985] :

$$\text{erro médio} = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)}{n} \quad (3.15)$$

onde x_i é o valor observado no instante i ;

\hat{x}_i é o valor previsto no instante i e

n corresponde ao número de previsões efetuadas.

$$\text{erro absoluto médio} = \frac{\sum_{i=1}^n |(x_i - \hat{x}_i)|}{n} \quad (3.16)$$

$$\text{erro quadrado médio} = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n} \quad (3.17)$$

$$\text{erro percentual absoluto} = \left| \frac{(x_i - \hat{x}_i)}{x_i} \right| (100) \quad (3.18)$$

$$\text{erro percentual médio} = \frac{\sum_{i=1}^n \frac{(x_i - \hat{x}_i)}{x_i}}{n} (100) \quad (3.19)$$

$$\text{erro percentual absoluto médio} = \frac{\sum_{i=1}^n \left| \frac{(x_i - \hat{x}_i)}{x_i} \right|}{n} (100) \quad (3.20)$$

Dessa forma, a verificação da adequação de um determinado modelo supostamente representativo da série histórica de dados é dependente da medida de erro adotada para efetuar essa validação.

CAPÍTULO IV

4 - REDES NEURAIS

Este capítulo descreve a estrutura e comportamento de redes neurais biológicas, de modo a estabelecer uma analogia com os modelos de redes neurais artificiais. Na sequência, uma possível definição de rede neural, sua organização, características e histórico da abordagem baseada nestes modelos são apresentados. Também é descrita a forma de representação do conhecimento em uma rede neural e como o conhecimento contido em sua estrutura é utilizado no processo de aprendizado, e conseqüente realização de previsões.

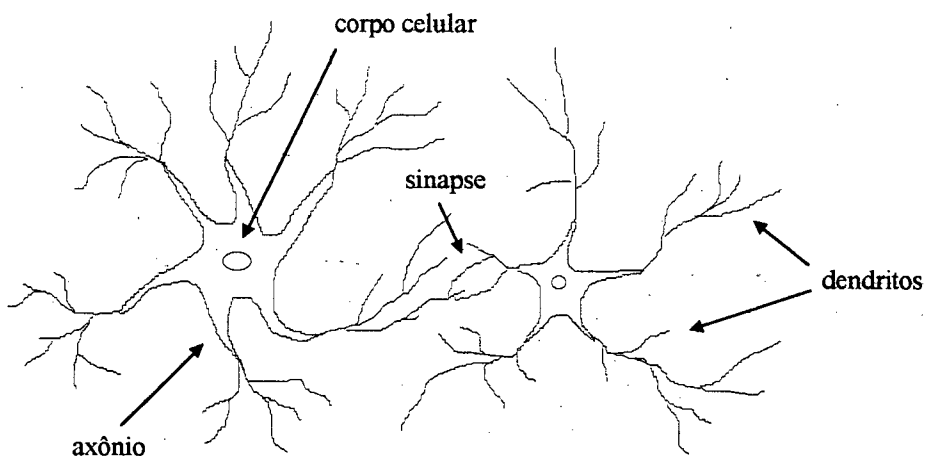
4.1. Redes Neurais Biológicas

O cérebro humano é o dispositivo mais complexo estudado pelo homem, sendo ainda muito pobremente compreendido [BEALE1990]. Ainda não se tem respostas satisfatórias para as questões mais fundamentais tais como "*o que é mente ?*" e "*como eu penso ?*". Entretanto, a compreensão geral da operação do cérebro já é possível.

O sistema nervoso humano é formado por uma rede de *neurônios* (figura 4.1), unidade básica do cérebro, responsável pelos fenômenos conhecidos como pensamento, emoção e cognição, além da execução das funções sensório-motoras e autônomas [WASSERMAN1989]. Cada neurônio compartilha várias características com outras células, mas possui capacidades singulares para receber, processar e transmitir sinais eletroquímicos ao longo das fibras nervosas, que compreendem o sistema de comunicação cerebral. [STEVENS1985] observa a célula nervosa como uma unidade de processamento analógica independente, capaz de estabelecer a comunicação com o sistema nervoso central.

Conforme [BEALE1990], um neurônio é conectado a vários outros neurônios através dos *dendritos* e do *axônio*. Os dendritos - uma complexa rede de prolongamentos - recebem impulsos nervosos de outros neurônios e os conduzem ao *corpo celular* ou *núcleo*. Essas informações são somadas, gerando novos impulsos. Caso o resultado dessa soma exceda um determinado limiar, o axônio transmitirá esse estímulo a outros neurônios através de fenômenos químicos denominados *sinapses*. A força sináptica da conexão neural ao refletir o nível de excitação ou inibição entre neurônios adjacentes, capacita o cérebro humano ao armazenamento do conhecimento e o conseqüente aprendizado.

Através das sinapses os neurônios se unem funcionalmente, formando *redes neurais* [BEALE1990].



Fonte : [WASSERMAN1989].

Figura 4.1. Neurônio biológico.

4.2. Redes Neurais Artificiais

A abordagem de *redes neurais artificiais* consiste em capturar os princípios básicos de manipulação de informação do cérebro humano e aplicar esse conhecimento na resolução de problemas que exigem aprendizado a partir da experiência [BEALE1990].

4.2.1. Definição

As redes neurais artificiais se fundamentam nos estudos sobre a estrutura do cérebro humano para tentar emular sua forma inteligente de processar informações. Apesar de se desconhecer a maneira pela qual o cérebro manipula informações complexas [BEALE1990], sabe-se que a modelagem do conhecimento contido em um problema específico pode ser representada através de inter-conexões entre células nervosas. Estruturalmente, a rede neural artificial, também conhecida por modelo conexionista de computação, se assemelha à rede neural biológica pela composição de seus neurônios e pela conexão entre eles [WASSERMAN1989].

“Um modelo conexionista é uma estrutura de processamento de informações distribuída e paralela. Ela é formada por unidades de processamento, comumente chamadas de nós, neurônios ou células, interconectadas por arcos unidirecionais, também chamados de ligações, conexões ou sinapses. Os nós possuem memória local e podem realizar operações de processamento de informação localizada. Cada célula possui uma única saída (axônio), a qual pode se ramificar em muitas ligações colaterais (cada ramificação possuindo o mesmo sinal de saída do neurônio). Todo o processamento que se realiza em cada unidade deve ser completamente local, isto é, deve depender apenas dos valores correntes dos sinais de entrada que chegam dos neurônios através das conexões. Estes valores atuam sobre os valores armazenados na memória local da célula” [HECHT1988].

4.2.2. Histórico das Redes Neurais Artificiais

De acordo com [MASSON1990], o primeiro modelo formalizado de redes neurais artificiais foi proposto por McCulloch e Pitts, em 1943. Essa estrutura abstraía a complexidade da atividade neural em sistemas neurais reais, assim como complicadas

características encontradas no corpo de neurônios biológicos, formando a base para a maioria dos modelos conexionistas desenvolvidos posteriormente.

[HEBB1949] publica *The Organization of Behavior*, assumindo que a aprendizagem do conhecimento representado em uma rede neural seja alcançada pelo fortalecimento das conexões entre neurônios adjacentes, sempre que esses estiverem excitados.

Em 1958, Frank Rosenblatt criou o *perceptron*, um modelo cognitivo que consistia de unidades sensoriais conectadas a uma única camada de neurônios de McCulloch e Pitts, capaz de aprender tudo o que pudesse representar. Rosenblatt demonstrou que, se fossem acrescentadas sinapses ajustáveis, as redes neurais de McCulloch e Pitts poderiam ser treinadas para classificar padrões em classes linearmente separáveis, convergindo em um número limitado de passos [MÁSSON1990].

No início da década de 60, Widrow e Hoff publicam um artigo no qual especificam um neurônio artificial baseado no modelo de McCulloch e Pitts, denominado *adaline*. Conforme [MÁSSON1990], a importância desse modelo está associada à regra de aprendizagem proposta, a regra Delta.

A publicação de *Perceptrons* de Minsky e Papert, em 1969, expôs as limitações do modelo de Rosenblatt, provando que tais redes não são capazes de resolver uma ampla classe de problemas devido às restrições de representação [MÁSSON1990] [WASSERMAN1989]. O impacto desta publicação foi devastador, praticamente desaparecendo o interesse em redes neurais artificiais nos anos seguintes. Somente poucos pesquisadores como Malsburg, Grossberg, Kohonen e Anderson permaneceram concentrando suas atividades na abordagem conexionista.

A partir dos anos 80, com o avanço da tecnologia e o fracasso da escola simbolista na solução de determinados tipos de problemas, as redes neurais artificiais passaram a atrair substancial atenção novamente [WASSERMAN1989].

Em 1982, a introdução do modelo conexionista conhecido pelo nome de seu idealizador, John Hopfield, permitiu esclarecer, pelas suas características computacionais e estabilidade, boa parte das dúvidas até então existentes em relação ao processo dinâmico executado por certas redes neurais. No mesmo ano, Kohonen publicou um artigo descrevendo uma rede neural artificial baseada em auto-organização e nas características de aprendizado adaptativo do cérebro humano [MÁSSON1990].

Hinton e Seynowsky, em 1983, estenderam o modelo de Hopfield com a incorporação de dinâmica estocástica. Este modelo de rede neural passou a ser conhecido como Máquina de Boltzmann [WASSERMAN1989].

Cerca de dois anos mais tarde, Rumelhart, Hinton e Williams aperfeiçoaram a idéia de perceptron, criando o algoritmo *Backpropagation* [RUMELHART1986], levando a uma explosão de interesse em redes neurais. Este algoritmo foi aplicado em uma grande variedade de problemas, como na identificação da estrutura de proteínas, hifenização de palavras em inglês, reconhecimento da fala, compressão de imagens e previsão de séries temporais [MÁSSON1990]. O sucesso deste algoritmo estimulou o desenvolvimento de muitas pesquisas em redes neurais artificiais e de uma variedade de modelos cognitivos.

4.2.3. Topologia das Redes Neurais Artificiais

De acordo com [MÁSSON1990], a topologia de uma rede neural artificial pode ser expressa através de um grafo dirigido com pesos $G = (V, A, W)$, onde V corresponde a um conjunto de vértices, A a um conjunto de arcos dirigidos e W a um conjunto de pesos para esses arcos. Cada vértice no grafo representa uma unidade de processamento.

4.2.3.1. Camadas

As pesquisas em redes neurais artificiais levaram ao desenvolvimento dos mais diversos modelos cognitivos, cada qual com suas particularidades e adequados a um tipo de situação.

A estruturação de uma rede neural em camadas é uma importante característica topológica desses modelos [MÁSSON1990]. Em uma rede neural estruturada em camadas, o conjunto de vértices V pode ser particionado em vários subconjuntos disjuntos $V = V^{(0)} \cup V^{(1)} \cup \dots \cup V^{(L)}$ de modo que as unidades de processamento da camada l somente apresentem conexões com as unidades das camadas $l+1$ e $l-1$, onde L corresponde ao número de camadas da rede neural artificial. [YOUNGOHC1991a] define uma rede *fully connected* como sendo ~~onde~~ aquela onde cada unidade de processamento da camada l estabelece conexão com todas as unidades de processamento da camada $l+1$.

Quanto ao número de camadas, as redes neurais podem ser dispostas em uma única camada, configuração mais simples de uma rede neural, ou em múltiplas camadas [BEALE1990] [WASSERMAN1989].

[LIPPMAN1987] ainda classifica as redes neurais artificiais em redes cíclicas, também chamadas de redes recorrentes, e redes acíclicas. A arquitetura de uma rede neural cíclica difere da acíclica por apresentar conexões entre as unidades de processamento pertencentes à mesma camada ou entre unidades de processamento de camadas diferentes cujas saídas passam a ser entradas na camada anterior. As redes recorrentes podem exibir propriedades muito similares à memória de curto termo dos seres humanos, onde o estado da saída da rede depende em parte da entrada anterior.

4.2.3.2. Conexões e Pesos

Os arcos do grafo são chamados de conexões e representam as sinapses entre os neurônios artificiais. A cada conexão no grafo está associado um peso $w_{ij}^{(l)}$, em analogia às sinapses de um modelo conexionista biológico, representando a força de ligação entre as unidades de processamento $v_i^{(l)}$ e $v_j^{(l-1)}$, onde i e j correspondem à posição - respectivamente nas camadas l e $l-1$ -, que essas unidades ocupam na rede [MÁSSON1990].

Conexões com pesos positivos, chamadas excitatórias, indicam o reforço na ativação do neurônio $v_i^{(l)}$; sinapses com pesos negativos, chamadas inibitórias, indicam a inibição na ativação do neurônio $v_i^{(l)}$. Assim, os neurônios artificiais, distribuídos no espaço e ligados por conexões, trocam sinais inibitórios ou excitatórios, competindo ou cooperando entre si. O comportamento inteligente emerge, então, da ação simultânea dessa coletividade, sem a necessidade de elementos centralizadores [MÁSSON1990].

4.2.3.3. Unidades Visíveis e Unidades Ocultas

A interface da rede neural artificial é definida pelas unidades de entrada (V_I), unidades de saída (V_O) e pelas unidades ocultas (V_H) [MÁSSON1990] [YOUNGOHC1991a].

$$V = V_I \cup V_O \cup V_H$$

O conjunto de unidades de entrada e de saída representam as unidades visíveis da rede, sendo dependentes da aplicação que se quer modelar [FREEMAN1992]. Em redes neurais estruturadas em camadas, a camada de entrada é tipicamente $V^{(0)}$ e a camada de saída, $V^{(L)}$.

As unidades ocultas são utilizadas para modificar os dados de entrada, de modo a suportar qualquer função requerida para a entrada ou saída [YOUNGOHC1991b],

impondo uma representação intermediária adicional dos dados de entrada para a saída desejada. Através das unidades ocultas, o modelo conexionista consegue representar abstrações que não poderiam ser diretamente realizadas a partir das unidades de entrada [JONES1987].

4.2.4.Representação do Conhecimento em uma Rede Neural Artificial

O domínio do conhecimento de um problema é representado em um modelo conexionista através das unidades de processamento, que abstraem a estrutura e o comportamento dos neurônios biológicos [BEALE1990].

A figura 4.2 ilustra a estrutura de uma unidade de processamento.

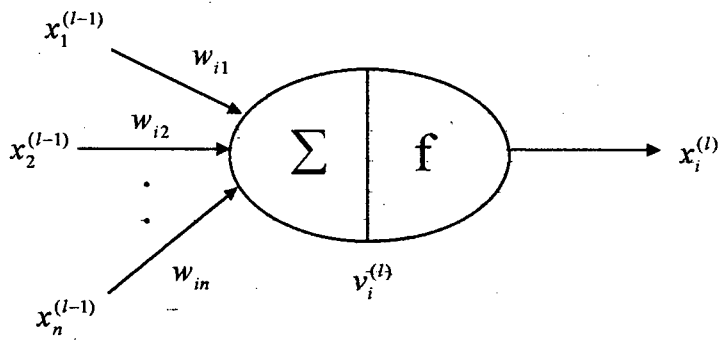


Figura 4.2. Unidade de processamento.

Uma unidade de processamento $v_i^{(l)}$ possui entradas $x_1^{(l-1)}$, $x_2^{(l-1)}$, ..., $x_n^{(l-1)}$, que correspondem aos estados dos neurônios $v_j^{(l-1)}$ com os quais está conectada.

A partir dessas entradas e do conjunto de pesos sinápticos $w_{ij}^{(l)}$, que refletem a força da unidade $v_j^{(l-1)}$ sobre a unidade $v_i^{(l)}$, é calculado o potencial *net* do neurônio $v_i^{(l)}$. Esse potencial no tempo t é determinado por uma *regra de propagação*, que geralmente equivale à soma linear da multiplicação das entradas pelos pesos [MÁSSON1990], conforme equação (4.1).

$$net_i^{(l)}(t) = \sum_{j=1}^{n^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)}(t) - \theta_i^{(l)}(t) \quad (4.1)$$

onde $x_j^{(l-1)}$ é o estado da j-ésima unidade,

$w_{ij}^{(l)}$ é a força sináptica entre a i-ésima unidade e a j-ésima unidade e,

$\theta_i^{(l)}$ é o limiar da i-ésima unidade, representando a força que as entradas das unidades conectadas à unidade $v_i^{(l)}$ precisam atingir para ativar esta unidade.

O potencial é modificado pela aplicação de uma *função de ativação g*, determinando o estado da unidade $v_i^{(l)}$ no instante $t+1$.

$$x_i^{(l)}(t+1) = g\left(net_i^{(l)}(t)\right) \quad (4.2)$$

De acordo com [MÁSSON1990], a função de ativação corresponde a um limiar que restringe a propagação do impulso nervoso à transposição de um certo nível de atividade, mapeando o potencial da unidade de processamento $v_i^{(l)}$ para um intervalo pré-especificado de saída.

Dentre as possíveis funções de ativação pode-se citar a *linear*, a *rampa*, a *salto* e a *sigmóide*.

A função *linear* (figura 4.3 a) é obtida pela equação $g\left(net_i^{(l)}(t)\right) = \alpha net_i^{(l)}(t)$, onde α é uma constante de proporcionalidade que regula a intensidade do $net_i^{(l)}(t)$.

A função *rampa* (figura 4.3 b) é limitada ao intervalo $[-y,+y]$ definida por

$$g\left(net_i^{(l)}(t)\right) = \begin{cases} +y, & \text{se } net_i^{(l)}(t) \geq y \\ net_i^{(l)}(t), & \text{se } |net_i^{(l)}(t)| < y \\ -y, & \text{se } net_i^{(l)}(t) \leq -y \end{cases}$$

onde y representa o ponto de saturação da função.

A figura 4.3 c) ilustra a função *salto* que admite valor +1 se o potencial da unidade de processamento $v_i^{(l)}$ for positivo e -1, caso contrário.

A função *sigmóide* (figura 4.3 d), também conhecida por função *logística*, é expressa matematicamente como

$$g(\text{net}_i^{(l)}(t)) = \frac{1}{1 + e^{-\text{net}_i^{(l)}(t)}} \quad (4.3)$$

sendo uma função contínua, monotonicamente crescente e que gera valores graduais e não lineares no intervalo $[0,1]$.

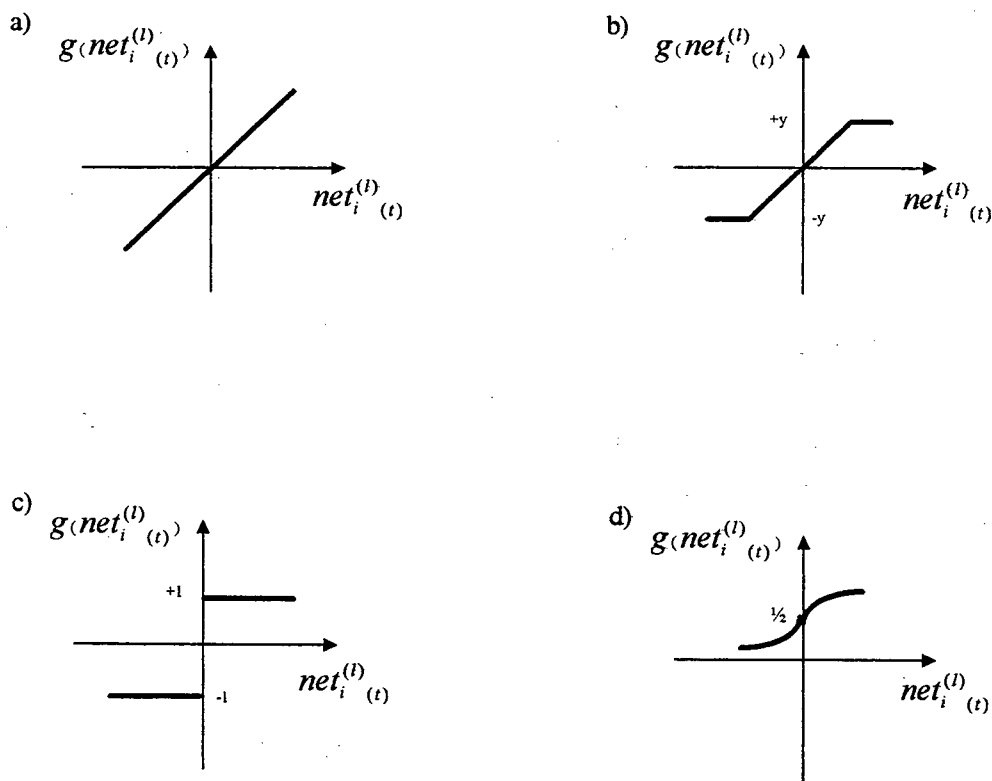


Figura 4.3. Funções de ativação.

Outra função de ativação comumente utilizada é a *tangente hiperbólica* [WASSERMAN1989]. Sua curva é similar à função sigmóide, mas simétrica na origem. Essa função é freqüentemente empregada por biólogos como um modelo matemático de ativação da célula nervosa.

Mesmo conhecendo completamente o comportamento de cada neurônio individual, a composição de várias unidades de processamento em uma estrutura de rede manifesta reações imprevisíveis. Desta forma, é a reunião do estado de ativação de todas as unidades de processamento que especifica o que está sendo representado na rede neural artificial em um determinado instante. É essa emergência de propriedades de conjunto que determina o interesse e a complexidade dos modelos conexionistas.

4.2.5. Características das Redes Neurais Artificiais

Por serem baseadas nas redes neurais biológicas, as redes neurais artificiais apresentam um surpreendente número de características observadas no processo cognitivo humano [WASSERMAN1989], como o aprendizado pela experiência, a generalização a partir de exemplos e a abstração de características essenciais de informações que contém fatos irrelevantes.

4.2.5.1. Aprendizado

[CARBONELL1989] define o conceito de aprendizado como a habilidade de realizar tarefas novas que não podiam ser realizadas anteriormente, ou melhorar a realização de tarefas antigas, como resultado de mudanças produzidas pelo processo de aprendizado.

As redes neurais artificiais podem modificar seu comportamento em resposta aos estímulos produzidos pelo ambiente, regulando a força da conexão entre unidades de

processamento adjacentes pela adaptação dos pesos sinápticos, reconhecendo as informações apresentadas às suas unidades visíveis [WASSERMAN1989].

4.2.5.2. Generalização

Segundo [WASSERMAN1989], um modelo conexionista é sensível às variações que podem ocorrer em informações procedentes de suas unidades de entrada, reconhecendo ruído e distorção. A capacidade da rede em se adaptar às novas situações, gerando valores de saída consistentes com os esperados, é vital para a aplicabilidade do modelo em um ambiente do mundo real.

Embora a maioria das pesquisas em redes neurais artificiais tenham concentrado seus esforços na redução dos tempos de aprendizagem, a característica mais importante de um modelo conexionista é a habilidade em generalizar sobre o domínio do problema [REFENES1993b].

O bom desempenho da generalização depende, entre outros fatores, do número de parâmetros livres da rede neural artificial [LECUN1989]. É desejável diminuir o tamanho das conexões sem, entretanto, reduzir o tamanho da rede ao ponto onde não se possa computar a função desejada.

4.2.5.3. Abstração

Alguns modelos de redes neurais artificiais são capazes de abstrair a essência do conjunto de dados a elas apresentados [WASSERMAN1989], permitindo, dessa forma, a classificação ou reconhecimento de padrões incompletos.

4.2.6. Aprendizado da Rede Neural Artificial

Dentre todas as características das redes neurais artificiais, nenhuma desperta tanto interesse quanto a sua habilidade em realizar o aprendizado [WASSERMAN1989].

Conforme [MÁSSON1990], o aprendizado em um modelo de redes neurais artificiais é decorrente do treinamento da rede através da apresentação de padrões às suas unidades visíveis.

O objetivo do treinamento consiste em atribuir os pesos sinápticos com valores apropriados, de modo a produzir o conjunto de saídas desejadas ou ao menos consistentes com um intervalo de erro estabelecido [FREEMAN1992]. Desta forma, o processo de aprendizado subsiste na busca de um espaço de pesos pela aplicação de alguma regra que defina esta aprendizagem [MÁSSON1990].

Em geral, as regras de aprendizado podem ser consideradas variantes da Regra de Hebb [HEBB1949]. Na essência, Hebb propõe que a sinapse conectando dois neurônios seja reforçada sempre que ambos os neurônios estiverem ativos.

Uma rede neural artificial que tenha a regra de Hebb como regra de aprendizado modifica os pesos sinápticos entre as conexões das unidades de processamento $v_i^{(l)}$ e $v_j^{(l-1)}$ proporcionalmente ao produto dos níveis de excitação desses neurônios, conforme equação (4.4).

$$\Delta w_{ij}^{(l)} = \eta \cdot x_i^{(l)} \cdot x_j^{(l-1)} \quad (4.4)$$

onde $\Delta w_{ij}^{(l)}$ corresponde a alteração no valor do peso $w_{ij}^{(l)}$ e

η é uma constante de proporcionalidade que reflete a evolução do processo de aprendizado pela busca no espaço de pesos.

Como adaptação à regra de Hebb, a regra Delta modifica os pesos de acordo com a variação entre a saída desejada e a observada no treinamento [MÁSSON1990].

A equação (4.5) atualiza os pesos associados aos arcos da rede neural artificial pela aplicação da regra Delta

$$\Delta w_{ij}^{(l)} = \eta \delta x_i^{(l)} \quad (4.5)$$

com $\delta = (d_i^{(l)} - x_i^{(l)})$, onde $d_i^{(l)}$ representa a saída desejada para a unidade de processamento $v_i^{(l)}$.

4.2.7. Estratégias de Aprendizado

Uma rede neural artificial deve ser ajustada para que a aplicação de um conjunto de entradas produza a saída desejada. Esse ajustamento, obtido pelo treinamento da rede, pode ser feito das seguintes formas [MÁSSON1990] [WASSERMAN1989]:

- sem treinamento : os valores dos pesos sinápticos são estabelecidos explicitamente.
- treinamento supervisionado : a rede é treinada pela apresentação dos vetores de entrada e seus respectivos vetores de saída, chamados de pares de treinamento.
- treinamento não supervisionado : o treinamento consiste da apresentação apenas dos vetores de entrada, a partir dos quais são extraídas as características desse conjunto de padrões, agrupando-os em classes. O treinamento não supervisionado pode ser observado como um processo autônomo ou auto-organizável.

4.2.8. Algoritmo de Aprendizado *Backpropagation*

Por muitos anos não se teve um algoritmo eficiente para treinar redes neurais artificiais de múltiplas camadas. Desde que as redes de uma única camada se mostraram limitadas naquilo que poderiam representar e, portanto, no que poderiam aprender, o desenvolvimento de modelos cognitivos deixou de ser um campo atraente e poucas pesquisas foram realizadas na área.

O algoritmo *backpropagation*, proposto por Werbos, Parker e Rummelhart, fez ressurgir o interesse em redes neurais artificiais, sendo o algoritmo de aprendizado mais largamente utilizado [MÁSSON1990] [REFENES1993c].

Conforme [BEALE1990], o *backpropagation* pode ser visto como uma generalização do método Delta para redes neurais de múltiplas camadas. Ao se apresentar um determinado padrão de entrada a uma rede neural não treinada e o respectivo padrão de saída, uma saída aleatória é produzida. A partir da saída produzida pela rede é calculado um erro, representando a diferença entre o valor obtido e o desejado. O objetivo consiste, então, em reduzir continuamente o erro até um determinado valor aceitável. Isto é alcançado pelo ajuste dos pesos entre as conexões dos neurônios pela aplicação da regra Delta Generalizada, que calcula o erro para alguma unidade particular e propaga esse erro para a camada anterior. Cada unidade tem seus pesos ajustados de modo a minimizar o erro da rede.

A minimização do erro no algoritmo *backpropagation* é obtida pela execução do *gradiente decrescente* na superfície de erros do espaço de pesos, onde a altura para qualquer ponto no espaço de pesos corresponde à medida do erro. O ajuste dos pesos inicia nas unidades de saída, onde a medida do erro está disponível, e procede com a retropropagação desse erro entre as camadas, ajustando os pesos até que a camada das unidades de entrada tenha sido processada. Para as unidades de saída, como são conhecidos os valores desejados e obtidos, o ajuste dos pesos sinápticos é relativamente simples; para as unidades das camadas ocultas, o processo não é tão trivial. Intuitivamente, as unidades ocultas que apresentarem erros grandes devem ter suas conexões bastante alteradas, enquanto que a mudança nos pesos daquelas que tiverem

suas saídas muito próximas das desejadas deverá ser pequena. Na realidade, os pesos para um neurônio particular devem ser ajustados na proporção direta ao erro da unidade de processamento a qual está conectado. Essa é a razão pela qual a retropropagação dos erros através da rede permite o correto ajuste dos pesos sinápticos entre todas as camadas do modelo conexionista.

Assim, é possível identificar duas fases distintas no processo de aprendizagem do *backpropagation* : aquela onde as entradas se propagam entre as camadas da rede, da camada de entrada até a camada de saída, e aquela em que os erros são propagados na direção contrária ao fluxo de entrada.

4.2.8.1. Etapas do Algoritmo *Backpropagation*

Seja o conjunto de treinamento $D = \{ (d_I^\lambda, d_O^\lambda) : \lambda = 1, \dots, \Lambda \}$, onde d_I^λ são os vetores de entrada, d_O^λ os correspondentes vetores de saída e Λ o número de vetores do conjunto de treinamento.

Conforme [FREEMAN1992] e [MÁSSON1990], o objetivo do treinamento consiste em ensinar à rede o mapeamento de todo vetor de entrada para o respectivo vetor de saída, isto é, estabelecer $x_o (w, d_I^\lambda) = d_O^\lambda, \forall 1 \leq \lambda \leq \Lambda$, pela configuração de valores apropriados para os pesos das conexões da rede.

O ajustamento desses pesos, executado sempre que a saída desejada d_o e a saída real x_o não forem coincidentes, é feito pela aplicação do método *gradiente decrescente* (equação (4.6)).

$$w_{ij}^{(l)}(t+1) = w_{ij}^{(l)}(t) + \Delta w_{ij}^{(l)} \quad (4.6)$$

com $\Delta w_{ij}^{(l)} = -\eta \frac{\partial \varepsilon}{\partial w_{ij}^{(l)}}$, onde η é a taxa de aprendizagem e ε , a medida do erro.

O erro ε é definido como uma soma sobre todos os padrões de treinamento, conforme equação (4.7).

$$\varepsilon = \sum_{\lambda=1}^{\Lambda} \varepsilon^{\lambda} \quad (4.7)$$

Como medida de erro para um padrão de treinamento específico λ tem-se

$$\varepsilon^{\lambda} = \frac{1}{2} \sum_j (d_{o,j}^{\lambda} - x_{o,j}^{\lambda})^2 \quad (4.8)$$

onde $d_{o,j}^{\lambda}$ é o j-ésimo componente do vetor de saída d_o^{λ} . O aprendizado é reduzido, portanto, à minimização da medida de erro em (4.7).

As etapas do *backpropagation* podem ser sumarizadas nos seguintes passos :

1. inicializar o conjunto de pesos W com valores pequenos e aleatórios;
2. apresentar o vetor de entrada do padrão de treinamento $x^{(0)} = d_i^{\lambda}$;
3. calcular a resposta da rede $x^{(L)} = x^{(L)}(w, x^{(0)})$;
4. calcular o erro para as unidades de saída

$$\delta_i^{(L),\lambda} = g'(net_i^{(L)}) (d_{o,i}^{\lambda} - x_i^{(L)}) \quad (4.9)$$

onde g' é a derivada da função de ativação em relação à $net_i^{(L)}$;

5. calcular o erro para as unidades ocultas

$$\delta_i^{(L),\lambda} = g'(net_i^{(L)}) \sum_k \delta_k^{(l+1),\lambda} w_{ki}^{(l+1)} \quad (4.10)$$

6. atualizar os pesos de todas as camadas

$$\Delta w_{ij}^{(l),\lambda} = \eta \delta_i^{(l),\lambda} x_j^{(l-1)} \quad (4.11)$$

Os padrões de treinamento são apresentados sucessivamente às unidades visíveis da rede neural artificial, até que um erro aceitável seja alcançado ou enquanto um número determinado de iterações não for satisfeito. O último conjunto de pesos observado entre as conexões das células é então mantido para testar a habilidade da rede em mapear a função de entrada para saída e a consequente validação do modelo de redes neurais artificiais [FREEMAN1992].

4.2.9. Considerações sobre o Algoritmo *Backpropagation*

O desempenho do algoritmo de aprendizagem *backpropagation* está condicionado à modelagem adotada na rede neural artificial e ao conjunto de dados utilizados no processo de ajuste dos pesos sinápticos entre as conexões da rede.

4.2.9.1. Dados para Treinamento

Conforme James Freeman [FREEMAN1992], não existe critério específico para seleção dos vetores de treinamento. É possível utilizar todos os dados disponíveis no treinamento do modelo conexionista, embora apenas um subconjunto desses dados talvez seja suficiente para que esse processo seja executado com sucesso. Os dados restantes podem ser usados para avaliar a capacidade de generalização do *backpropagation* no mapeamento de entradas nunca encontradas no treinamento para saídas consistentes.

4.2.9.2. Tamanho da Rede Neural Artificial

[REFENES1993c] coloca que o número de unidades de processamento das camadas de entrada e saída é usualmente determinado pela aplicação. No caso das camadas ocultas, a relação não é tão transparente. O ideal é utilizar o menor número possível de unidades ocultas para que a generalização não fique prejudicada

[RUMELHART1986]. Se o número de neurônios ocultos for muito grande, a rede acaba memorizando os padrões apresentados durante o treinamento. Contudo, se a arquitetura das camadas ocultas possuir unidades de processamento em número inferior ao necessário, o algoritmo *backpropagation* pode não conseguir ajustar os pesos sinápticos adequadamente, impedindo a convergência para uma solução.

A experiência ainda é a melhor indicação para a definição da topologia de um modelo conexionista [SURKAN1990].

4.2.9.3. Pesos e Parâmetros de Aprendizado

[FREEMAN1992] sugere que os pesos das conexões entre as camadas de uma rede neural sejam inicializados com valores aleatórios e pequenos para que se evite a saturação da função de ativação e a conseqüente incapacidade de realizar a aprendizagem.

À medida que o treinamento evolui, os pesos sinápticos podem passar a assumir valores maiores, forçando a operação dos neurônios na região onde a derivada da função de ativação é muito pequena. Como o erro retropropagado é proporcional a esta derivada, o processo de treinamento tende a se estabilizar, levando a uma paralisação da rede sem que a solução tenha sido encontrada. Isto pode ser evitado pela aplicação de uma taxa de aprendizagem menor. Teoricamente, o algoritmo de aprendizado exige que a mudança nos pesos seja infinitesimal [RUMELHART1986]. Entretanto, a alteração dos pesos nessa proporção é impraticável, pois implicaria em tempo de treinamento infinito. Em vista disso, é recomendável que a taxa de aprendizado assuma valor maior no início do treinamento e, à medida em que se observe decréscimo no erro da rede, essa taxa também seja diminuída. Diminuindo progressivamente a taxa de atualização dos pesos, o *gradiente decrescente* está apto a alcançar uma solução melhor [BEALE1990].

Outra maneira de aumentar a velocidade de convergência da rede neural artificial treinada pelo algoritmo *backpropagation* é a adoção de um método chamado *momentum* [BEALE1990] [FREEMAN1992] [RUMELHART1986] [WASSERMAN1989]. O

propósito desse método consiste em adicionar, quando do cálculo do valor da mudança do peso sináptico, uma fração proporcional à alteração anterior. Assim, a introdução desse termo na equação de adaptação dos pesos tende a aumentar a estabilidade do processo de aprendizado, favorecendo mudanças na mesma direção. A equação (4.12) especifica o ajuste das conexões entre unidades de processamento pela aplicação do termo *momentum*.

$$\Delta w_{ij}^{(l),\lambda} = \eta \delta_i^{(l),\lambda} x_j^{(l-1)} + \alpha (w_{ij}^{(l),\lambda} - w_{ij}^{(l),\lambda}{}_{(t-1)}) \quad (4.12)$$

onde α representa o termo *momentum*, $0 < \alpha < 1$.

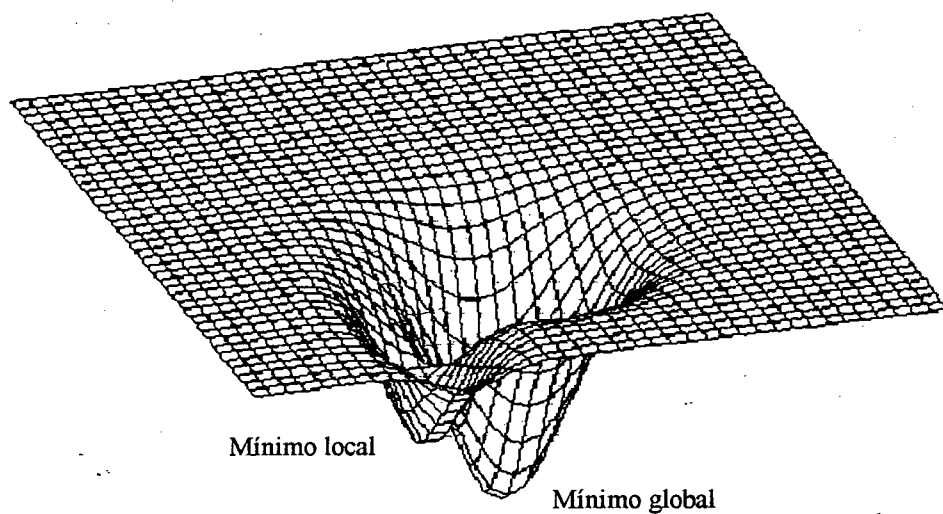
4.2.9.4. Mínimo Local

O *backpropagation* utiliza a heurística do *gradiente decrescente* para ajustar os pesos entre as sinapses, seguindo a curva da superfície dos erros em direção a um ponto mínimo [WASSERMAN1989]. Superfícies de erros convexas, por apresentarem um único mínimo, permitem que este método atinja o mínimo global. Nas superfícies de erros não convexas e altamente convolutas, normalmente encontradas em problemas práticos, a solução alcançada pode não ser a ótima. Nestes casos, haverá que ser utilizado algum algoritmo de otimização global.

Assim que um mínimo é encontrado, seja global ou local, o aprendizado cessa [FREEMAN1992]. Se a rede alcançar um mínimo local (figura 4.4), do seu ponto de vista limitado, todas as direções em sua volta representam valores maiores que o alcançado e, conseqüentemente, a convergência para o mínimo global não é atingida. Nesse caso, a magnitude do erro da rede pode ser muito alta e, portanto, inaceitável.

Caso a rede neural encerre o aprendizado antes que uma solução satisfatória seja obtida, o redimensionamento do número de unidades ocultas ou da taxa de aprendizagem e do termo *momentum* podem ser suficientes para resolver o problema. De acordo com [FREEMAN1992], outra possibilidade para se tentar encontrar o mínimo global é

realizar o treinamento a partir de um conjunto de pesos inicial diferente daquele utilizado anteriormente.



Fonte : [KOVÁCS1996].

Figura 4.4. Mínimo local.

Através da figura 4.4, que ilustra um corte em uma superfície de erros hipotética no espaço de pesos, é possível observar um ponto de mínimo local. Tanto à direita, quanto à esquerda, os valores são maiores que esse mínimo.

CAPÍTULO V

5 - REDES NEURAS ARTIFICIAIS E A PREVISÃO DO PREÇO DE AÇÕES

Neste capítulo é avaliada a possibilidade de previsão dos preços das ações através de técnicas de redes neurais artificiais. É proposto um modelo de rede neural e critérios para avaliação desse modelo. Ainda, é abordado um método de identificação de regularidades presentes nas séries temporais modeladas.

5.1. Análise de Viabilidade

A importância de técnicas de modelagem de redes neurais artificiais no reconhecimento de padrões é evidenciada através de um amplo espectro de aplicações, sendo, provavelmente, a área onde as habilidades destes modelos conexionistas foram mais exploradas [WHITE1988].

Diante do sucesso alcançado pelas redes neurais na identificação do caos determinístico [WHITE1988] surge a possibilidade de investigar a aplicabilidade destes modelos na extração de regularidades não-lineares presentes em séries temporais econômicas. Mais especificamente, no mercado de ações o interesse é dirigido para a detecção de regularidades no movimento da cotação das ações no mercado.

Contra otimismo de que modelos conexionistas possam representar o desempenho observado no mercado acionário está a *hipótese do mercado eficiente*.

A *hipótese do mercado eficiente* [LEVY1969] [WHITE1988] assera que a posse de um conjunto de informações públicas disponíveis em um determinado instante de

tempo não altera o retorno¹ esperado de investimento no mercado. Formalmente, $E(r_t | I_{t-1}) = E(r_t)$, onde $E(r_t | I_{t-1})$ corresponde ao valor esperado do retorno da ação no tempo t , condicionado às informações I conhecidas no tempo $t-1$. Baseado nesta hipótese, o movimento da cotação de uma ação seria completamente imprevisível a partir de informações tais como o preço e o volume passados da ação ou de qualquer outra ação.

Segundo [LEVY1969], a propriedade de ajustamento instantâneo das informações públicas aos preços das ações em um mercado eficiente implica na independência entre variações de preços sucessivos na seqüência das negociações. Um mercado que apresenta tal comportamento é, por definição, um mercado *random walk*. De acordo com o modelo *random walk*, uma série de mudanças de preços não tem memória e, portanto, a história passada da série não pode ser usada para prever o futuro de uma maneira significativa.

Apesar da força da *hipótese do mercado eficiente*, ela ainda é uma teoria e qualquer teoria pode ser refutada com evidência apropriada [WHITE1988]. Conforme [LO1988], estudos empíricos sobre o mercado demonstraram que o movimento das cotações das ações não segue o modelo *random walk*. “Ocorre que em vários testes realizados no mercado, tanto aqui no Brasil como no exterior, verificou-se ... que os preços das próximas negociações são fortemente dependentes dos preços das negociações anteriores, sendo possível, então, prever tendências de preços a partir da observação dos movimentos passados” [SÁ1987].

É a partir desta constatação que a modelagem de redes neurais artificiais se torna uma ferramenta importante na investigação de regularidades no preço das ações mascaradas pelas oscilações do mercado.

¹ O retorno r_t de uma ação é definido como $r_t = (p_t - p_{t-1} + d_t) / p_{t-1}$, onde p_t representa o preço de fechamento da ação no tempo t e d_t corresponde aos dividendos pagos no tempo t [WHITE1988].

5.2. Modelo de Rede Neural Artificial Proposto

As arquiteturas de redes neurais artificiais têm sido muito exploradas nos últimos anos devido às suas interessantes habilidades em realizar o aprendizado, generalizando sobre um determinado domínio de conhecimento [KIMOTO1990] [REFENES1993b]. Entre os numerosos algoritmos de aprendizado propostos pelos modelos conexionistas, o *backpropagation* é o procedimento adotado na maioria das aplicações [MÁSSON1990]. A flexibilidade em manipular os mais diversos problemas, alcançando soluções satisfatórias, garantiu o sucesso desse algoritmo.

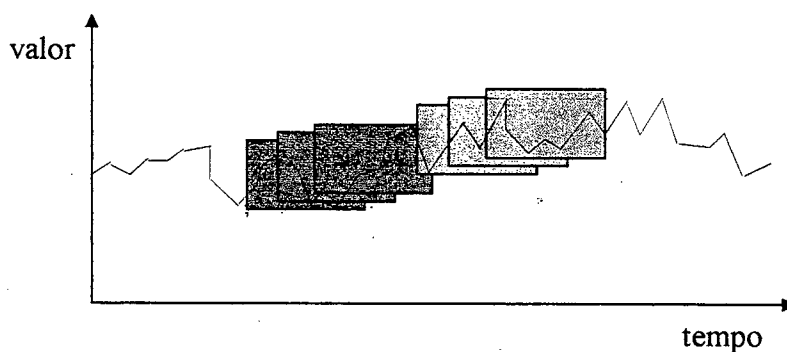
Acredita-se que o *backpropagation* seja um procedimento de aprendizagem efetivo quando o mapeamento da entrada para a saída contenha tanto regularidades quanto exceções, sendo em princípio capaz de solucionar qualquer problema de classificação não-linear [HINTON1987]. Esta característica é particularmente importante para a previsão dos valores futuros de séries temporais, onde muitas das tarefas consistem no tratamento de problemas de classificação não-linear.

[KAMIJO1990] [KIMOTO1990] [SURKAN1990] [WHITE1988] e [YOUNGOHC1991b] obtiveram resultados muito significativos na previsão de séries temporais econômicas com a aplicação do algoritmo desenvolvido por [RUMELHART1986].

Entretanto, ainda não está disponível qualquer metodologia que oriente a construção de uma rede de múltiplas camadas cujo processo de aprendizagem seja realizado pelo *backpropagation*. Assim, o projeto de uma rede neural é um processo exaustivo em busca de uma topologia adequada e de valores apropriados para os parâmetros de aprendizado [REFENES1993a].

5.3. Identificação de Regularidades em uma Série Temporal

De acordo com [REFENES1993a], técnicas de previsão de séries temporais não baseadas em modelos matemáticos dependem da descoberta de regularidades empíricas entre as observações de uma série histórica de dados. A partir dessa consideração é proposto um método para a identificação de regularidades em uma série temporal. A idéia básica consiste na definição de duas *janelas* J^I e J^O de tamanhos fixos n e m , respectivamente, para visualizar as observações da série em diferentes intervalos de tempo, conforme figura 5.1. Para um determinado tamanho de *janela*, a suposição é que a seqüência de valores J_0^I, \dots, J_n^I está de alguma forma relacionada com a seqüência J_0^O, \dots, J_m^O e essa relação, embora desconhecida, está definida completamente no conjunto de dados. Para que o comportamento da série temporal seja capturado, J^I e J^O são deslocados ao longo da série pela aplicação de um passo fixo de tamanho p .



Fonte : [REFENES1993a].

Figura 5.1. *Janelamento* de uma série temporal.

A escolha do tamanho das janelas e do passo de deslocamento é crítica à habilidade em identificar padrões recorrentes no tempo e, portanto, aproximar acuradamente o relacionamento não-observável na série temporal [REFENES1993a] [REFENES1993c].

5.4. Critérios de Avaliação das Previsões

Como a escolha da arquitetura da rede neural artificial e a seleção dos parâmetros que influenciam no aprendizado e, conseqüentemente, na obtenção de uma previsão acurada, são decorrentes de investigação empírica dos resultados, torna-se necessário estabelecer alguns critérios de avaliação do desempenho da rede.

Dois critérios fundamentais a serem observados para a validação do modelo cognitivo proposto são a convergência e a capacidade de generalização da rede neural artificial.

A convergência concerne a capacidade do algoritmo de aprendizado em identificar regularidades não-observáveis na série temporal econômica, sob quais condições isso ocorre e quais os requisitos computacionais dispensados para a sua realização [REFENES1993a]. Visto que o aprendizado do *backpropagation* é executado pela aplicação do *gradiente decrescente* na superfície de erros do espaço de pesos, o alcance do mínimo global não pode ser garantido. Assim, o processo de treinamento do modelo conexionista pode divergir e a solução nunca ser alcançada; ou ainda, para um número de iterações pré-determinado, o erro da rede pode ser extremamente alto. Dessa forma, as propriedades relacionadas com a convergência do *backpropagation* devem ser sempre investigadas experimentalmente [REFENES1993c].

A generalização mede a habilidade da rede neural artificial em gerar previsões acuradas para observações da série temporal não presentes no conjunto de treinamento do modelo, extrapolando sobre o domínio do conhecimento [REFENES1993a] [REFENES1993c]. Caso a topologia da rede tenha sido adequadamente projetada, os parâmetros de aprendizado corretamente estimados e as observações apresentadas durante o treinamento forem representativas do comportamento da série temporal, espera-se que a rede consiga capturar as características essenciais dessas observações, generalizando, portanto, para qualquer situação. De acordo com [REFENES1993b], a capacidade de generalização da rede é de maior importância para avaliar o poder preditivo do modelo conexionista.

CAPÍTULO VI

6 - APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NA PREVISÃO DO PREÇO DE AÇÕES

O objetivo deste capítulo consiste na aplicação da abordagem de modelos conexionistas na manipulação de séries de cotações de ações e na investigação da habilidade desses modelos em prever os preços futuros das ações. O modelo de redes neurais artificiais proposto é especificado e é feita a análise dos resultados obtidos com o aprendizado, considerando os critérios estabelecidos no capítulo anterior. As previsões são comparadas com aquelas obtidas pelos métodos convencionais de previsão de Médias Móveis e Alisamento Exponencial Simples. Também é feita uma crítica quanto à estabilidade dos valores previstos pelas redes neurais artificiais. Ainda, a aplicação dos modelos conexionistas na previsão dos preços de ações, considerando uma carteira simplificada de ações, é analisada.

6.1. Definição do Problema

O propósito deste trabalho consiste na identificação de regularidades não-lineares presentes em séries de preços de ações pela aplicação de técnicas de redes neurais artificiais. Assumindo que tais regularidades possam ser detectadas e que elas persistem na série histórica das cotações das ações, é possível realizar a previsão dos valores futuros dessas ações, investigando, então, a habilidade preditiva dos modelos conexionistas na manipulação de séries temporais econômicas irregulares, comportamento este normalmente observado no mercado acionário.

O *software* NeuralWorks Explorer foi utilizado para efetuar a modelagem das redes neurais artificiais. A este aplicativo foi conectado um programa desenvolvido em C++ que realiza o tratamento dos dados empregados no processo de aprendizagem dos modelos conexionistas e fornece uma interface com o NeuralWorks.

6.2. Modelagem do Problema

A concretização de previsões acuradas de uma determinada série de preços de ações depende da modelagem adequada da rede neural empregada na representação do conhecimento contido nessa série.

A percepção de inter-relações complexas entre os parâmetros que influenciam o desempenho de uma rede neural e as inúmeras possibilidades de projeto da arquitetura do modelo conexionista disponibilizam ampla liberdade para a construção de algum modelo que produza o aprendizado efetivo das informações nele representadas.

A adoção do algoritmo *backpropagation*¹ para capturar as informações relevantes na série de cotações das ações e permitir a realização de previsões, implica na avaliação detalhada desses fatos.

6.2.1. Arquitetura do Modelo

A arquitetura de uma rede neural artificial é definida pelo agrupamento de suas unidades de processamento, ou seja, pela disposição dos neurônios artificiais e pelo modo de estabelecimento das conexões entre eles.

Para que o algoritmo de aprendizado *backpropagation* possa ser aplicado é necessário que a arquitetura do modelo de redes neurais seja estruturada em múltiplas camadas *fully connected feed forward* [FREEMAN1992] [JONES1987] [REFENES1993b]. O desafio está na identificação do número apropriado de camadas para que uma solução satisfatória seja alcançada, assim como do número de unidades de processamento de cada camada do modelo conexionista.

Tipicamente, a rede treinada pelo *backpropagation* apresenta uma camada de entrada, uma camada de saída e ao menos uma camada oculta. Não há limitação quanto

¹ Descrito na secção 4.2.8.

ao número de camadas ocultas; entretanto, estudos demonstram que três camadas ocultas, no máximo, são suficientes para que a rede consiga capturar a não-linearidade presente nos padrões de treinamento [BEALE1990].

A correta especificação do número de unidades de processamento em cada camada da rede neural também é relevante para o sucesso da aprendizagem.

As unidades visíveis são dependentes da aplicação que se quer modelar e, portanto, mais simples de serem dimensionadas [FREEMAN1992].

Como a apresentação das cotações da ação de uma série específica é feita às unidades de entrada, o número destas unidades será determinado em função da quantidade de cotações que serão apresentadas simultaneamente à rede neural durante uma iteração do algoritmo *backpropagation*. Admitindo que os preços diários da ação observados ao longo de uma semana qualquer são suficientes para prever o valor futuro da ação, o número de neurônios para a camada de entrada ficou fixado em cinco, em analogia ao número de dias úteis da semana. Entretanto, não existe relação pré-determinada da posição ocupada por um neurônio da camada de entrada com algum dia em particular, de sorte que cada unidade de processamento pode representar qualquer dia da semana. Naturalmente, as observações são apresentadas em seqüência cronológica de suas realizações.

Visto que a saída da rede neural é justamente a previsão de algum valor futuro da série das ações, a camada de saída do modelo é estruturada com apenas uma unidade de processamento, representando esta previsão.

Já o tamanho das camadas ocultas deve ser projetado mantendo o compromisso com a convergência para uma solução, sem, no entanto, prejudicar a habilidade de generalização da rede neural. [RUMELHART1986] afirma que a rede que melhor generaliza os dados é a menor rede apta a realizar o treinamento dos dados; todavia, não orienta quanto ao dimensionamento ideal de um modelo cognitivo. [WHITE1988] enfatiza que a modelagem das camadas ocultas deve considerar o menor número possível de unidades de processamento, de modo a evitar a memorização da seqüência treinada.

Assim, como a especificação do número de unidades de processamento das camadas ocultas não é um processo determinístico, a preferência de um modelo a outro será dada pela confrontação dos resultados obtidos nas diferentes configurações analisadas.

Supondo que uma rede neural tenha sido modelada contendo apenas uma camada oculta com três neurônios artificiais, o treinamento dessa rede na busca da previsão do preço p_{t+1} de uma ação no tempo $t+1$ seria ilustrado pela figura 6.1.

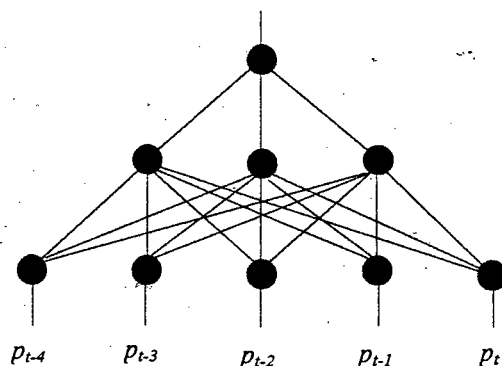


Figura 6.1. Rede neural de múltiplas camadas *fully connected* com 5, 3 e 1 neurônios artificiais em cada camada, respectivamente. (Notação : 5 - 3 - 1)

A topologia da rede neural artificial é muito importante na determinação das características cognitivas da rede e de sua habilidade em generalizar sobre o domínio do conhecimento representado [REFENES1993a] que, neste caso, é caracterizado pelos preços de fechamento da ação de alguma empresa.

6.2.2. Parâmetros do Modelo

A necessidade de obtenção de estabilidade estatística dos resultados exige o ajuste apurado dos parâmetros que influenciam o desempenho do modelo conexionista proposto. A idéia é, então, identificar intervalos de valores para esses parâmetros que possibilitem alcançar previsões estatisticamente estáveis e demonstrar que essa situação persiste para diferentes conjuntos de treinamento e teste.

6.2.2.1. Função de Ativação

As redes de múltiplas camadas somente oferecem vantagens sobre as de uma única camada se existir uma função não-linear entre suas camadas. Assim, a escolha da função de ativação adequada é fundamental para que a solução convirja rapidamente e se possa ter êxito na generalização do conhecimento modelado [REFENES1993c].

De acordo com [REFENES1993a], as funções sigmóides são utilizadas na maioria das aplicações de redes neurais artificiais. A frequência de utilização desse tipo de função é devida ao desempenho satisfatório observado na aprendizagem da rede, por serem funções mais complexas e diferenciáveis.

Refenes e Alipi [REFENES1991] provaram que as funções sigmóides simétricas são capazes de superar a velocidade de convergência de funções sigmóides assimétricas em até dez vezes. A partir desta constatação, a função tangente hiperbólica (figura 6.2) foi selecionada como função de ativação das redes neurais modeladas para realizar a previsão do mercado acionário.

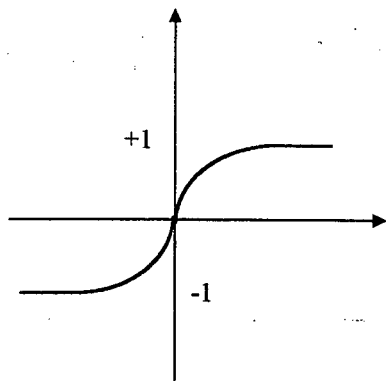


Figura 6.2. Função tangente hiperbólica (tanH).

Como esta função fornece resultados no intervalo $[-1,+1]$, os dados usados durante o aprendizado do algoritmo *backpropagation* foram assim escalonados; desta forma, a aplicação de $\tan H(p_{\text{máx}}) = +1$ e $\tan H(p_{\text{mín}}) = -1$, onde $p_{\text{máx}}$ e $p_{\text{mín}}$ correspondem aos valores máximo e mínimo observados entre os preços das ações da série analisada.

6.2.2.2. Taxa de Aprendizado

Outro parâmetro que interfere na convergência para uma solução é a taxa de aprendizado ou coeficiente de plasticidade sináptica. Sua influência está relacionada à mudanças nos pesos entre as conexões das unidades de processamento : taxas de aprendizado muito pequenas (próximas de zero) implicam em tempo de treinamento longo; contudo, taxas de aprendizado muito grandes (próximas de 1.0) podem provocar oscilações em torno da solução [REFENES1993a].

De acordo com [BEALE1990], caso a taxa pela qual os pesos são atualizados seja decrementada progressivamente, o *gradiente decrescente* está habilitado a encontrar uma melhor solução.

Desta forma, procurando conciliar tempo de treinamento e convergência do resultado da aprendizagem, a taxa de aprendizado adotada nesta modelagem assume valor maior no início do processo de ajuste dos pesos, decrescendo exponencialmente na ordem de 0.50 à medida que o treinamento evolui. Ainda, para cada camada do modelo conexionista há um valor específico para a taxa de aprendizado, iniciando em 0.90 para a primeira camada oculta, 0.70 para a segunda, 0.50 para a terceira e 0.15 para a camada de saída, caso a topologia do modelo apresente esta estrutura.

6.2.2.3. Momentum

O termo *momentum*², como alternativa para diminuir a possibilidade de oscilações durante a execução do algoritmo *backpropagation*, não permaneceu constante para o conjunto de redes neurais artificiais avaliadas, assumindo valores no intervalo [0.2,0.7].

² Conforme seção 4.2.9.

6.2.2.4. Época

A velocidade de convergência do aprendizado também é afetada pela frequência de atualização dos pesos que conectam as unidades de processamento entre diversas camadas da rede neural [REFENES1993b]. Ao invés de se efetuar esta atualização após cada iteração do algoritmo de aprendizado, os desvios observados entre a previsão obtida e a desejada podem ser acumulados durante a apresentação sucessiva de um certo número de padrões de treinamento. Somente com o decorrer de uma *época*, ou seja, quando todos aqueles padrões tiverem sido apresentados às unidades visíveis do modelo conexionista, é que a atualização dos pesos da camada de saída e das camadas ocultas é processada.

Visto que as séries dos preços das ações normalmente apresentam comportamento bastante irregular, os modelos de previsão propostos foram treinados com a versão estocástica [REFENES1993b] [REFENES1993c] do *backpropagation*, onde os pesos são atualizados a cada iteração deste algoritmo.

6.2.3. Aprendizagem Móvel

Para realizar com sucesso a previsão de séries temporais econômicas, tais como o preço de ações, nas quais as regras de previsão estão sendo alteradas continuamente, o aprendizado da rede neural precisa ser adaptado de modo a acompanhar essas mudanças.

Adotando, na modelagem de redes neurais artificiais, a técnica de *janelamento* proposta por [REFENES1993a] para a identificação de regularidades em séries temporais, $J^I \rightarrow J^O$ pode ser usado como o vetor ou padrão de treinamento. Dessa forma, se estabelece um relacionamento entre a seqüência de preços contidas em J^I com aquela contida em J^O .

Seguindo o mesmo raciocínio desenvolvido para o dimensionamento das unidades visíveis do modelo conexionista, os tamanhos n e m das *janelas* J^I e J^O ,

respectivamente, foram condicionados à suposição de que o preço da ação em determinado dia está relacionado com os observados nos cinco dias precedentes. Esta hipótese pode ser assumida ao se considerar que não ocorrem alterações drásticas entre observações consecutivas de uma série histórica de cotações de ações. Desta forma, fixou-se $n = 5$ e $m = 1$.

Quanto ao tamanho do passo de deslocamento, não se estabeleceu um valor específico para o treinamento da rede neural. A detecção de regularidades não-lineares em uma série de cotações de ações foi comprovada pelo treinamento aleatório ao longo de toda a série temporal. Genericamente, a apresentação das observações para treinamento é dada por

$$J^I = \{p_{t-5}, \dots, p_{t-1}\} \rightarrow J^O = \{p_t\}$$

onde p_t é o preço da ação no tempo t .

Para a etapa de teste da rede proposta, o tamanho do passo de deslocamento foi fixado em 1. Desta forma, à medida que uma nova observação se torna disponível, a previsão para o período seguinte pode ser realizada. Mais formalmente, tem-se

$$J^I = \{p_{t-4}, \dots, p_t\} \rightarrow J^O = \{\hat{p}_{t+1}\}$$

onde \hat{p}_{t+1} é o preço previsto para a ação no tempo $t+1$.

6.2.4. Dados para Aprendizagem

O ajuste sucessivo dos pesos entre as conexões dos neurônios artificiais é alcançado pela exposição dos preços das ações às unidades de entrada e saída da rede. Concluída a etapa de treinamento, a capacidade da rede neural proposta em capturar as regularidades presentes nos dados é avaliada. Assim, para que as previsões dos preços

futuros possam ser realizadas, os preços das ações devem ser novamente apresentados à camada de entrada do modelo conexionista.

Para tanto, foram coletados os preços de fechamento diários de um conjunto de ações ordinárias nominativas e preferenciais nominativas de algumas empresas nacionais entre janeiro de 1986 a junho de 1990.

Pelo fato de ser comum em séries econômicas a existência de tendências [MORETTIN1981], os dados utilizados na aprendizagem das redes neurais foram pré-processados através da aplicação da função *logaritmo natural*, de modo a torná-los os mais regulares possíveis, conforme orienta [KIMOTO1990].

6.2.5. Tempo de Treinamento

O tempo de treinamento corresponde ao número de apresentações das cotações da ação de uma determinada companhia à rede neural artificial. Como o poder preditivo da rede neural pode ser observado como uma função do tempo de treinamento [REFENES1993b], um número variado de iterações foi abordado, de modo a permitir melhor compreensão dos resultados obtidos e avaliação mais confiável do desempenho do modelo de previsão.

6.2.6. Horizonte de Previsão

As previsões dos valores futuros das séries de preços das ações foram realizadas adotando como horizonte de previsão o instante de tempo imediatamente subsequente à origem. Tal procedimento está fundamentado em [SOUZA1989], que estabelece que a otimalidade de uma previsão não pode ser garantida para períodos superiores ao instante seguinte de tempo.

6.3. Análise dos Resultados

A partir dos critérios relacionados na seção anterior, as redes neurais artificiais foram modeladas e o aprendizado efetuado. A convergência e generalização dos modelos desenvolvidos e a estabilidade dos resultados obtidos foram analisadas.

6.3.1. Convergência

A convergência de um modelo conexionista é avaliada pela capacidade deste modelo em produzir previsões acuradas para a série de preços apresentada durante o processo de ajuste dos pesos entre as conexões das camadas da rede neural. Muitos casos foram analisados, centenas de redes foram construídas, treinadas e testadas. Na impossibilidade de se apresentar todos estes estudos, uma série de preços de ações de uma determinada empresa (figura 6.3) foi escolhida aleatoriamente entre o universo de ações disponíveis para demonstrar a habilidade de convergência dos modelos conexionistas na previsão de séries temporais³.

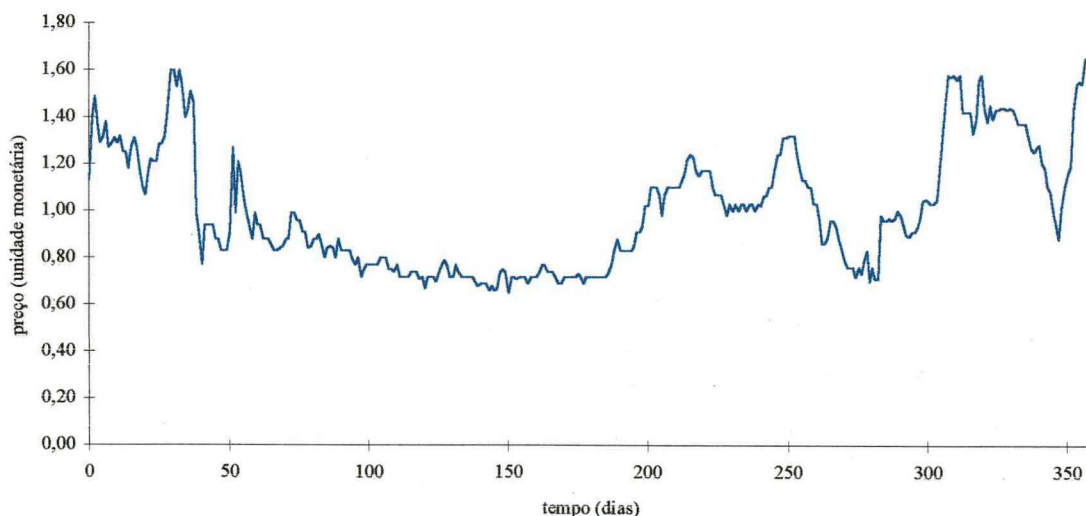


Figura 6.3. Gráfico do preço de fechamento diário de uma ação.

³ Analogamente, ações de algumas outras empresas estão apresentadas no Anexo.

A série ilustrada pelo gráfico anterior representa o preço de fechamento diário da ação selecionada⁴ para um período de tempo de 360 dias, equivalente a 1½ ano. Pode-se verificar que esta série apresenta comportamento bastante irregular, oscilando entre quedas e subidas de preço, sendo muito representativa das situações normalmente observadas no mercado acionário.

A figura 6.4 ilustra as previsões das cotações da ação obtidas com o treinamento e teste de uma rede neural artificial 5 - 10 - 5 - 1, durante 20000 iterações, para um valor de *momentum* = 0.4. Através deste gráfico, é possível constatar a habilidade do referido modelo conexionista em gerar previsões bastante acuradas. Os valores previstos são muito próximos dos realmente observados, tanto em termos de tendência quanto em valores absolutos.

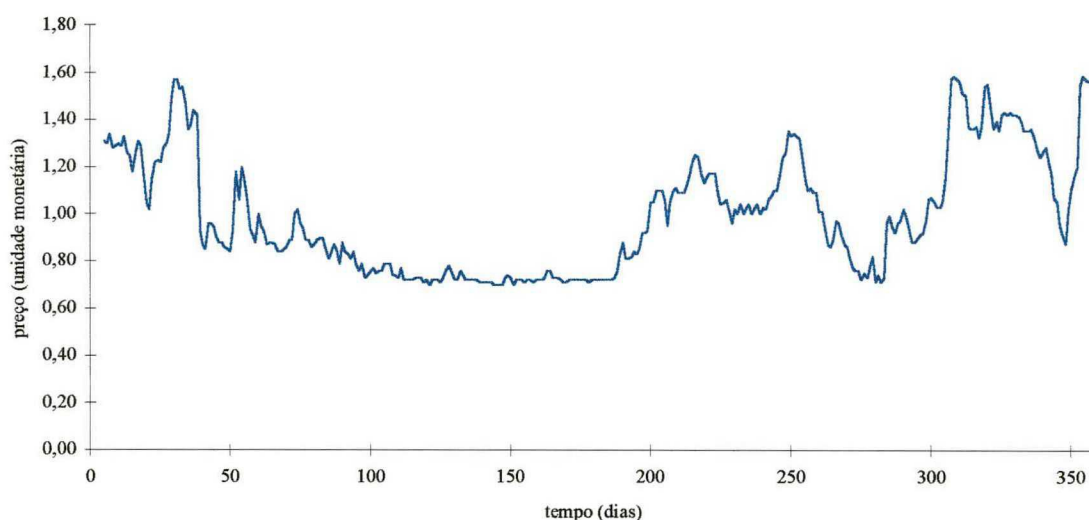


Figura 6.4. Gráfico dos preços previstos para a ação pelo modelo de redes neurais, considerando a propriedade de convergência.

A figura 6.5 permite melhor visualização da convergência das previsões, sobrepondo em um único gráfico, os valores reais e os obtidos no aprendizado da rede neural artificial.

⁴ ITAU PN do período de 02/01/86 a 29/06/87.

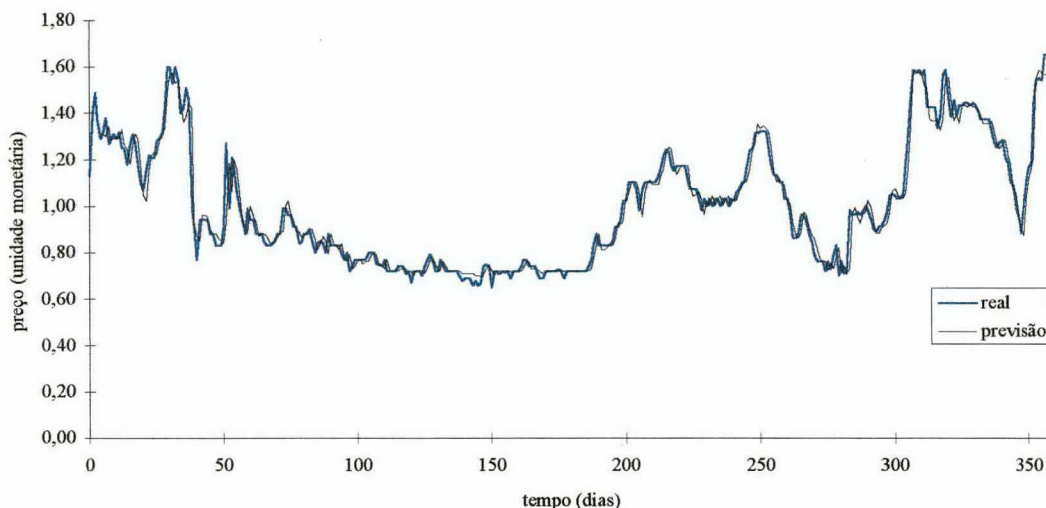


Figura 6.5. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência.

A convergência do modelo conexionista também pode ser avaliada através da análise de um gráfico de dispersão entre as observações e as previsões. A curva ideal para um gráfico de dispersão é uma reta com 45° em relação ao eixo das abscissas, passando pela origem. Isto porque, se o valor observado é, por exemplo, 1,20, o ideal seria que o previsto também fosse 1,20; caso fosse observado o valor 0,80, a previsão para este valor deveria ser a mesma, e assim sucessivamente. Os pontos (1,20;1,20), (0,80;0,80) definem a curva descrita acima. Analisando o gráfico de dispersão representado pela figura 6.6 é possível perceber que o comportamento daquelas dispersões é praticamente aquele considerado o ideal.

Outra possibilidade de verificar se as previsões realizadas convergiram para um intervalo de erro estabelecido como aceitável é traçar o gráfico de distribuição de frequência para estes erros. Calculando o erro percentual absoluto⁵ das previsões obtidas em relação às observações da série de preços de ações e distribuindo estes valores no gráfico ilustrado pela figura 6.7, pode-se inferir a respeito da acuidade das previsões. Pela análise deste gráfico, percebe-se que a maioria dos erros ficou distribuída entre zero e 3%; mais precisamente, 56% dos erros apurados. Apenas 6% dos desvios apresentaram valores superiores ou iguais a 10%.

⁵ Conforme equação (3.18).

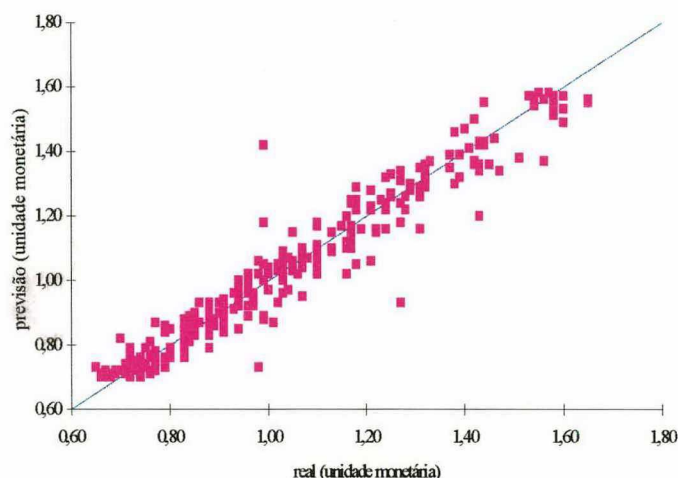


Figura 6.6. Gráfico de dispersão entre o preço real da ação e a previsão deste preço pelo modelo conexionista, considerando a propriedade de convergência.

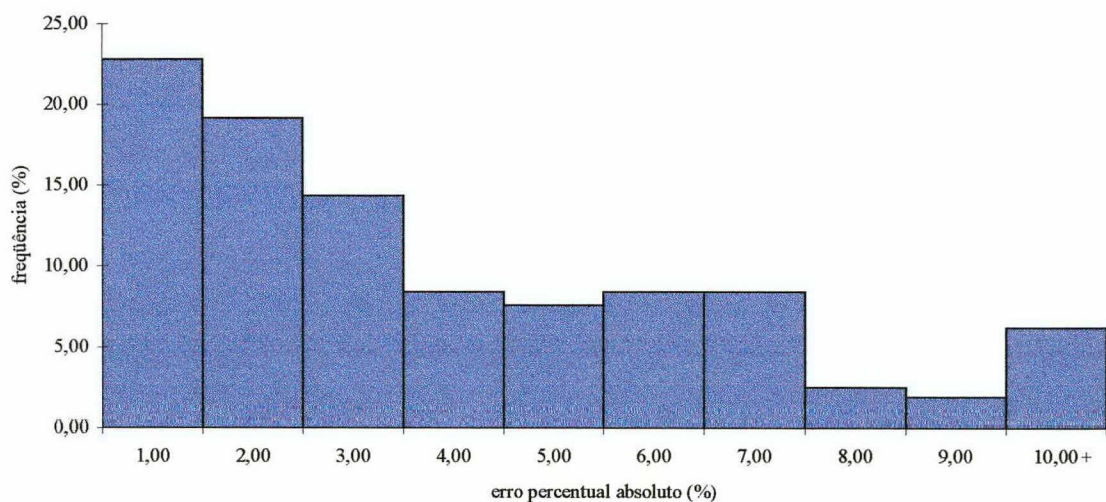


Figura 6.7. Distribuição de frequência dos erros da previsão da rede neural, considerando a propriedade de convergência.

Os resultados obtidos pela aplicação de técnicas de redes neurais artificiais foram comparados com aqueles gerados pela utilização de dois métodos de previsão convencionais frequentemente utilizados : Média Móvel e Alisamento Exponencial Simples. O horizonte de previsão para o método de Médias Móveis foi fixado em 5, seguindo o mesmo princípio adotado na modelagem de redes neurais, onde a cotação da ação nos últimos 5 dias é usada para prever o preço da ação no próximo período. O

valor da constante α^6 do método de Alisamento Exponencial Simples foi estimado em 0,60, a partir da análise dos resultados obtidos pela aplicação deste método na série de preços considerada.

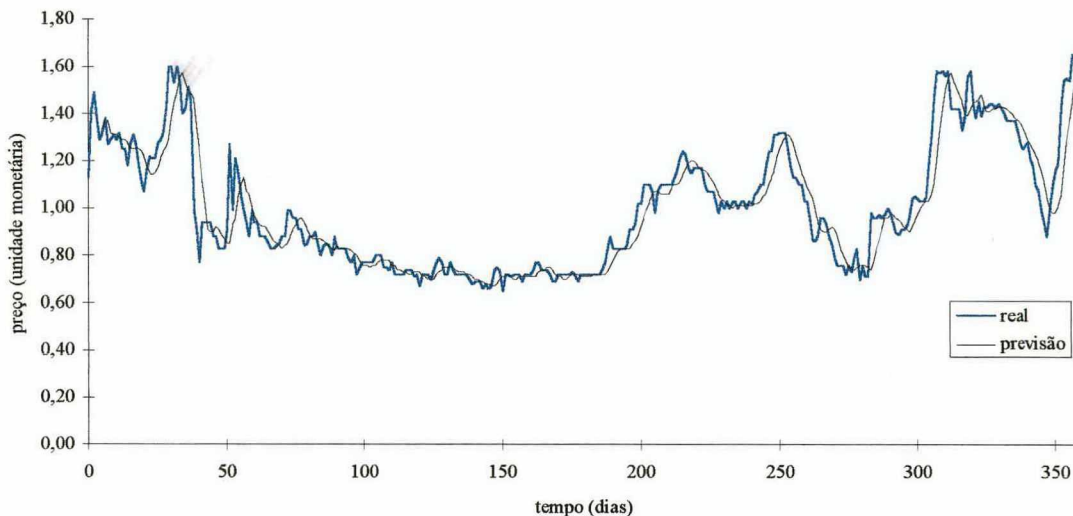


Figura 6.8. Gráfico comparativo entre os preços reais e os previstos pela Média Móvel, considerando a propriedade de convergência.

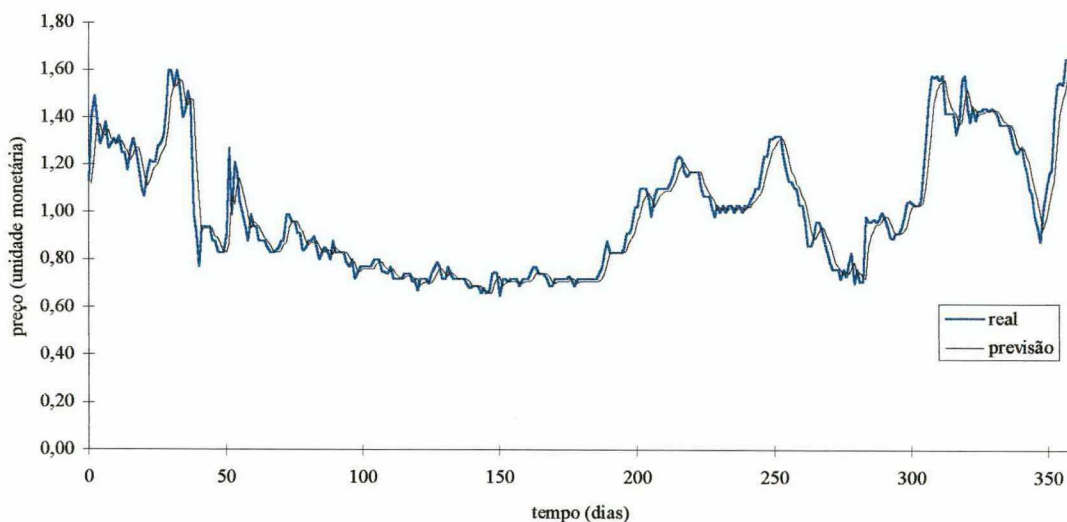


Figura 6.9. Gráfico comparativo entre preços reais e os previstos pelo Alisamento Exponencial Simples, considerando a propriedade de convergência.

⁶ Conforme equação (3.3).

As figuras 6.8 e 6.9 representam, respectivamente, as previsões geradas pela aplicação da Média Móvel e Alisamento Exponencial Simples. Comparando estes gráficos com o ilustrado pela figura 6.5, verifica-se que os modelos conexionistas geraram previsões mais acuradas que as dos referidos métodos convencionais de previsão. Esta constatação é reforçada pela análise da distribuição de frequência dos erros percentuais absolutos para aqueles métodos, onde percebe-se pelos gráficos das figuras 6.10 e 6.11 a grande concentração dos desvios no intervalo de erro maior.

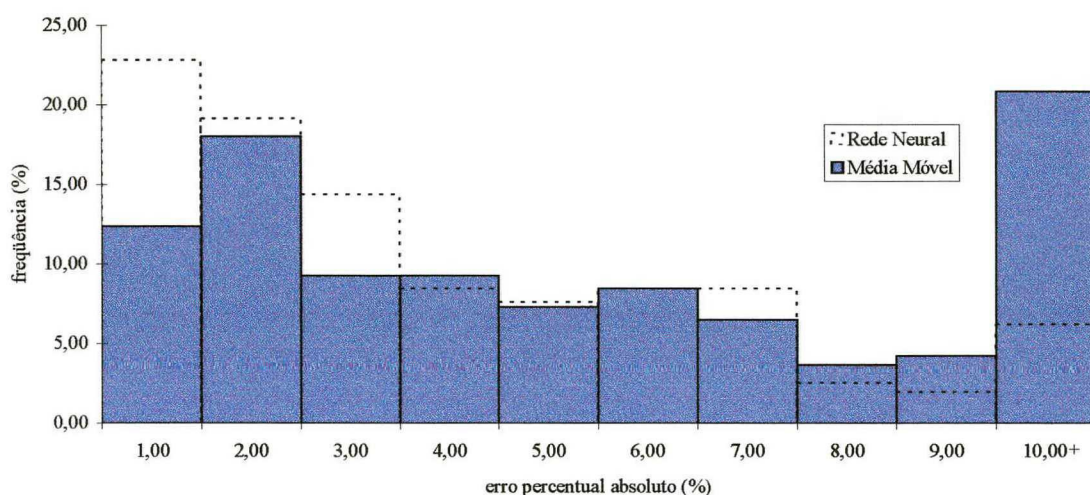


Figura 6.10. Distribuição de frequência dos erros da previsão da Média Móvel, considerando a propriedade de convergência.

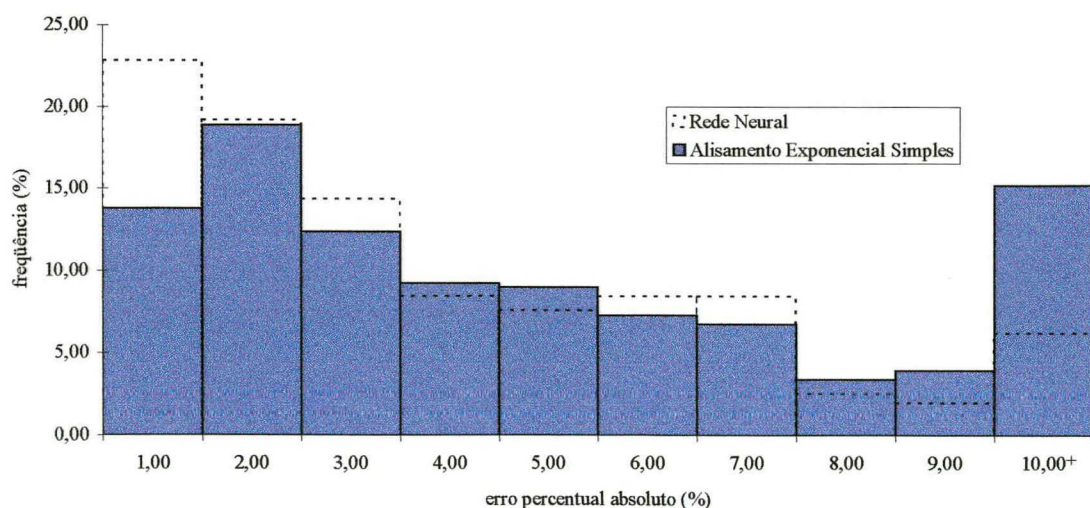


Figura 6.11. Distribuição de frequência dos erros da previsão do Alisamento Exponencial Simples, considerando a propriedade de convergência.

6.3.2. Generalização

A propriedade de generalização de um modelo conexionista é avaliada pela habilidade deste modelo em capturar as regularidades presentes no conjunto de treinamento e extrapolar esse conhecimento para uma nova situação. Através da capacidade da rede neural artificial em abstrair as informações relevantes da série de preços das ações apresentadas durante o ajuste dos pesos entre as conexões da rede neural, é possível realizar a previsão da cotação futura de determinada ação. Desta forma, os preços subsequêntes àqueles usados para o treinamento da rede neural artificial descrita na seção anterior foram empregados para avaliar a habilidade do modelo conexionista em generalizar sobre o domínio do problema ao qual foi submetido durante o processo de ajuste dos pesos sinápticos.

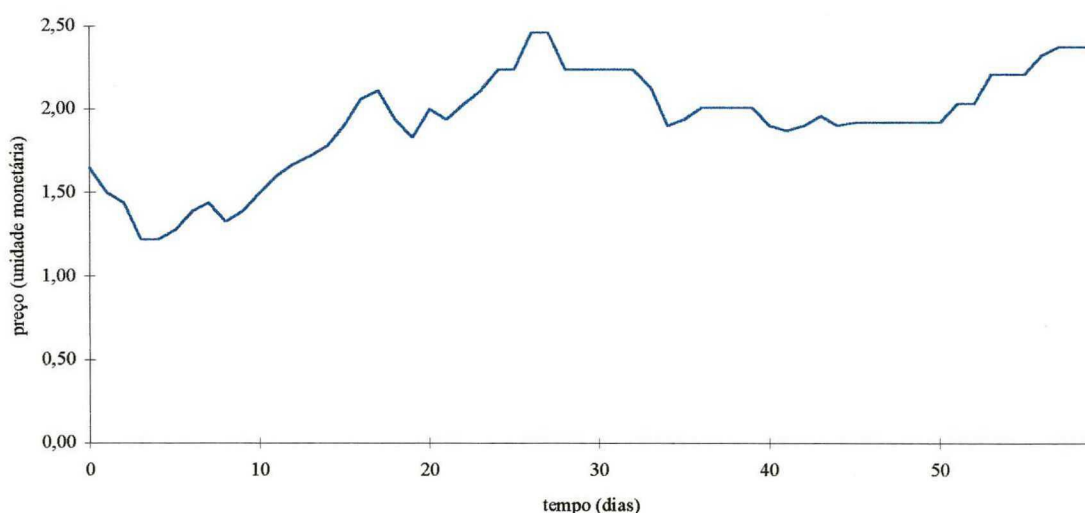


Figura 6.12. Gráfico do preço de fechamento diário de uma ação, referente ao período de tempo imediatamente subsequente ao treinamento.

A figura 6.12 ilustra os preços diários de fechamento da série histórica selecionada, referentes ao período de tempo de três meses após o treinamento da rede neural. Apresentando, sucessivamente, estas observações à rede treinada, é possível, então, realizar a previsão do preço futuro da ação (figura 6.13).

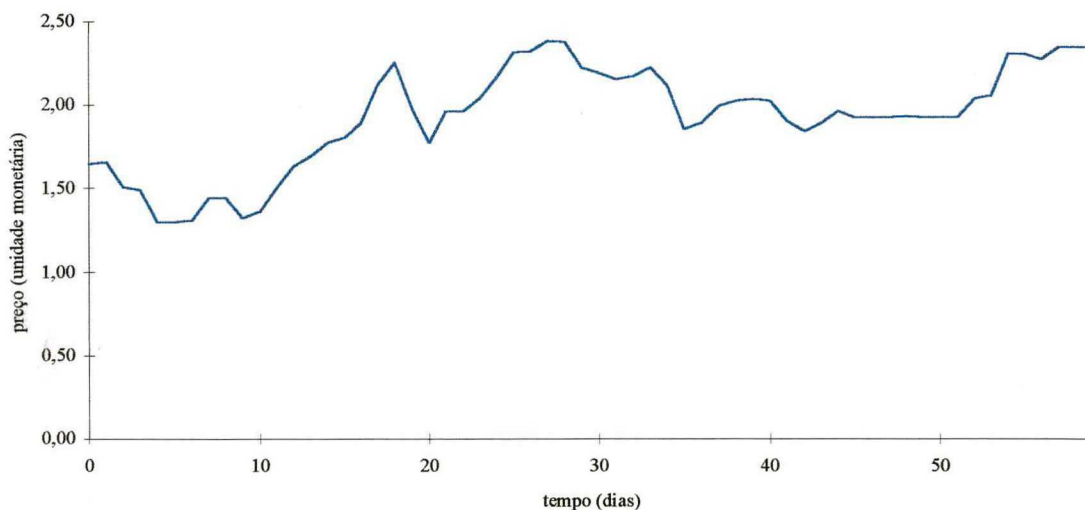


Figura 6.13. Gráfico dos preços previstos para a ação pelo modelo de redes neurais, considerando a propriedade de generalização.

O gráfico anterior e aquele representado pela figura 6.14 demonstram a capacidade da rede neural artificial em capturar as regularidades não-lineares presentes no conjunto de treinamento, fornecendo previsões acuradas para o futuro próximo.

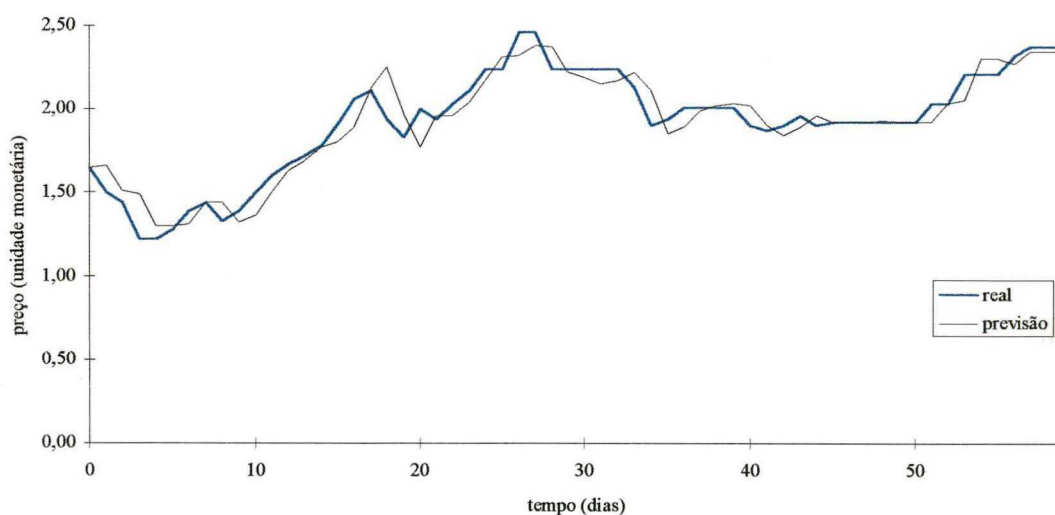


Figura 6.14. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização.

Identicamente, o gráfico de dispersão (figura 6.15) ilustra a capacidade de ajustamento do modelo de redes neurais artificiais na previsão dos preços da ação analisada.

É importante observar que a rede neural apresentou um desempenho relativamente satisfatório durante o período de tempo decorrido entre o seu treinamento e a realização de previsões, sem que houvesse necessidade de ajuste dos pesos entre as conexões do modelo à medida que novas observações se tornaram disponíveis.

A distribuição de frequência dos erros (figura 6.16) das previsões obtidas pela aplicação de técnicas de redes neurais traduz a habilidade destes modelos em capturar as regularidades contidas no conjunto de treinamento e extrapolar essas características para uma nova situação.

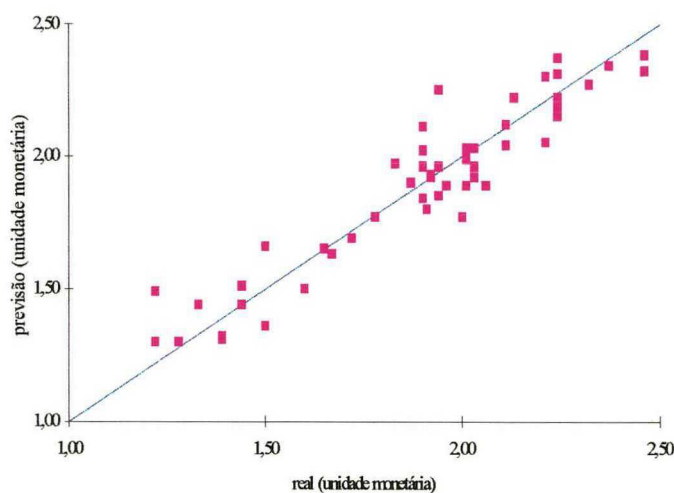


Figura 6.15. Gráfico de dispersão entre o preço real da ação e a previsão deste preço pelo modelo conexionista, considerando a propriedade de generalização.

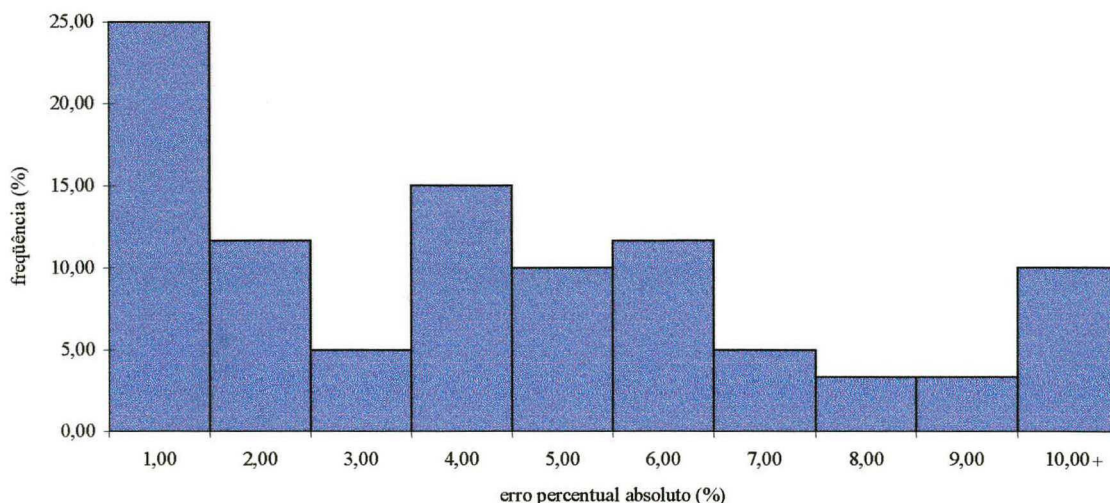


Figura 6.16. Distribuição de frequência dos erros da previsão da rede neural, considerando a propriedade de generalização.

Analogamente à convergência, a propriedade de generalização dos modelos conexionistas também pode ser evidenciada pela comparação com as previsões geradas pela aplicação da Média Móvel (figura 6.17) e Alisamento Exponencial Simples (figura 6.18). Analisando os gráficos, constata-se maior acuidade dos valores previstos pela rede neural em relação a estes métodos convencionais de previsão de séries temporais. Observa-se também a dificuldade destes modelos em reagir rapidamente às mudanças verificadas no comportamento da série de preços da ação avaliada.

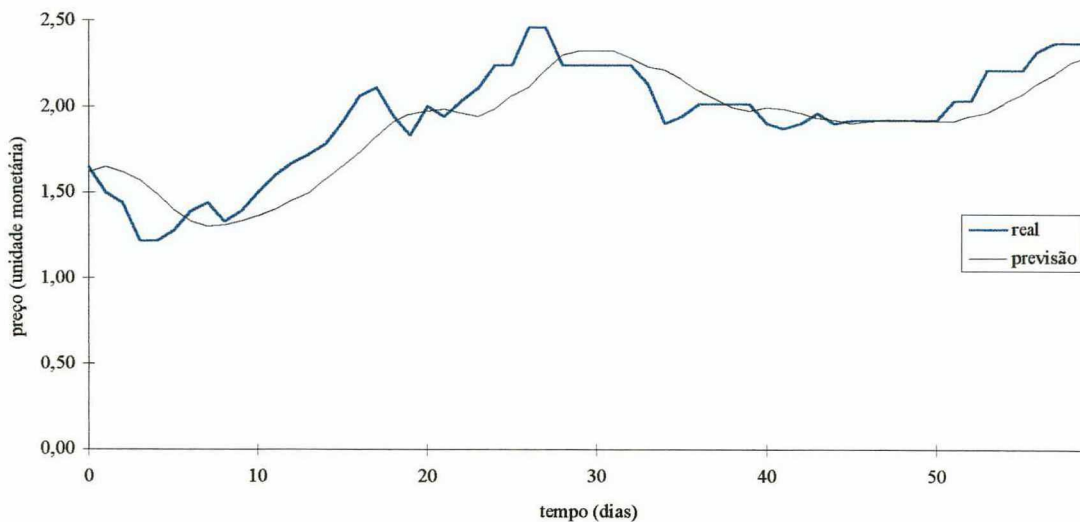


Figura 6.17. Gráfico comparativo entre os preços reais e os previstos pela Média Móvel, considerando a propriedade de generalização.

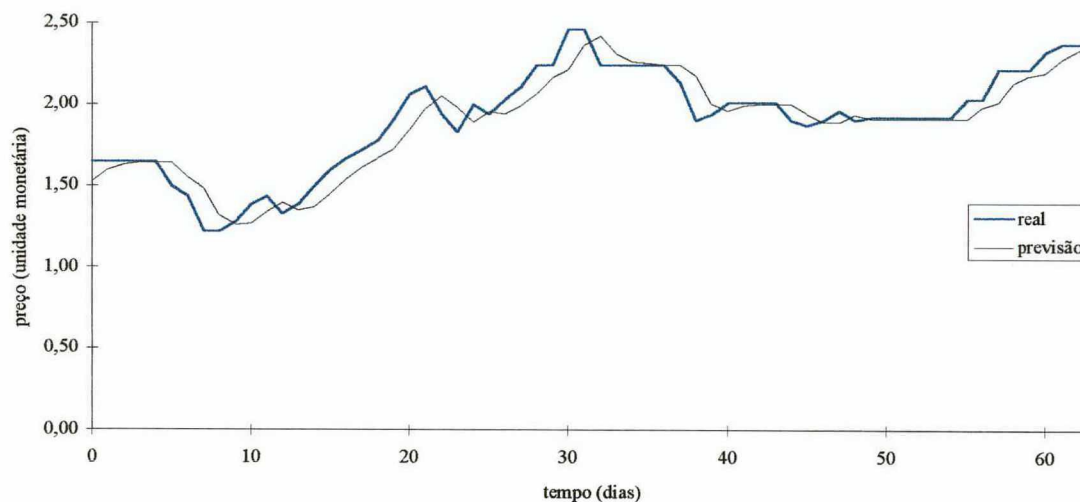


Figura 6.18. Gráfico comparativo entre os preços reais e os previstos pelo Alisamento Exponencial Simples, considerando a propriedade de generalização.

Analisando os gráficos de distribuição de frequência dos erros das previsões da Média Móvel e do Alisamento Exponencial Simples, figuras 6.19 e 6.20, respectivamente, verifica-se novamente muitos desvios concentrados no intervalo de erro máximo, em oposição ao observado na distribuição de frequência dos erros das previsões obtidas pela aplicação do modelo de redes neurais artificiais.

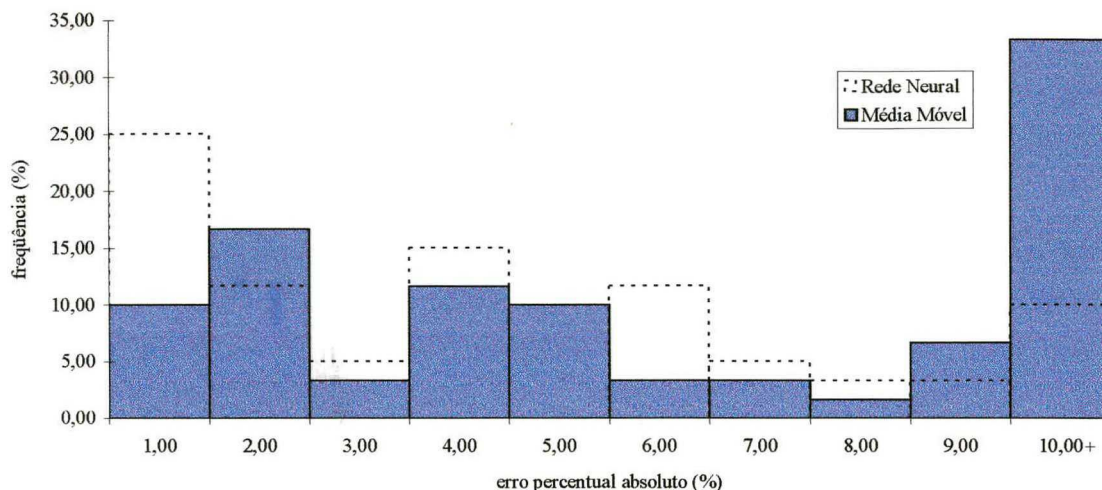


Figura 6.19. Distribuição de frequência dos erros da previsão da Média Móvel, considerando a propriedade de generalização.

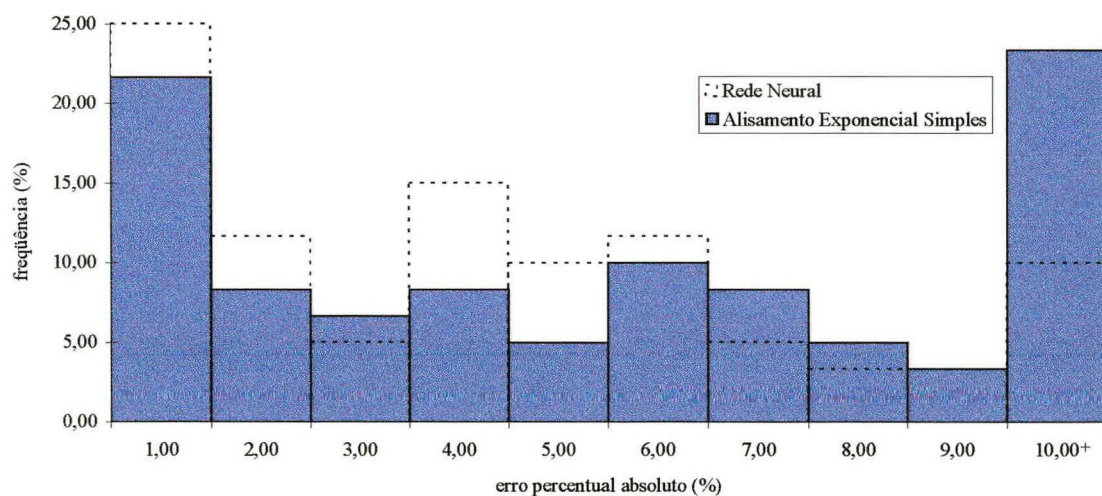


Figura 6.20. Distribuição de frequência dos erros da previsão do Alisamento Exponencial Simples, considerando a propriedade de generalização.

6.3.3. Estabilidade

A estabilidade dos resultados obtidos pela rede neural foi analisada através da avaliação de fatores que influenciam no desempenho de um modelo conexionista, como a sua arquitetura, o tempo dispendido para treinamento e as características de diferentes

séries treinadas, os termos do *gradiente decrescente*⁷ e as condições iniciais do aprendizado.

6.3.3.1. Quanto à Arquitetura do Modelo Conexionista

Na tabela 6.1 está relacionado o erro percentual absoluto médio⁸ das previsões realizadas pela aplicação de técnicas de redes neurais artificiais para diversas topologias destas redes, variando o número de camadas ocultas e também o número de unidades de processamento nestas camadas. Os padrões foram apresentados de 2000 a 80000 vezes à rede neural, para um valor fixo de *momentum* = 0.4. Como esperado, à medida que se observou a evolução do treinamento, o erro entre o valor previsto e o observado diminuiu, independente da topologia empregada. Entretanto, conforme a arquitetura proposta, a convergência para uma solução aceitável foi alcançada com diferente número de iterações para o algoritmo *backpropagation*.

Iteração	Topologia			
	5 - 5 - 1	5 - 10 - 5 - 1	5 - 20 - 10 - 1	5 - 15 - 10 - 5 - 1
2000	11.02	7.75	7.00	11.23
5000	4.62	7.70	8.41	15.08
10000	3.96	4.59	4.64	5.60
15000	3.99	4.06	4.88	4.92
20000	3.64	3.93	4.17	3.84
40000	3.66	3.65	3.83	3.66
80000	3.71	3.61	3.82	3.64

Tabela 6.1. Erro percentual absoluto médio das previsões de padrões **usados** no treinamento da rede neural.

⁷ Descrito na seção 4.2.8.

⁸ Conforme equação (3.20).

Os valores representados na tabela 6.2 representam os desvios entre as previsões e os preços observados na série subsequente àquela utilizada no ajuste dos pesos da rede neural artificial.

Iteração	Topologia			
	5 - 5 - 1	5 - 10 - 5 - 1	5 - 20 - 10 - 1	5 - 15 - 10 - 5 - 1
2000	10.01	6.50	6.05	9.37
5000	4.92	6.41	5.17	12.73
10000	4.90	4.75	4.48	5.38
15000	4.59	4.44	5.12	4.87
20000	4.03	4.42	4.44	4.32
40000	4.18	4.20	4.33	4.40
80000	4.26	4.12	4.26	4.33

Tabela 6.2. Erro percentual absoluto médio das previsões de padrões **não usados** no treinamento da rede neural.

Percebe-se, pela comparação dos erros, que o tamanho da camada escondida pode variar em um intervalo relativamente grande sem degradação significativa do desempenho da generalização, sobretudo quando o número de iterações é maior.

6.3.3.2. Quanto ao Tempo de Treinamento

Analisando as tabelas 6.1 e 6.2, verifica-se que a magnitude dos erros decresceu à medida que aumentou o número de iterações do treinamento dos modelos conexionistas. No entanto, é possível detectar a existência de regiões onde se observa um desempenho temporário pior. Acredita-se que esta situação não possa ser interpretada como super treinamento, visto que elas ocorreram com um número de iterações relativamente pequeno - entre 5000 e 15000 iterações; provavelmente a sua existência implica em a rede estar ainda subtreinada e que soluções melhores sejam alcançadas com um tempo de

treinamento maior. A partir de 20000 iterações, os desvios sofreram alterações muito pequenas, indicando que o treinamento exaustivo do modelo conexionista não implica, necessariamente, em previsões mais acuradas.

A figura 6.21 ilustra os erros das previsões para cada observação do conjunto de padrões usados na avaliação da generalização de uma rede neural treinada durante 2000 e 80000 iterações.

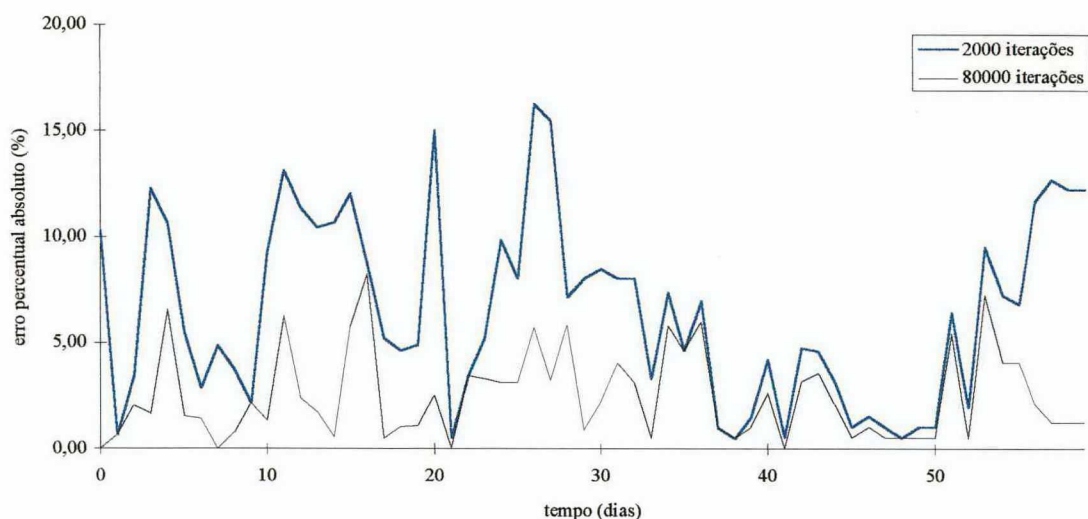


Figura 6.21. Gráfico do erro das previsões em relação ao número de iterações do treinamento.

6.3.3.3. Quanto aos Diferentes Conjuntos de Treinamento

Através da análise da figura 6.22, observa-se a capacidade das redes neurais artificiais em se adaptar às diferentes situações, fornecendo previsões igualmente acuradas.

Comparando o erro percentual absoluto médio das previsões realizadas para os 60 e 120 dias imediatamente subsequentes àqueles empregados no treinamento da rede descrita na seção 6.3.1, percebe-se que existe diferença na magnitude dos desvios para cada um destes conjuntos de preços; entretanto, com a evolução do aprendizado,

independentemente das características dos dados empregados, os erros tendem a assumir valores menores, estabilizando em determinado nível.

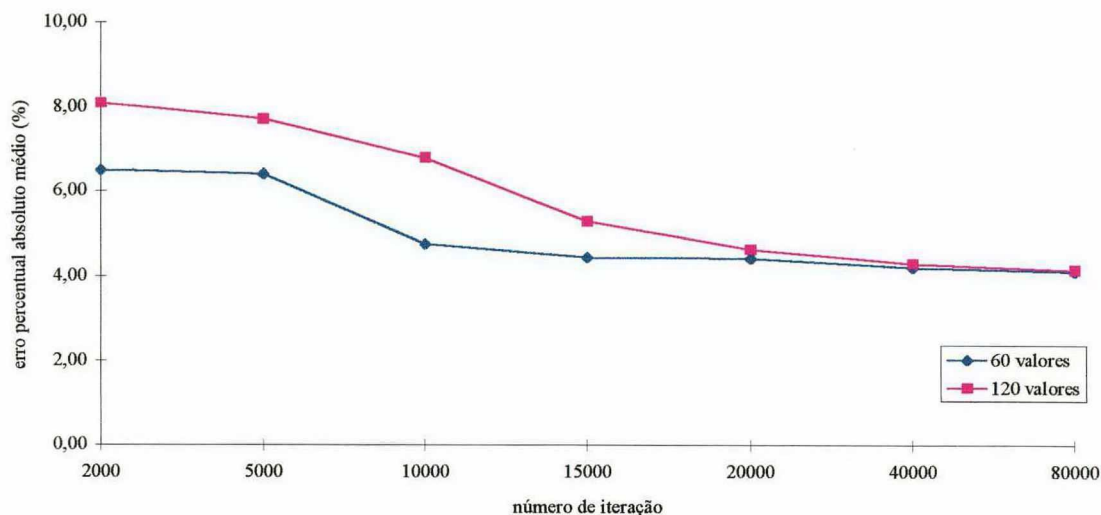


Figura 6.22. Gráfico do erro das previsões para diferentes conjuntos de treinamento.

Ainda, as tabelas 6.3 e 6.4 representam os intervalos de confiança⁹ de 95% para o erro percentual absoluto médio das previsões obtidas a partir dos preços de diversas ações **apresentados** e **não apresentados**, respectivamente, durante o treinamento da rede neural artificial.

Erro Percentual Absoluto Médio					
Rede Neural		Média Móvel		Alisamento Exponencial Simples	
limite inferior	limite superior	limite inferior	limite superior	limite inferior	limite superior
2.72	3.99	7.42	9.32	5.51	7.10
2.88	4.13	6.22	8.30	4.33	5.93
3.18	4.04		6.70	4.04	5.15
3.70	4.65	5.32 5.27	7.50	4.19	4.75
4.19	4.75	5.61	6.33	5.26	5.85

Tabela 6.3. Intervalo de confiança para o erro percentual absoluto médio das previsões de cinco conjuntos de padrões **usados** no treinamento da rede neural.

⁹ Conforme [HOFFMANN1991].

Erro Percentual Absoluto Médio					
Rede Neural		Média Móvel		Alisamento Exponencial Simples	
limite inferior	limite superior	limite inferior	limite superior	limite inferior	limite superior
1.41	4.98	3.40	7.58	2.23	5.67
2.69	3.71	5.16	6.83	3.71	5.04
3.07	5.17	5.36	8.30	3.93	6.09
3.13	4.32	5.06	6.90	3.76	5.13
4.53	14.20	9.11	22.12	6.06	15.97

Tabela 6.4. Intervalo de confiança para o erro percentual absoluto médio das previsões de cinco conjuntos de padrões **não usados** no treinamento da rede neural.

Apesar dos intervalos de erro se sobreporem em alguns casos, observa-se claramente que os limites inferiores e superiores (para todos os conjuntos) obtidos pelas redes neurais artificiais apresentam menor magnitude que aqueles obtidos com a aplicação dos métodos de Média Móvel e Alisamento Exponencial Simples, o que qualifica positivamente o emprego de técnicas de redes neurais artificiais na previsão.

6.3.3.4. Quanto aos Termos do *Gradiente Decrescente*

Outro ponto a ser analisado é a estabilidade dos termos do *gradiente decrescente* dos modelos conexionistas aplicados na previsão do mercado acionário. Como a taxa de aprendizado foi fixada para cada camada das redes neurais, é suficiente verificar a influência dos possíveis valores a serem assumidos pelo *momentum* no desempenho da rede neural artificial. Na tabela 6.5 estão representados os erros percentuais absolutos médios das previsões dos valores futuros da série analisada para diferentes números de iterações do algoritmo *backpropagation*. Avaliando essas medidas, a afirmação de que exista algum valor de *momentum* que forneça previsões mais acuradas que outro não pode ser feita, sobretudo para períodos de treinamento mais longos. Entretanto,

adotando valores de *momentum* no intervalo [0.2,0.4] é observável que a convergência para uma solução aceitável é alcançada com menor número de iterações.

Iteração	<i>Momentum</i>					
	0.2	0.3	0.4	0.5	0.6	0.7
2000	5.01	5.32	5.07	8.02	8.58	9.50
5000	5.26	5.33	4.44	6.81	5.40	8.68
10000	4.31	3.86	4.50	3.81	4.22	5.01
15000	3.82	3.91	4.44	5.08	3.79	5.01
20000	3.66	3.86	4.89	3.89	3.79	3.67
40000	3.67	3.66	3.68	3.67	3.66	3.60
80000	3.68	3.69	3.67	3.65	3.61	3.62

Tabela 6.5. Erro percentual absoluto médio das previsões de padrões usados no treinamento da rede neural para diferentes valores do termo *momentum*.

As mesmas considerações podem ser feitas para a tabela 6.6, que representa os erros das previsões de padrões não apresentados durante o treinamento da rede neural.

Iteração	<i>Momentum</i>					
	0.2	0.3	0.4	0.5	0.6	0.7
2000	4.55	4.65	5.00	6.08	7.39	6.54
5000	4.65	4.37	5.16	8.41	5.70	7.90
10000	4.11	4.14	4.33	4.22	4.76	4.90
15000	4.37	4.46	4.80	6.15	4.15	4.35
20000	4.15	4.30	4.23	4.02	3.97	4.32
40000	4.09	4.18	4.11	4.21	4.06	4.13
80000	3.96	4.16	4.12	4.03	4.05	4.27

Tabela 6.6. Erro percentual absoluto médio das previsões de padrões **não usados** no treinamento da rede neural para diferentes valores do termo *momentum*.

6.3.3.5. Quanto às Condições Iniciais

Embora o algoritmo *backpropagation* seja sensível aos valores das condições iniciais [REFENES1993c], isto é, aos valores de inicialização dos pesos das conexões entre as camadas da rede, não se detectou alterações que pudessem ser consideradas significativas ao se comparar previsões obtidas a partir de diferentes inicializações dos pesos da rede.

6.4. Aplicação do Modelo de Previsão considerando uma Carteira Simplificada de Ações

A aplicação dos modelos conexionistas na previsão do mercado acionário ainda pode ser avaliada através do desempenho obtido por esses modelos ao investir em uma carteira simplificada de ações, onde é considerado um número reduzido - mas de diversos setores produtivos - de ações ordinárias e preferenciais.

Admitindo a possibilidade de investimento em um conjunto de ações de cinco empresas (A, B, C, D e E¹⁰), o objetivo consiste em aplicar os recursos disponíveis na ação que, de acordo com a previsão da rede neural artificial, terá o maior retorno¹¹.

A tabela 6.7. apresenta os preços observados das ações selecionadas no decorrer de um determinado mês. A partir desses preços, o retorno efetivo para cada uma das ações foi calculado. Na seqüência, as redes neurais artificiais referentes a cada ação foram treinadas com a apresentação dos preços das ações anteriores ao período considerado. Após o ajuste dos pesos dos modelos conexionistas, foram realizadas as previsões das cotações das ações para o período analisado.

Em seguida, para cada dia, foram determinadas as ações sujeitas a maior retorno, conforme o que realmente foi observado e o que foi previsto pelas redes neurais. Consideradas essas informações, calculou-se o retorno acumulado máximo que seria obtido caso fossem feitos investimentos naquelas ações que realmente apresentaram maior retorno e, o retorno acumulado máximo alcançado com a aplicação dos recursos nas ações que, segundo os modelos conexionistas, maximizariam o retorno global.

Assumindo a inexistência de custos das transações, a tabela 6.8. apresenta os resultados que seriam obtidos no decorrer do período por um investidor dispondo de 1000,00 unidades monetárias a serem aplicadas no mercado de ações. É importante observar que na estratégia de investimento adotada não foi considerada a possibilidade de aplicação dos recursos em mais de uma ação simultaneamente.

¹⁰ A, B, C, D e E correspondem, respectivamente, a ITAU PN, MANNESMANN ON, METAL LEVE PN, PETROBRÁS ON e SADIA PN.

¹¹ Conforme disposto na seção 5.1.

Tempo (dias)	Preço (unidade monetária)				
	A	B	C	D	E
1	3.81	0.05	17.09	136.28	0.08
2	3.81	0.05	18.47	132.57	0.08
3	3.90	0.05	18.47	134.43	0.08
4	4.16	0.05	19.40	140.92	0.08
5	4.33	0.05	21.25	143.70	0.08
6	4.94	0.05	23.09	144.62	0.09
7	4.68	0.05	23.09	142.31	0.09
8	4.68	0.06	19.40	141.84	0.10
9	4.50	0.06	18.95	139.06	0.09
10	4.38	0.06	22.17	136.28	0.09
11	4.28	0.06	21.25	139.06	0.08
12	4.33	0.06	22.17	145.55	0.09
13	4.68	0.07	21.25	145.55	0.09
14	4.53	0.07	21.25	149.72	0.09
15	4.70	0.08	24.02	154.82	0.09
16	4.70	0.08	24.94	151.11	0.09
17	4.53	0.07	24.02	148.33	0.10
18	4.53	0.07	24.02	145.55	0.10
19	4.70	0.08	24.94	155.75	0.10
20	4.79	0.08	24.02	155.29	0.10

Tabela 6.7. Preço das ações para um período de tempo de vinte dias.

Para efeitos de comparação, apresenta-se o montante capitalizado com a aplicação em Caderneta de Poupança no mesmo período. Apesar do índice de atualização dos depósitos em Caderneta de Poupança ser um índice mensal, este foi pró-rateado entre o número de dias envolvidos na transação, de modo a estabelecer uma analogia com o processo de atualização das aplicações em ações.

Tempo (dias)	Ação de retorno máximo		Retorno acumulado máximo		Retorno acumulado Poupança	Valor (unidade monetária)		
	Real	Previsão	Real	Previsão		Real	Previsão	Poupança
0			1,0000	1,0000	1,0000	1.000,00	1.000,00	1.000,00
1	C	C	1,0807	1,0807	1,0000	1.080,70	1.080,70	1.000,00
2	A	A	1,1063	1,1063	1,0083	1.106,30	1.106,30	1.008,34
3	A	A	1,1800	1,1800	1,0167	1.180,00	1.180,00	1.016,75
4	C	C	1,2926	1,2926	1,0252	1.292,60	1.292,60	1.025,23
5	A	A	1,4747	1,4747	1,0338	1.474,70	1.474,70	1.033,78
6	B	C	1,4747	1,4747	1,0424	1.474,70	1.474,70	1.042,40
7	B	B	1,7696	1,7696	1,0511	1.769,60	1.769,60	1.051,09
8	B	D	1,7696	1,7349	1,0599	1.769,60	1.734,90	1.059,86
9	C	C	2,0703	2,0297	1,0687	2.070,30	2.029,70	1.068,70
10	D	D	2,1125	2,0711	1,0776	2.112,50	2.071,10	1.077,61
11	E	C	2,3766	2,1608	1,0866	2.376,60	2.160,80	1.086,60
12	B	A	2,7727	2,3354	1,0957	2.772,70	2.335,40	1.095,66
13	D	C	2,8521	2,3354	1,1048	2.852,10	2.335,40	1.104,79
14	B	C	3,2595	2,6399	1,1140	3.259,50	2.639,90	1.114,01
15	C	C	3,3844	2,7410	1,1233	3.384,40	2.741,00	1.123,30
16	E	E	3,7604	3,0455	1,1327	3.760,40	3.045,50	1.132,67
17	A	B	3,7604	3,0455	1,1421	3.760,40	3.045,50	1.142,11
18	B	A	4,2976	3,1598	1,1516	4.297,60	3.159,80	1.151,64
19	A	A	4,3799	3,2203	1,1612	4.379,90	3.220,30	1.161,24

Tabela 6.8. Valores obtidos com a aplicação de recursos no mercado acionário e em Caderneta de Poupança.

Analisando estes resultados, é possível verificar que os investimentos orientados pelas previsões dos modelos conexionistas representam 73,52% do valor que poderia ser alcançado com a maximização dos retornos. Ou seja, o investidor teria perfeito conhecimento do preço futuro das ações.

O gráfico da figura 6.23 permite melhor visualização do comportamento dos valores deste estudo ao longo do tempo, incluindo também o retorno acumulado mínimo que seria obtido com a aplicação dos recursos nas ações que apresentaram o menor retorno para o período considerado.

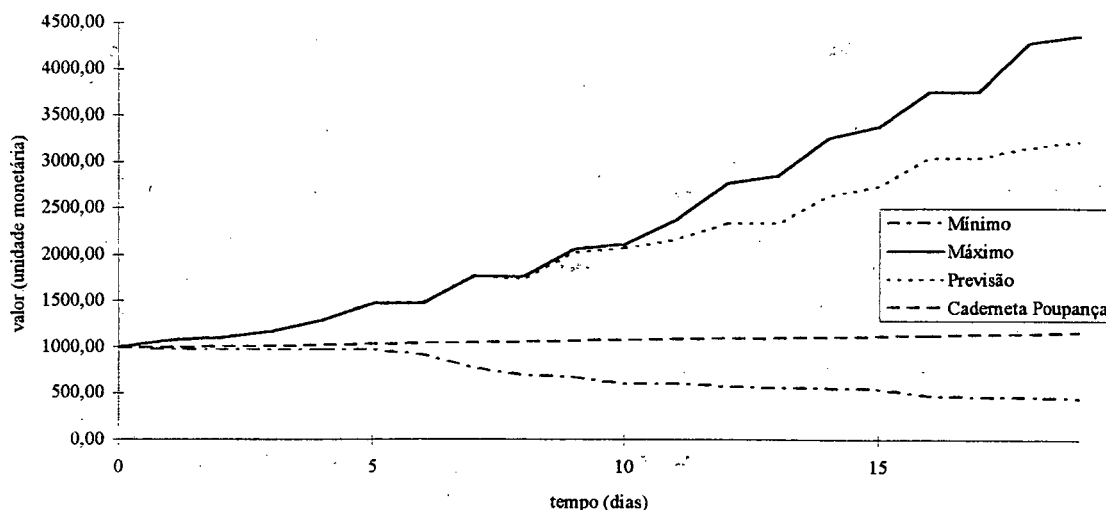


Figura 6.23. Gráfico evolutivo de recursos aplicados no mercado de ações ou em Caderneta de Poupança.

6.5. Considerações Finais

A aplicação de redes neurais artificiais na previsão do preço das ações foi avaliada através da análise empírica dos resultados obtidos com o aprendizado desses modelos pela adoção do algoritmo *backpropagation*. A sensibilidade desse algoritmo aos parâmetros e à arquitetura dos modelos de redes neurais artificiais implicou na investigação detalhada desses fatores, de modo que o modelo proposto pudesse convergir para uma solução satisfatória em um espaço de tempo aceitável.

As propriedades de convergência e generalização da rede neural, que representam a capacidade dos modelos conexionistas em produzir previsões acuradas para as séries de preços de ações apresentadas durante o processo de ajustamento dos pesos sinápticos da rede neural artificial e para aquelas apresentadas após esse ajuste, foram analisadas e as previsões obtidas comparadas com as geradas pelos métodos de previsão convencionais de Médias Móveis e Alisamento Exponencial Simples. Os erros das previsões obtidas a partir da aplicação de técnicas de redes neurais artificiais invariavelmente apresentaram magnitudes inferiores àqueles observados pela aplicação dos métodos convencionais considerados. Esta situação se confirmou para diferentes

séries de cotações de ações, reforçando a habilidade das redes neurais artificiais em identificar regularidades não-lineares presentes nas séries de preços das ações e a adequação desses modelos à previsão dos valores futuros de séries temporais econômicas irregulares, como o mercado acionário.

CAPÍTULO VII

7 - CONCLUSÕES E RECOMENDAÇÕES

O propósito deste trabalho consistiu na identificação de regularidades não-lineares presentes em séries de preços de ações pela aplicação de técnicas de redes neurais artificiais. Assumindo que essas regularidades pudessem ser detectadas e que persistiriam na série histórica das cotações das ações, a previsão dos valores futuros de algumas ações foi realizada. A partir da análise dos resultados foi possível investigar o potencial preditivo dos modelos conexionistas na manipulação das observações das séries de preços das ações, notadamente uma situação atípica em relação às séries temporais normalmente utilizadas para avaliar a preditibilidade desses modelos.

7.1. Conclusões

A acuidade das previsões geradas pelas redes neurais artificiais caracterizam os modelos conexionistas como uma ferramenta alternativa para a realização das previsões de séries econômicas. Através da aplicação da abordagem desses modelos na manipulação de séries de cotações de ações é possível alcançar previsões extremamente próximas dos valores observados, tanto em termos de tendência quanto em termos de valores absolutos. Invariavelmente, as redes neurais artificiais apresentaram desvios de menor magnitude quando comparados com aqueles obtidos pela aplicação dos métodos de previsão convencionais considerados.

Os estudos para avaliação da preditibilidade dos modelos conexionistas demonstraram a relevância do número de iterações para o algoritmo de aprendizado *backpropagation* em capturar as regularidades presentes nas séries temporais econômicas analisadas e produzir previsões confiáveis dessas séries, assim como a topologia empregada. Entretanto, desde que a modelagem da rede neural tenha sido feita de maneira criteriosa, com a evolução do processo de ajuste dos pesos sinápticos entre

as conexões da rede neural, a convergência para uma solução satisfatória pode ser alcançada em um espaço de tempo aceitável.

O treinamento exaustivo dos modelos de redes neurais artificiais não implica, necessariamente, em melhora contínua das previsões obtidas. Para tempos de treinamento mais longos, o erro verificado entre o preço da ação e o valor previsto para esse preço tende a estabilizar em um determinado nível que parece intransponível. Naturalmente, o referido nível de erro depende da série histórica de cotações considerada. Esse comportamento foi observado para as diferentes séries temporais analisadas.

A arquitetura do modelo conexionista deve ser projetada mantendo o compromisso entre a propriedade de convergência e a capacidade de generalizar adequadamente em um ambiente na qual a saída não é uma função exatamente igual à entrada, mas apresenta alguma flutuação em torno desse valor. A inclusão de unidades ocultas em número superior ao necessário pode levar à “memorização” da seqüência de preços de ações empregadas no treinamento e, portanto, prejudicar a realização de previsões dos preços futuros dessas ações. Por outro lado, uma rede neural artificial com estrutura mais simples que a necessária pode não alcançar aproximação satisfatória mesmo para as cotações apresentadas durante o processo de ajuste dos pesos sinápticos. Séries de preços de ações muito irregulares exigem maior número de unidades ocultas. A preferência de uma topologia à outra depende da característica da série em análise e deve ser sempre investigada experimentalmente.

O modelo de rede neural artificial proposto apresentou desempenho relativamente satisfatório durante o tempo decorrido entre o seu treinamento e a realização de previsões dos preços das ações, sem que houvesse necessidade de ajuste dos pesos entre as conexões do modelo. A princípio, as redes neurais artificiais são capazes de manipular o problema de instabilidade estrutural entre as observações de uma série temporal. No entanto, à medida que se está extrapolando o conhecimento adquirido no treinamento do modelo conexionista e não se tenha confirmação de permanência da estrutura da série de preços analisada, é conveniente realizar novo ajuste dos pesos sinápticos assim que

novas observações se tornem disponíveis, de modo a adaptar a rede neural artificial a essa situação.

Técnicas de redes neurais artificiais aplicadas na previsão da cotação das ações apresentaram maior sincronismo com o mercado, reagindo mais rapidamente às mudanças observadas no comportamento das séries de preços que os métodos de previsão convencionais adotados como “*benchmark*”. Isto confirma a dificuldade destes modelos em manipular informações decorrentes de ambientes extremamente dinâmicos, como o mercado acionário.

7.2. Recomendações

A partir da realização deste trabalho foi possível identificar alguns pontos que poderiam ser objeto de pesquisa futura :

- especificação de alguma metodologia que oriente a realização da previsão dos preços das ações para períodos superiores ao instante seguinte de tempo, procurando identificar as tendências gerais e os pontos de inflexão¹ mais relevantes na série de cotações das ações. Através da identificação de pontos de inflexão, o investidor poderá programar o momento ideal para comprar e vender ações;
- avaliar o desempenho dos modelos conexionistas na previsão dos preços das ações considerando, além da série de preços analisada, outra(s) série(s) temporal(ais) que tenha(m) influência significativa no mercado de ações²;
- analisar as sazonalidades *dia-da-semana*, *mês-do-ano* e outras anomalias que possam ser relevantes no comportamento dos retornos dos preços das ações;

¹ Descrito na seção 3.3.

² Na realidade, este estudo procurou estabelecer um relacionamento entre as séries de preços de ações empregadas na previsão e a série do preço do dólar americano, mas os resultados não foram positivos.

- avaliar o impacto da distribuição de dividendos no comportamento da série histórica das cotações das ações no mercado;
- estabelecer na abordagem conexionista algum critério para introduzir o risco associado ao investimento em determinada ação, ou combinação de ações, de modo que um investidor maximize o retorno esperado de seu investimento, aliado ao menor nível de risco da aplicação no mercado acionário.

REFERÊNCIAS BIBLIOGRÁFICAS

- [BARBANCHO1970] BARBANCHO, Alfonso G. **Fundamentos e Possibilidades da Econometria**. Rio de Janeiro : Forum Editora, 1970. p. 18 - 32.
- [BEALE1990] BEALE, R; JACKSON, T. **Neural Computing : an Introduction**. Bristol : Adam Hilger, 1990.
- [BOX1976] BOX, George P; JENKINS, Gwilym, M. **Time Series Analysis : Forecasting and Control**. San Francisco : Holden-Day, 1976. p. 1 - 19. c
- [CARBONELL1989] CARBONELL, J. G. **Introduction : Paradigms for Machine Learning**. [S.L. : s.n.], 1989. v. 40. p. 1 - 9.
- [CASTRO1979] CASTRO, Hélio Oliveira Portocarrero de. **Introdução ao Mercado de Capitais**. 10. ed. Rio de Janeiro : Instituto Brasileiro de Mercado de Capitais - IBMEC, 1979.
- [GRANGER1977] GRANGER, C. W. J.; NEWBOLD, Paul. **Forecasting Economic Time Series**. New York : Academic Press, 1977. c
- [FREEMAN1992] FREEMAN, James A; SKAPURA, David M. **Neural Networks Algorithms, Applications and Programming Techniques**. [S.L.] : Addelfon-Wefley, 1992. p. 89 - 106.
- [HEBB1949] HEBB, D. O. **The Organization of Behavior**. New York : John-Wiley & Sons Inc, 1949.
- [HECHT1988] HECHT-NIELSEN, R. **Applications of Counterpropagation Networks**. [S.L.] : Neural Networks Journal, 1988. p. 131-140.

- [HINTON1987] HINTON, Geoffrey. **Connectionist Learning Procedures**. [S.L. : s.n.]. Computer Science Department, Carnegie-Melon University, 1987.
- [HOFFMANN1991] HOFFMANN, Rodolfo. **Estatística para Economistas**. 2. ed. São Paulo : Livraria Pioneira Editora, 1991.
- [JONES1987] JONES, William P.; HOSKINS, Josiah. **Back-Propagation : a Generalized Delta Learning Rule**. BYTE, [S.L.], October 1987. p. 155 - 162.
- [JUDGE1988] JUDGE, George G. et al. **Introduction to the Theory and Practice of Econometrics**. 2th edition. New York : John Wiley & Sons Inc, 1988. p. 673 - 750.
- [KAMIJO1990] KAMIJO, Ken-ichi; TANIGAWA, Tetsuji. **Stock Price Pattern Recognition : A Recurrent Neural Network Approach**. In : Proceedings of the IEEE International Joint Conference on Neural Networks. San Diego : [s.n.], 1990. p. 215 - 221.
- [KIMOTO1990] KIMOTO, Takashi et al. **Stock Market Prediction System with Modular Neural Networks**. In : Proceedings of the IEEE International Joint Conference on Neural Networks. San Diego : [s.n.], 1990. p. 11- 16.
- [KOVÁCS1996] KOVÁCS, Z. L. **Redes Neurais Artificiais : Fundamentos e Aplicações**. São Paulo : Editora Acadêmica, 1996.
- [LECUN1989] LE CUN, Y. **Generalisation and Network Design Strategies**. [S.L. : s.n.]. Technical Report CRG-TR-89-4, Department of Computer Science, University of Toronto, 1989.
- [LEVY1969] LEVY, Robert A. **The Relative Strength Concept of Common Stock Price Forecasting**. New York : Investors Intelligence Inc, 1969.

[LIPPMAN1987] LIPPMAN, R. P. **An Introduction to Computing with Neural Nets.** IEEE ASSP Magazine, [S.L.], 1987.

[LO1988] LO, A.; MACKINLEY, A. C. **Stock Prices do not follow Random Walks : Evidence from a simple specification test.** Review of Financial Studies, [S.L.], 1988.

[MÁSSON1990] MÁSSON, Egill; WANG, Yih-Jeou. **Introduction to Computation and Learning in Artificial Neural Networks.** European Journal of Operational Research, North-Holand , 47, 1990. p. 1 - 28.

[MENTZER1984] MENTZER, J. T.; COX, J. E. **Familiarity, Application and Performance of Sales Forecasting Techniques.** Journal of Forecasting, [S.L.], 1984. p. 27 - 36.

[MORETTIN1981] MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Modelos para Previsão de Séries Temporais.** In : 13º Colóquio Brasileiro de Matemática. Rio de Janeiro : [s.n.], 1981.

[NELSON1973] NELSON, Charles R. **Applied Time Series Analysis.** San Francisco : Holden-Day, 1973. ✓

[OLIVEIRA1986] OLIVEIRA, Miguel Delmar de. **Introdução ao Mercado de Ações.** Rio de Janeiro : Comissão Nacional de Bolsas de Valores - CNBV, 1986.

[REFENES1991] REFENES, A. N.; ALIPPI, C. **Histological Image Understanding by Error Backpropagation.** In : Microprocessing and Microprogramming. North-Holland, 1991. p. 437 - 446.

[REFENES1993a] REFENES, A. N. et al. **Financial Modelling using Neural Networks.** London : [s.n.]. Department of Computer Sciences, University College London, 1993. p. 1 - 31.

- [REFENES1993b] _____. **Stock Performance Modeling using Neural Networks : A comparative study with regression models.** London : [s.n.]. Department of Computer Sciences, University College London, 1993. p. 1 - 18.
- [REFENES1993c] _____. **Currency Exchange Rate Prediction and Neural Network Design Strategies.** Neural Computing and Applications Journal, London, 1993. p. 46 - 58.
- [RUDGE1993] RUDGE, Luiz Fernando; CAVALCANTE, Francisco. **Mercado de Capitais.** 2. ed. Belo Horizonte : Comissão Nacional de Bolsas de Valores - CNBV, 1993.
- [RUMELHART1986] RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. **Learning Internal Representations by Error Propagation.** In : Parallel Distributed Processing : exploration in the microstructure of cognition. Cambridge : MIT Press, 1986.
- [SÁ1987] SÁ, Geraldo Tosta de. **Mercado de Ações & Bolsas de Valores.** Rio de Janeiro : Aplicação Editora Técnica Ltda, 1987.
- [SANVICENTE1988] SANVICENTE, Antônio Zoratto; MELLAGI, Armando Filho. **Mercado de Capitais e Estratégias de Investimento.** São Paulo : Editora Atlas, 1988.
- [SOUZA1989] SOUZA, Reinaldo Castro. **Modelos Estruturais para Previsão de Séries Temporais : Abordagens Clássica e Bayesiana.** In : 17º Colóquio Brasileiro de Matemática. Rio de Janeiro, 1989.
- [STEVENS1985] STEVENS, John K. **Reverse Engineering the Brain.** BYTE, [S.L.], April 1985. p. 287 - 289.

[SURKAN1990] SURKAN, Alvin J; SINGLETON, Clay. **Neural Networks for Bond Rating Improved by Multiple Hidden Layers**. In : Proceedings of the IEEE International Joint Conference on Neural Networks. San Diego : [s.n.], 1990. p. 163 - 168.

[VANHORNE1972] VAN HORNE, James C. **Funções e Análise das Taxas de Mercado de Capitais**. São Paulo : Editora Atlas, 1972.

[WASSERMAN1989] WASSERMAN, Philip D. **Neural Computing : Theory and Practice**. New York : Van Nostrand Reinhold, 1989.

[WHEELWRIGHT1985] WHEELWRIGHT, Steven C.; MAKRIDAKIS, Spyros. **Forecasting Methods for Management**. 4th edition. New York : John Wiley & Sons Inc, 1985.

[WHITE1988] WHITE, Halbert. **Economic Prediction using Neural Networks : The case of the IBM Daily Stock Returns**. In : Proceedings of the IEEE International Joint Conference on Neural Networks. San Diego : [s.n.], 1988. p. 451- 458.

[YOUNGOHC1991a] YOUNGOHC, Yoon et al. **Comparison of Discriminant Analysis vs Artificial Neural Networks**. Journal of the Operational Research Society, [S.L.], 44, 1991. p. 52 - 55.

[YOUNGOHC1991b] YOUNGOHC, Yoon; SWALES, George. **Predicting Stock Price Performance : A Neural Network Approach**. In : Proceedings of the IEEE 24th Annual International Conference of Systems Sciences. [S.L. : s.n.], 1991. p. 156 - 162.

NEURAL Computing : NeuralWorks Professional II / PLUS and NeuralWorks Explorer. [S.L.] : NeuralWare Inc, 1991.

REVISTA Investimentos. [S.L. : s.n.]. <http://www.arras.com.br>, 1996.

ANEXO

Nesta seção são apresentados alguns gráficos comparativos entre os preços reais de determinadas ações e os valores previstos pela aplicação dos modelos conexionistas na manipulação dessas séries temporais.

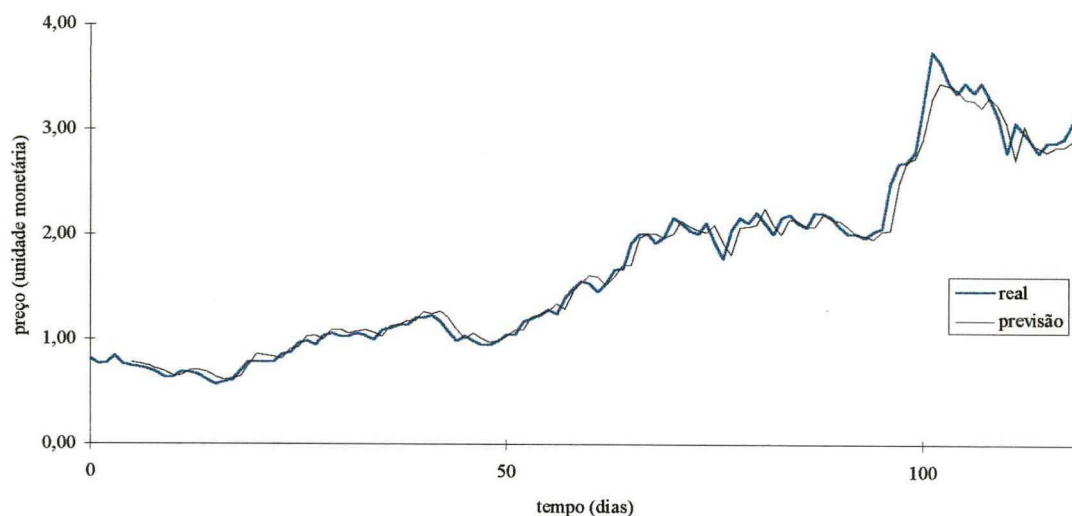


Figura 1. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência, para a ação MANNESMANN ON, do período de 16/11/88 a 19/05/89.

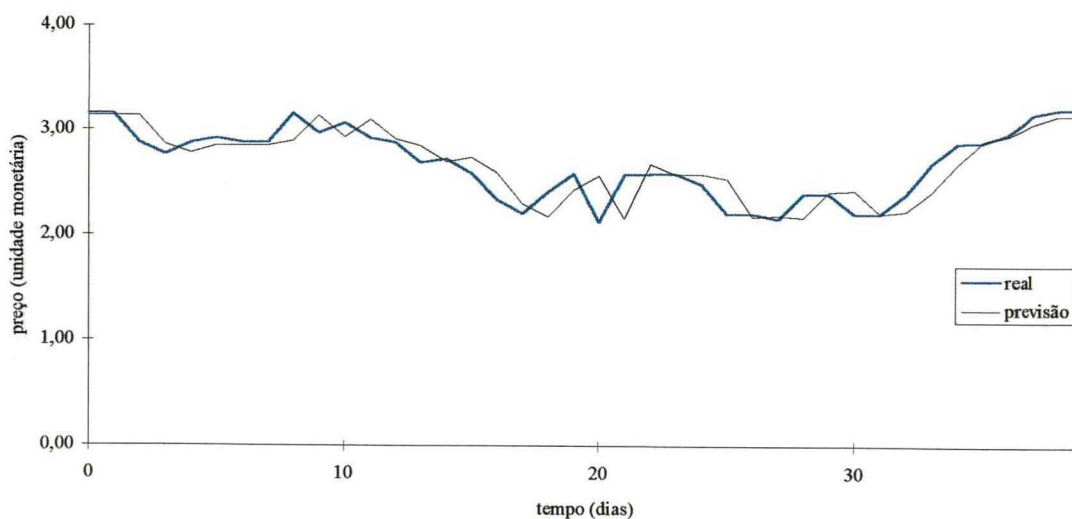


Figura 2. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização, para a ação MANNESMANN ON, do período de 15/05/89 a 18/07/89.

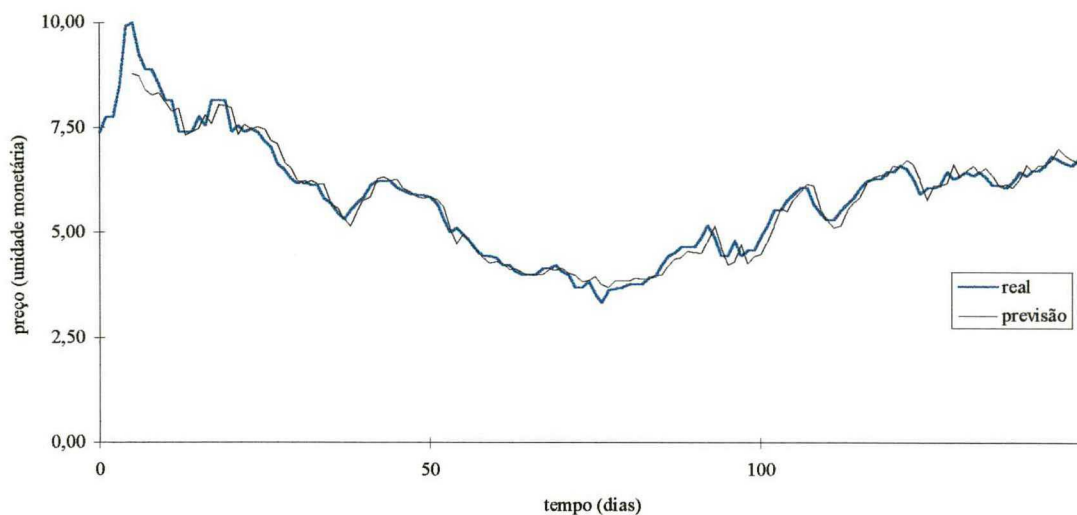


Figura 3. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência, para a ação METAL LEVE PN, do período de 22/10/86 a 12/06/87.

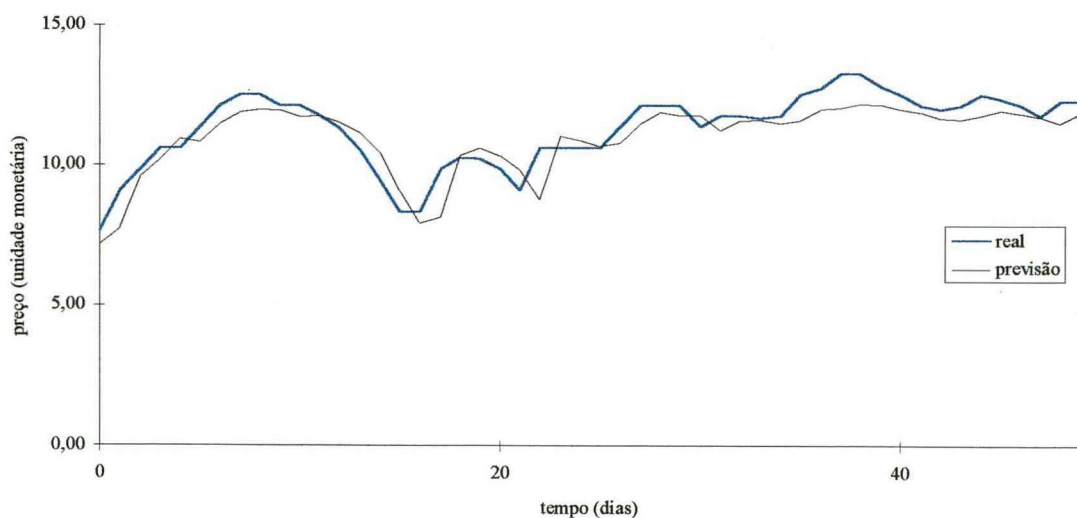


Figura 4. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização, para a ação METAL LEVE PN, do período de 08/06/87 a 24/08/87.

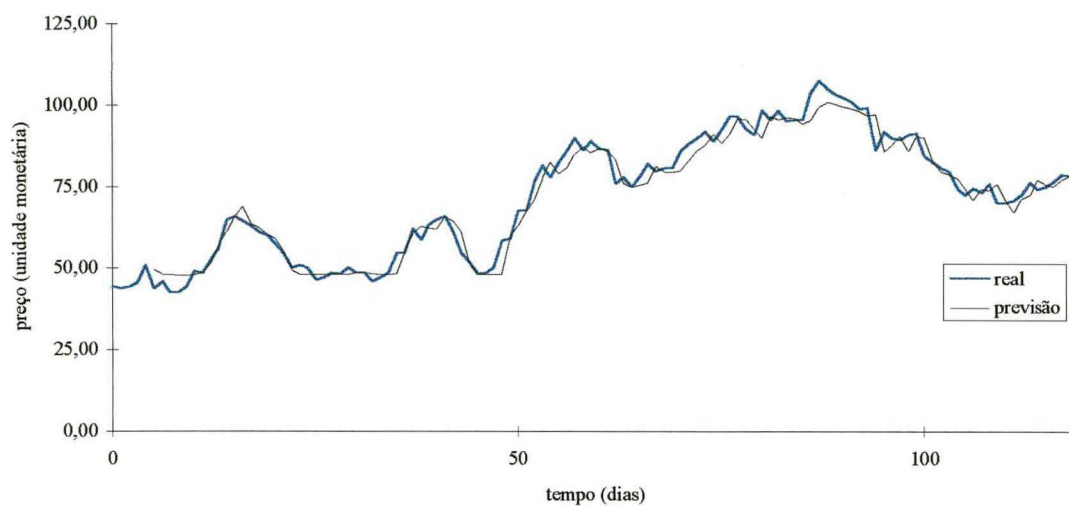


Figura 5. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência, para a ação PETROBRÁS ON, do período de 23/03/87 a 22/09/87.

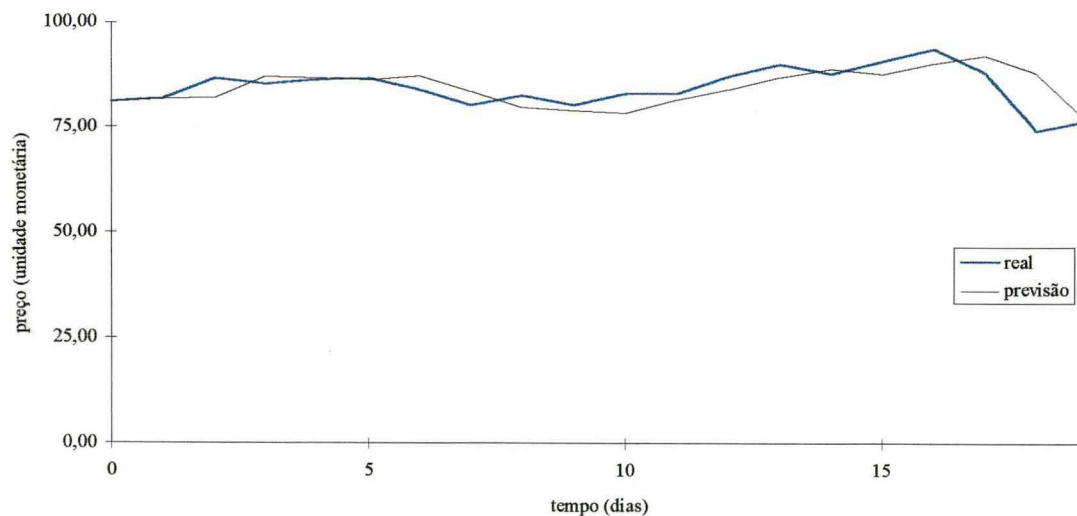


Figura 6. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização, para a ação PETROBRÁS ON, do período de 16/09/87 a 21/10/87.

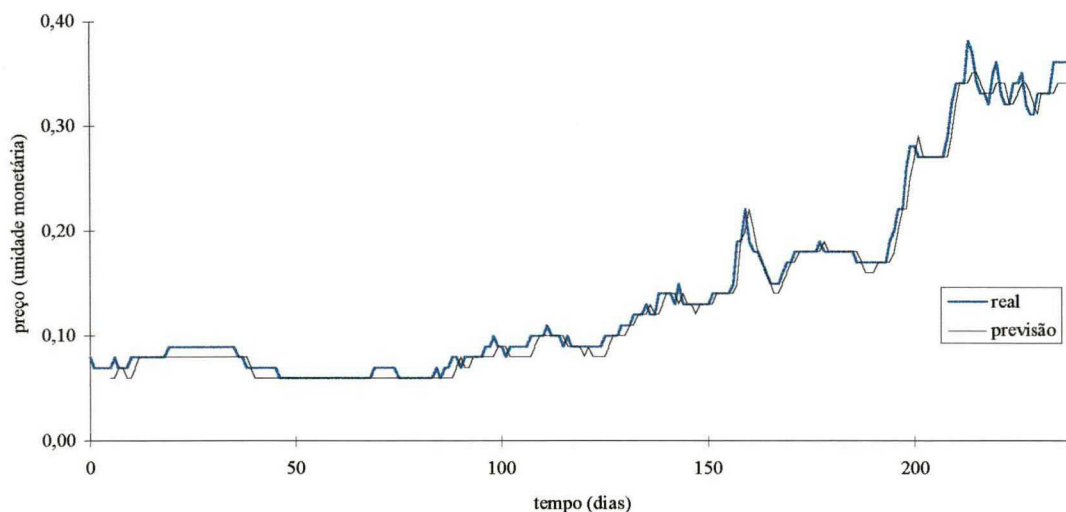


Figura 7. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de convergência, para a ação SADIA PN, do período de 27/08/87 a 22/08/88.

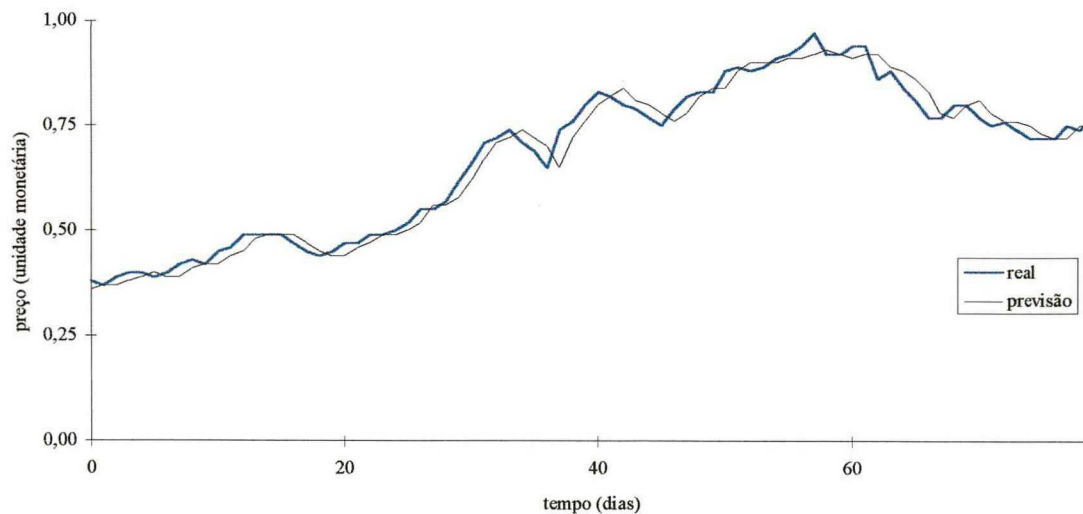


Figura 8. Gráfico comparativo entre os valores reais e os previstos pela rede neural, considerando a propriedade de generalização, para a ação SADIA PN, do período de 16/08/88 a 16/12/88.