

TESE

1139

ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ

**Um Modelo de Previsão de Indicadores Industriais  
de Vendas Utilizando Redes Neurais Artificiais**

LEANDRO ROSA DOS SANTOS

ITAJUBÁ - Dezembro

2001



# **ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ**

**Pró-Diretoria de Pesquisa e Pós-Graduação**

**Programa de Pós-Graduação em Engenharia de Produção**

## **UM MODELO DE PREVISÃO DE INDICADORES INDUSTRIAIS DE VENDAS UTILIZANDO REDES NEURAIS ARTIFICIAIS**

**Leandro Rosa dos Santos**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Produção como requisito parcial a obtenção do título de **Mestre em Ciências em Engenharia de Produção**

**Orientador : Prof. Pedro Paulo Balestrassi, Dr.**

**Itajubá, dezembro de 2001**

CLASS. 004.032.26 (043.2)

CUTI: R. 5237 m

TOMBO. 1139



Ministério da Educação  
**ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ**  
Reconhecida Lei 3232 - 05/01/1917

## **A N E X O I**

### **PRONUNCIAMENTO DA BANCA EXAMINADORA**

A Banca Examinadora, abaixo assinada, nomeada pela Portaria nº407 de 20 de dezembro de 2001, considerando o resultado do Julgamento da Prova de Defesa Pública da Dissertação de Mestrado intitulada: **“Um Modelo de Previsão de Indicadores Industriais Utilizando Redes Neurais Artificiais”** apresenta pronunciamento no sentido de que o Coordenador dos Cursos de Pós-Graduação em Engenharia de Produção da Escola Federal de Engenharia de Itajubá solicite ao DRA (Departamento de Registro Acadêmico) a expedição do título de **Mestre em Ciências em Engenharia de Produção, na Área de Concentração Qualidade**, satisfeitas as demais exigências regimentais, a **Leandro Rosa dos Santos**.

Itajubá, 20 de Dezembro de 2001.

José Roberto Alves de Mattos  
1º Examinador - FAENQUIL

Sebastião Carlos da Costa  
2º Examinador - EFEI

Luiz Fernando Barca  
3º Examinador - EFEI

Pedro Paulo Balestrassi  
4º Examinador - EFEI- Orientador



Ministério da Educação  
**ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ**  
Reconhecida Lei 3232 - 05/01/1917

**ANEXO II**

**FOLHA DE JULGAMENTO DA BANCA EXAMINADORA**

Título da Dissertação: **“Um Modelo de Previsão de Indicadores Industriais Utilizando Redes Neurais Artificiais”**

Autor: **Leandro Rosa dos Santos**

**JULGAMENTO**

Examinadores	Conceito	Rubrica
1º	A+	
2º	A+	
3º	A+	
4º	A+	

Resultado Médio: Conceito:     A+    , ou seja,     Aprovado      
    com Distinção    

Observações: \_\_\_\_\_

Itajubá, 20 de Dezembro de 2001.

José Roberto Alves de Mattos  
1º Examinador - FAENQUIL

Sebastião Carlos da Costa  
2º Examinador - EFEI

Luiz Fernando Barca  
3º Examinador - EFEI

Pedro Paulo Balestrassi  
4º Examinador – (Orientador) - EFEI

## Agradecimentos

Agradeço ao professor Pedro Paulo, meu orientador, pelo incentivo, motivação e suporte intelectual sem o qual este trabalho não teria sido tão desafiador.

Ao professor Sdnei de Brito Alves, o responsável pelo meu gosto pela pesquisa, e um grande amigo desde os tempos da minha graduação, sem o qual este trabalho nem teria sido imaginado.

Aos meus pais pelo apoio e segurança necessária, sem os quais em alguns momentos este trabalho não teria sido concluído.

Agradeço também ao meu sogro e minha sogra, que por muitas vezes me deram o suporte necessário para que este trabalho não fosse interrompido.

E a minha esposa, Cláudia, a qual o amor e a compreensão pela perda de momentos irrecuperáveis de convívio sustentaram meu ânimo para concluir este desafio.

## Sumário

<b>1</b>	<b><u>INTRODUÇÃO</u></b> .....	<b>1</b>
1.1	<u>INTRODUÇÃO</u> .....	1
1.2	<u>DESCRIÇÃO DO PROBLEMA</u> .....	1
1.3	<u>OBJETIVO PRINCIPAL</u> .....	2
1.4	<u>OBJETIVOS SECUNDÁRIOS</u> .....	2
1.5	<u>ORGANIZAÇÃO E METODOLOGIA DO TRABALHO</u> .....	2
<b>2</b>	<b><u>REVISÃO BIBLIOGRÁFICA</u></b> .....	<b>5</b>
2.1	<u>ECONOMIA</u> .....	5
2.1.1	<u>Introdução</u> .....	5
2.1.2	<u>Microeconomia</u> .....	5
2.1.3	<u>Macroeconomia</u> .....	5
2.2	<u>ECONOMETRIA</u> .....	12
2.2.1	<u>Introdução</u> .....	12
2.2.2	<u>Modelo Econométrico de Análise de Regressão Múltipla</u> .....	12
2.3	<u>REDES NEURAIS ARTIFICIAIS</u> .....	14
2.3.1	<u>Introdução</u> .....	14
2.3.2	<u>O que são as Redes Neurais Artificiais (RNAs)</u> .....	14
2.3.3	<u>Modelo de Neurônio Artificial</u> .....	18
2.3.4	<u>Treinamento</u> .....	21
2.3.5	<u>Generalização</u> .....	26
2.3.6	<u>Redimensionamento da Rede Neural</u> .....	27
2.3.7	<u>Uma Pequena Análise de resultados da Literatura</u> .....	31
2.4	<u>PROJETO DE EXPERIMENTOS</u> .....	32
2.4.1	<u>Introdução</u> .....	32
2.4.2	<u>Planejamento</u> .....	32
2.4.3	<u>Caracterização</u> .....	33
2.4.4	<u>Otimização</u> .....	33
2.4.5	<u>Verificação</u> .....	33
2.4.6	<u>Tipos de Projetos de Experimentos</u> .....	33
2.4.7	<u>Aplicações de Projetos de Experimentos e Redes Neurais Artificiais</u> .....	34
<b>3</b>	<b><u>MODELO PROPOSTO</u></b> .....	<b>35</b>
3.1	<u>REDE NEURAL</u> .....	35
3.1.1	<u>Camada de Entrada</u> .....	35
3.1.2	<u>Camada Intermediária</u> .....	39
3.1.3	<u>Camada de Saída</u> .....	39
3.2	<u>TREINAMENTO</u> .....	40
3.2.1	<u>Determinação da Configuração ótima para treinamento da Rede Neural</u> .....	40

<b>4</b>	<b><u>RESULTADOS OBTIDOS</u></b> .....	<b>48</b>
4.1	<u>AMBIENTE DE IMPLEMENTAÇÃO</u> .....	48
4.2	<u>CARACTERÍSTICAS DA BASE DE DADOS UTILIZADA</u> .....	48
4.3	<u>AVALIAÇÃO DO MODELO PROPOSTO</u> .....	48
4.4	<u>CONJUNTOS DE TREINO E TESTE</u> .....	48
4.5	<u>RESULTADOS OBTIDOS</u> .....	49
4.5.1	<i>Bens de Consumo Duráveis</i> .....	50
4.5.2	<i>Bens de Consumo Não Duráveis</i> .....	50
4.5.3	<i>Bens de Consumo Intermediários</i> .....	51
<b>5</b>	<b><u>CONCLUSÕES</u></b> .....	<b>52</b>
5.1	<u>INTRODUÇÃO</u> .....	52
5.2	<u>DESENVOLVIMENTO E CONTRIBUIÇÕES</u> .....	52
5.2.1	<i>Introdução</i> .....	52
5.2.2	<i>Revisão</i> .....	52
5.2.3	<i>Definição do Modelo Proposto</i> .....	53
5.2.4	<i>Resultados Obtidos</i> .....	53
5.3	<u>DISCUSSÕES</u> .....	53
5.3.1	<i>Redes Neurais</i> .....	54
5.3.2	<i>Modelo Econométrico</i> .....	54
5.3.3	<i>Limitações do trabalho</i> .....	55
5.4	<u>CONCLUSÕES</u> .....	55
5.5	<u>DESENVOLVIMENTOS FUTUROS</u> .....	56
<b>6</b>	<b><u>APÊNDICE I – TABELAS DE TREINO E TESTE</u></b> .....	<b>57</b>
6.1	<u>BENS DE CONSUMO DURÁVEIS</u> .....	57
6.2	<u>BENS DE CONSUMO NÃO-DURÁVEIS</u> .....	58
6.3	<u>BENS DE CONSUMO INTERMEDIÁRIOS</u> .....	59
<b>7</b>	<b><u>APÊNDICE II – CÓDIGO DE TREINAMENTO E TESTES DA REDE NEURAL</u></b> .....	<b>60</b>
<b>8</b>	<b><u>APÊNDICE III – RESULTADOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA</u></b> .....	<b>61</b>
8.1	<u>BENS DE CONSUMO DURÁVEIS</u> .....	61
8.2	<u>BENS DE CONSUMO NÃO DURÁVEIS</u> .....	63
8.3	<u>BENS DE CONSUMO INTERMEDIÁRIOS</u> .....	65
<b>9</b>	<b><u>APÊNDICE IV – CÓDIGO FONTE EXEMPLO PARA TREINAMENTO DE UMA REDE NEURAL DE TRÊS CAMADAS</u></b> .....	<b>67</b>
<b>10</b>	<b><u>REFERENCIAS BIBLIOGRÁFICAS</u></b> .....	<b>70</b>
<b>11</b>	<b><u>BIBLIOGRAFIA COMPLEMENTAR</u></b> .....	<b>73</b>

## Resumo

O presente trabalho propõe um *Modelo de Previsão de Indicadores Industriais de Vendas utilizando a técnica Redes Neurais Artificiais (RNA)*. Tal modelo abrange um amplo escopo de aplicações pois destina-se a um grande número de atividades de análise/simulação/previsão. A capacidade das redes neurais de mapear uma função que generalize um padrão de comportamento, é utilizada de forma a determinar a interação de diversas variáveis macroeconômicas no desempenho industrial. Esta capacidade permite utilizar o modelo apresentado não apenas como um meio de prever o comportamento do mercado mas também como uma forma de simular diversas situações possíveis e utilizar os resultados como parâmetros a serem investigados. A técnica em questão não demonstrou uma capacidade superior as técnicas atuais de previsão, mas o trabalho indica um campo de pesquisa fértil e desafiador. O fato de poder-se adicionar novos dados de entrada, re-mapear a função de previsão continuamente com novos dados e de utilizar a técnica combinada com os modelos atuais de previsão indica sua utilidade. Para análise dos resultados do modelo proposto mapeou-se redes neurais para três segmentos macroeconômicos : bens duráveis, bens não duráveis e bens intermediários. A configuração mais apropriada para treinamento da rede neural utilizada foi determinada através de um projeto de experimentos. A avaliação do método foi realizada comparando-se as respostas do modelo proposto (redes neurais artificiais - RNA) e as obtidas através da análise regressão múltipla (ARM) utilizando as mesmas variáveis independentes de entrada e obtendo-se a mesma variável dependente na saída.

## Abstract

The purpose of this work is present a *Sales Industrial Indicators Previewer based on Artificial Neural Network Model*. This model has a wide field of applications because can be used on analysis/simulation and preview. The artificial neural network capacity of to map one function that can generalize a standard behavior is used in order to determine the interaction of many macroeconomics variables on industrial performance. This capacity give the opportunity to use this model not only way to preview the marketing behavior but a other way to simulate many possible market situations that can be used to investigate possible market changes. The model present on this work does not demonstrate a superior capacity than multiple regression analysis, but this work indicates a wide field of challenger research. The fact that to add new data on input layer, re-map the approximation function and to combine this model with others statistical models of preview demonstrate its utility. The analysis of the purposed model was executed over three different marketing clusters: durable goods, not durable goods and intermediary goods. The better configuration to training the artificial neural network was determined by a design of experiments. The evaluation of the model results was executed by the analysis of the difference between its results and multiple regression analysis results using the same numbers of independent variables and calculating the same dependent variable.

## Lista de tabelas

<u>TABELA 2.1 – TERMINOLOGIA CORRESPONDENTE ENTRE ESTATÍSTICA E REDES NEURAIS</u> .....	17
<u>TABELA 3.1 – PRINCIPAIS VARIÁVEIS DO PROCESSO DE TREINAMENTO DE UMA REDE NEURAL DE 3 CAMADAS</u> .....	41
<u>TABELA 3.2 – RESULTADOS OBTIDOS DURANTE OS EXPERIMENTOS DA REDE NEURAL</u> .....	42
<u>TABELA 3.3 – RESULTADOS ÓTIMOS DO PROJETO DE EXPERIMENTOS</u> .....	45
<u>TABELA 3.4 – CONFIGURAÇÃO ESCOLHIDA PARA TREINAMENTO DA REDE NEURAL</u> .....	46
<u>TABELA 4.1 – RESULTADOS PARA BENS DE CONSUMO DURÁVEIS</u> .....	50
<u>TABELA 4.2 – RESULTADOS PARA BENS DE CONSUMO NÃO DURÁVEIS</u> .....	50
<u>TABELA 4.3 – RESULTADOS PARA BENS INTERMEDIÁRIOS</u> .....	51
<u>TABELA 6.1 – BENS DE CONSUMO DURÁVEIS, FONTE – REVISTA CONJUNTURA ECONÔMICA, FGV, OUTUBRO/2000 E MARÇO/2001</u> .....	57
<u>TABELA 6.2 – BENS DE CONSUMO NÃO DURÁVEIS, FONTE – REVISTA CONJUNTURA ECONÔMICA, FGV, OUTUBRO/2000 E MARÇO/2001</u> .....	58
<u>TABELA 6.3 – BENS INTERMEDIÁRIOS, FONTE – REVISTA CONJUNTURA ECONÔMICA, FGV, OUTUBRO/2000 E MARÇO/2001</u> .....	59
<u>TABELA 8.1 – RESULTADOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA PARA OS DADOS DA SÉRIE DE BENS DE CONSUMO DURÁVEIS</u> .....	61
<u>TABELA 8.2 – RESULTADOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA PARA OS DADOS DA SÉRIE DE BENS DE CONSUMO NÃO DURÁVEIS</u> .....	63
<u>TABELA 8.3 – RESULTADOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA PARA OS DADOS DA SÉRIE DE BENS INTERMEDIÁRIOS</u> .....	65

## Lista de figuras

<u>FIGURA 2.1 – DEMANDA AGREGADA E OFERTA AGREGADA (FONTE BRUNER ET AL., 1999). AS1 - CURVA DE OFERTA AGREGADA, AD1 – CURVA DE DEMANDA AGREGADA, E1 – PONTO DE EQUILÍBRIO</u> .....	6
<u>FIGURA 2.2 – MODELO DE PERCEPTRON</u> .....	18
<u>FIGURA 2.3 – MODELO DE UMA REDE PERCEPTRON-MULTICAMADA DE 3 CAMADAS</u> .....	19
<u>FIGURA 2.4 – INTERPRETAÇÃO GEOMÉTRICA DO MÉTODO DE DESCIDA PELO GRADIENTE</u> .....	22
<u>FIGURA 2.5 – DETECÇÃO DA SOLUÇÃO ÓTIMA PELO ALGORITMO DE RETROPROPAGAÇÃO DE ERRO (FONTE MATLAB VERSÃO 5.1)</u> .....	25
<u>FIGURA 3.1 – MODELO DE REDE NEURAL PERCEPTRON-MULTICAMADA DE 3 CAMADAS PROPOSTA</u> .....	35
<u>FIGURA 3.2 – PRODUÇÃO INDUSTRIAL X TAXA DE JUROS DEFASADA DE 4 MESES (FONTE – CANUTO, 2001)</u> .....	37
<u>FIGURA 3.3 – VARIAÇÃO DO DESEMPREGO X VARIAÇÃO DO PIB (FONTE - CANUTO, 2001)</u> .....	38
<u>FIGURA 3.4 – GRÁFICO DE EFEITOS MAIS IMPORTANTES PARA BENS DE CONSUMO DURÁVEIS</u> .....	42
<u>FIGURA 3.5 – GRÁFICO DE EFEITOS PRINCIPAIS PARA BENS DE CONSUMO NÃO DURÁVEIS</u> .....	43
<u>FIGURA 3.6 – GRÁFICO DE EFEITOS PRINCIPAIS PARA BENS INTERMEDIÁRIOS</u> .....	43
<u>FIGURA 3.7 – ANÁLISE DE OTIMIZAÇÃO DA RESPOSTA, PONTO ÓTIMO</u> .....	44
<u>FIGURA 3.8 – ANÁLISE DE OTIMIZAÇÃO DA RESPOSTA PARA SITUAÇÃO FORA DO PONTO ÓTIMO</u> .....	45
<u>FIGURA 3.9 – ANÁLISE DE OTIMIZAÇÃO DA RESPOSTA PARA PONTO FORA DO ÓTIMO</u> .....	46
<u>FIGURA 8.1 – GRÁFICO DE RESÍDUOS PARA A SÉRIE DE BENS DE CONSUMO DURÁVEIS</u> .....	62
<u>FIGURA 8.2 – GRÁFICO DE RESÍDUOS PARA A SÉRIE DE BENS DE CONSUMO NÃO DURÁVEIS</u> .....	64
<u>FIGURA 8.3 – GRÁFICO DE RESÍDUOS PARA A SÉRIE DE BENS INTERMEDIÁRIOS</u> .....	66

# 1 INTRODUÇÃO

## 1.1 Introdução

Boa Previsão é uma das principais exigências para o sucesso da indústria. Sem um plano de vendas razoavelmente preciso é difícil conseguir fazer entregas nas datas previstas, ou conseguir uma utilização otimizada da mão-de-obra, equipamentos e capital.

A previsão de vendas é um estudo que pode ser desenvolvido por vários métodos, mas independentemente do método abordado deverá sempre haver uma previsão de vendas para o planejamento a longo prazo.

Baseado na previsão de vendas a indústria pode elaborar um plano mestre de produção o qual será a linha mestra da empresa na determinação de sua política de compras, contratação e investimento. Uma previsão de vendas mal implementada pode causar impactos profundos e irreversíveis ao futuro de qualquer empresa.

## 1.2 Descrição do Problema

O objetivo principal de previsão é, em última análise, aplicar a melhor informação disponível em direção as metas da empresa. Considerando que a programação de produção depende da previsão de vendas, é conveniente considerar o risco de efetuar-se a programação da produção baseada em previsões. Se as vendas forem menores que as previsões, o resultado é uma baixa rotação de estoque, ou mesmo produtos invendáveis em estoque. Caso contrário a produção será menor do que era possível vender, e a produção será efetuada em regime de urgência, o que fatalmente elevará os custos industriais. Desta forma um critério rigoroso e bem elaborado de previsão deve ser modelado, utilizado e continuamente realimentado com informações de forma a permitir a melhor previsão possível.

As informações que fundamentam um estudo de previsão podem ser de dois tipos:

- Qualitativas – Pesquisa de mercado, opinião de consultores, percepção quanto a políticas governamentais, do comportamento dos agentes econômicos e etc;
- Quantitativas – Dados históricos que quantificam o desempenho dos produtos em épocas anteriores;

É importante salientar que os dados quantitativos ilustram e objetivam as previsões e a precisão das mesmas. Entretanto, a instabilidade econômica faz com que os modelos econômicos e estatísticos adotados sofram variações frequentes, sendo pequeno o seu período

de atuação, o que pode inviabilizar as previsões baseadas única e exclusivamente em dados quantitativos. Neste caso as informações qualitativas serão fundamentais.

O trabalho aqui proposto destina-se a aplicar o método de *Redes Neurais Artificiais à Previsão de Indicadores Industriais de Vendas* com o objetivo de fornecer uma visão do comportamento do mercado as variações na economia.

### 1.3 Objetivo Principal

O trabalho tem como objetivo principal : *Propor um modelo de Previsão de Indicadores Industriais de Vendas utilizando-se Redes Neurais Artificiais*. O Modelo será alimentado pelas vendas de períodos anteriores e também por variáveis macroeconômicas que descrevem as políticas governamentais e as interações entre os diversos agentes econômicos.

### 1.4 Objetivos Secundários

De forma a alcançar o objetivo principal do trabalho deve-se atingir vários objetivos secundários. Os objetivos secundários deste trabalho são :

- Revisão da técnica de redes neurais artificiais (RNA);
- Revisão de macroeconomia;
- Revisão da técnica de análise de regressão múltipla (ARM);
- Revisão da técnica de projeto de experimentos;
- Definição das variáveis de entrada e saída do modelo proposto;
- Coleta de dados para treinamento e teste do modelo proposto;
- Determinação da configuração ótima para treinamento da rede neural;
- Treinamento e teste do modelo proposto;
- Comparação dos resultados obtidos entre o modelo proposto e o modelo de análise de regressão múltipla;

### 1.5 Organização e Metodologia do Trabalho

A metodologia empregada no desenvolvimento do trabalho é baseada na coleta das vendas reais de três grandes grupos de produtos e de diversos indicadores macroeconômicos utilizados pelo governo de forma a estimular ou reprimir o crescimento econômico. O uso de dados reais foi escolhido principalmente devido aos seguintes fatores :

- Disponibilidade de dados de vendas para diversos setores industriais, coletados através de órgãos governamentais e privados;
- Disponibilidade de dados de indicadores macroeconômicos, coletados de órgãos governamentais e privados, que descrevem políticas econômicas do governo e a interação dos diversos agentes econômicos;
- Lidar com as respostas reais das vendas em função das variações dos indicadores macroeconômicos, pelo mercado ou pelo governo;
- Verificar, na prática, a adequação do modelo proposto de previsão em situações reais, verificando sua capacidade de generalização para diferentes setores da indústria;

Baseado nestes fatores e em seqüência a esta introdução o trabalho está dividido em 5 capítulos.

No capítulo 2 é realizada uma revisão da literatura, dividida em :

- Economia : O objetivo desta seção é realizar uma revisão dos principais modelos macroeconômicos e entender as relações entre as políticas do governo e suas repercussões na economia. Nesta seção também se relaciona alguns indicadores macroeconômicos com suas funções nas políticas governamentais, tais como, taxa de juros, inflação e etc.
- Econometria : O objetivo desta seção é realizar uma revisão da técnica de regressão múltipla e sua aplicação na economia. A técnica de regressão múltipla foi utilizada para comparação dos resultados obtidos com o modelo proposto.
- Redes neurais artificiais : O objetivo desta seção é realizar uma revisão da técnica de redes neurais artificiais, descrevendo suas características, as dificuldades de implementação e suas principais aplicações atuais assim como os resultados obtidos com esta técnica em diversas áreas.
- Projeto de experimentos : O objetivo desta seção é realizar uma revisão da técnica de projeto de experimentos. Esta técnica foi utilizada para determinação da configuração ótima de treinamento da rede neural artificial do modelo proposto.

No capítulo 3 é descrito como o modelo foi determinado. Neste capítulo descreve-se o modelo de rede neural utilizado, definindo-se seus parâmetros :

- Número de neurônios da camada de entrada;
- Número de neurônios da camada de saída;
- Definição dos parâmetros para treinamento da rede neural;
  - Número de neurônios da camada intermediária;
  - Função de transferência;
  - Erro total de treinamento;

Para a definição dos parâmetros de treinamento da rede neural foi realizado um projeto de experimento, onde foram analisados, os três parâmetros acima descritos e determinados os níveis de interação entre eles assim como o ponto de melhor rendimento da rede neural.

No capítulo 4 uma rede neural baseada no modelo proposto é treinada com os dados reais para os três segmentos de produtos analisados neste trabalho. Após o treinamento a rede neural foi testada de forma a verificar sua capacidade de generalização aos dados de treinamento.

No capítulo 5 os resultados obtidos através do modelo proposto são comparados com os resultados da análise de regressão múltipla sobre os mesmos dados. São realizadas então as principais considerações a respeito do desenvolvimento dessa pesquisa, discussões e conclusões.

O apêndice I apresenta os conjuntos de dados reais de vendas para os setores de bens de consumo duráveis, bens de consumo não duráveis e bens intermediários e divide os dados em dois conjuntos, o conjunto de treinamento e o conjunto de testes utilizados para verificar o modelo aqui proposto.

O apêndice II apresenta o programa desenvolvido, no MATLAB versão 5.1, para treinamento e teste do modelo de rede neural proposto.

O apêndice III apresenta os resultados da análise de regressão múltipla para os dados do apêndice I.

Finalmente no apêndice IV é apresentado o código fonte, em C++, para treinamento de uma rede neural perceptron multicamadas de 3 camadas treinada pelo algoritmo de retropropagação de erro (error backpropagation).

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Economia

#### 2.1.1 Introdução

A economia de um país está sujeita a pressões oriundas de diversas fontes, internas e externas. Os fatores que geram instabilidade e variações são geralmente não controlados e susceptíveis a percepções do mercado.

Para controlar a economia o governo utiliza-se de diversas ferramentas, tais como a taxa de juros, o nível de renda e as taxas de impostos, (Bruner et al., 1999). As políticas do governo atuam de forma a estabilizar a economia e tem repercussão sobre a demanda e oferta agregada de produtos, o que é a base para definir os parâmetros macroeconômicos a serem utilizados no modelo de rede neural proposta neste trabalho.

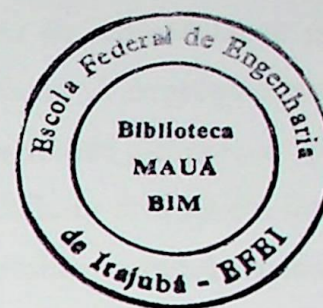
#### 2.1.2 Microeconomia

A Microeconomia freqüentemente diz respeito a escolhas. Qual a produção da minha empresa? Qual o preço de venda? Quais os volumes de vendas? Essas decisões não são tomadas no vácuo. São afetadas, literalmente, por centenas e milhares de fatores, suscetíveis de redução a um punhado de determinantes básicos. Ademais, esses determinantes ou fatores geralmente interagem uns com os outros.

#### 2.1.3 Macroeconomia

Enquanto a microeconomia trata das decisões dos indivíduos e das empresas, a macroeconomia versa sobre a situação agregada da economia. Ao se indagar sobre a situação da economia ou do mercado de ações, geralmente tem-se em mente temas macroeconômicos. A macroeconomia diz respeito ao *desemprego*, à *inflação*, ao *crescimento*, às *taxas de juros* e às *políticas governamentais* que influenciam esses elementos. Como o impacto das políticas depende, em última instância, do comportamento dos indivíduos e das empresas, a macroeconomia não pode ignorar a análise microeconômica.

Esta revisão focaliza as principais relações macroeconômicas e os dois principais instrumentos de política econômica: a *política fiscal* e a *política monetária*, de forma a escolher entre os diversos indicadores macroeconômicos disponíveis aqueles que melhor representam estas políticas e que serão utilizados no modelo proposto para previsão das vendas para os setores industriais aqui estudados (Lanzana, 2001).



Como ponto de partida, os eventos econômicos são suscetíveis de representação por um gráfico de oferta e demanda. A Figura 2.1 mostra um desses gráficos. A expressão mais comum é Produto Interno Bruto (PIB). O preço, no caso, é algum índice de preços que represente os milhões de bens e serviços produzidos na economia; o deflator do PIB é o mais amplo. As curvas de demanda e oferta agregadas geralmente são determinadas pelos mesmos fatores que influenciam a demanda e oferta individuais, mas a interpretação destes fatores deve ser realizada de modo mais abrangente. Observe que, a partir de um certo ponto QF, a curva da oferta se torna quase vertical. Esse ponto representa o nível de produção mais elevado possível na economia, no qual ocorre o pleno emprego de todos os recursos.

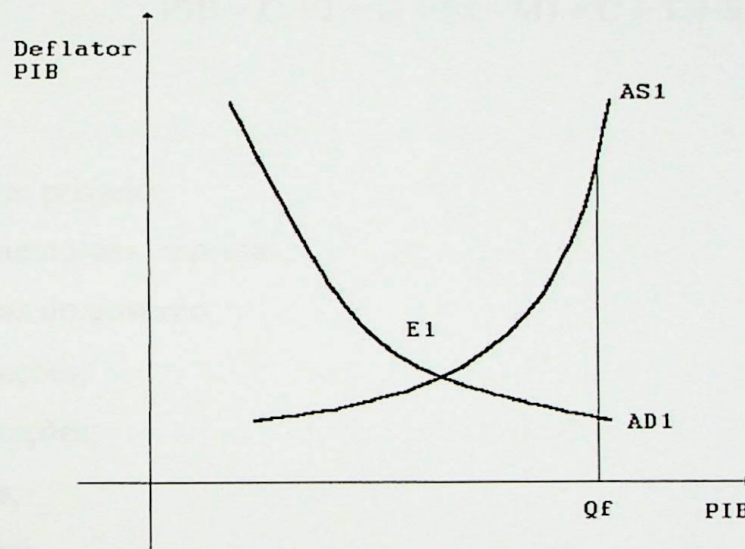


Figura 2.1 – Demanda agregada e oferta agregada (Fonte Bruner et al., 1999). AS1 - Curva de Oferta Agregada, AD1 – Curva de Demanda Agregada, E1 – Ponto de equilíbrio

### 2.1.3.1 Contabilidade Nacional

Para compreender as relações entre demanda agregada e oferta agregada e ser capaz de melhor avaliar a política econômica, os economistas começaram a desenvolver o conceito de contabilidade nacional na década de 1920. O conceito tenta medir o resultado ou produto da economia. A chave para a análise e utilização da contabilidade nacional é a compreensão de que o seu fundamento é uma identidade ou relação que, por definição, é verdadeira. Essa identidade indica que o produto total, que denomina-se PIB (Produto Interno Bruto), é igual à renda total que, por sua vez, é igual à distribuição de renda entre as várias formas de despesa e poupança:

$$\text{Produto} = \text{Renda} = \text{Despesa} \quad (2.1.1)$$

O produto interno bruto ou PIB já foi definido como o valor total das mercadorias e serviços produzidos na economia durante certo período, (Bruner et al., 1999). Uma vez que o valor de qualquer coisa é igual ao custo mais a margem de lucro, e o custo e o lucro representam a renda, conclui-se que o PIB pode ser mensurado sob uma perspectiva de renda ou de produto. Além disso, como os que auferem a renda podem consumi-la, poupá-la ou pagar impostos, a renda total deve ser igual ao PIB. Finalmente, como a economia possui vários tipos de despesas, o seu total também corresponde ao PIB. Logo, a identidade da renda nacional pode ser apresentada sob a seguinte forma:

$$\text{PIB} = C + I + G + (X - M) = C + T + S \quad (2.1.2)$$

Onde:

C = consumo privado;

I = investimento das empresas;

G = despesas do governo;

X = exportações;

M = importações;

T = tributos;

S = poupança.

Exceto quanto a algumas questões de estimativa, não há nada de controvertido sobre essas relações. Todos os economistas, independente de seu ponto de vista político, aceitam essas relações. A controvérsia entre os economistas é se o nível existente é basicamente o de pleno emprego e qual o impacto, se houver, das políticas governamentais sobre o PIB.

Se a economia está abaixo do nível de pleno emprego. Há recursos ociosos. As fábricas estão funcionando abaixo da capacidade. Existem trabalhadores desempregados e dispostos a receber menor remuneração. Com o excesso de capacidade e de disponibilidade de materiais, caem os custos de fabricação. Com a queda dos salários e dos outros custos, os fabricantes se mostram dispostos a aumentar a produção (a curva de oferta se desloca para fora); no entanto, precisam reduzir os preços para vender o aumento da produção. A diminuição dos preços não é problema para os fabricantes, em face da redução dos custos. Com a queda nos preços, os indivíduos, as empresas e o governo querem e podem comprar mais, aumentando o volume da demanda. Esse movimento cessa somente no ponto de pleno

emprego. Nesse nível de produto, não há excesso de recursos ou pressões pela queda dos salários ou dos preços.

Para os que acreditam no poder dos mercados, o exemplo é um modelo elegante e convincente. E não há dúvida de que a situação tem alguma validade empírica. Quando há excesso de capacidade, os preços de fato caem, ou ao menos sobem a índices mais baixos. No entanto, há, em geral, provas muito fortes de que a economia pode operar e de fato opera em níveis inferiores ao pleno emprego, por períodos de tempo prolongados (Bruner et al., 1999).

O caso oposto começa com a visão de que os preços e salários, embora exerçam menores pressões durante os períodos de baixa atividade econômica, raramente chegam a cair o bastante para manter a economia em pleno emprego. A razão pela qual a economia se encontra abaixo do pleno emprego é a demanda insuficiente para sustentar maiores índices de PIB. Os indivíduos dispõem de menor renda porque estão desempregados e, em consequência, gastam menos. As empresas que já estão com capacidade ociosa diminuem os investimentos em bens de capital.

Tais reduções fazem com que a demanda agregada se desloque para a esquerda e para baixo, e a economia entra em estagnação.

De forma a minimizar os impactos destas oscilações econômicas os governos tentam através do manejo das políticas monetárias e fiscal gerenciar a demanda agregada e desta forma controlar a economia (Lanzana, 2001).

Estas incursões do governo são percebidas pelo mercado e muitas vezes são informações valiosas utilizadas na previsão de vendas pelas indústrias. O trabalho proposto utiliza-se de alguns instrumentos das políticas governamentais para prever o comportamento do mercado a suas variações.

### 2.1.3.2 Política Fiscal

A política fiscal se refere ao uso das despesas e da tributação do governo para influenciar o nível da demanda agregada. As atividades governamentais exercem impacto direto e indireto sobre a demanda. Ao gastar dinheiro em programas como defesa, educação, rodovias e exploração espacial, o governo compra bens e serviços. Com isso, a demanda agregada se desloca para cima e para a direita, elevando os preços e os níveis do produto. Para atender aos níveis altos da demanda, as empresas precisam expandir a produção, por meio do aumento das horas extras ou da contratação de novos trabalhadores. De qualquer maneira, os salários aumentam e a renda adicional permite que os indivíduos elevem as suas despesas. Com isso, ocorre um aumento em  $C$ , ou seja, no consumo. Em algum ponto, as

empresas enfrentarão problemas de insuficiência de capacidade e adquirirão novos equipamentos industriais e construirão novas fábricas. O aumento em C e I induz a um deslocamento da demanda agregada para fora, superior à primeira movimentação, decorrente do aumento em G, despesas do governo. A mudança total no produto daí resultante é um múltiplo da mudança inicial na demanda. Ao falarem de política fiscal, os economistas se referem a esse efeito multiplicador.

A política fiscal trata das despesas do governo e do recolhimento de tributos. Os governos também são capazes de estimular a economia, por meio da diminuição dos tributos, em vez do aumento das despesas. Com a redução dos tributos, há um aumento na renda disponível dos indivíduos e das empresas, e uma parcela desse incremento se destina ao consumo. Essas despesas deslocam a demanda agregada para fora, dando início ao ciclo de aumento da renda e da despesa acima descrito.

Muitos programas e ferramentas tornam a política fiscal anticíclica. Quando se reduz o ritmo da atividade econômica, o governo adota uma política fiscal mais expansionista; quando o crescimento econômico está muito acelerado, a política fiscal passa a ser contracionista. Esses efeitos anticíclicos são denominados estabilizadores automáticos.

As duas atividades fundamentais em que funcionam os estabilizadores automáticos são os programas de benefícios sociais e o sistema tributário progressivo. Os programas de benefícios sociais oferecem alguma forma de pagamento pelo governo com base na renda dos indivíduos, por exemplo o valor real do salário mínimo. À medida que sobem os níveis de renda, cada vez menos pessoas fazem jus a tais benefícios, caindo as despesas governamentais. Nos sistemas tributários progressivos, a arrecadação tributária sobe como o aumento da renda. Com a progressividade das alíquotas dos tributos, a arrecadação aumenta a uma taxa crescente. O impacto dos estabilizadores automáticos consiste na redução das despesas governamentais e no aumento da arrecadação tributária, em decorrência da elevação da renda. O mecanismo coloca um freio na economia, evitando o excesso de aquecimento. Ao contrário, quando a economia se desacelera ou entra em recessão, cai a renda, aumentam as despesas do governo e reduz-se a arrecadação tributária, proporcionando o necessário estímulo à economia.

A política fiscal sozinha não tem condições de manter a estabilidade econômica. Os efeitos da política fiscal não são imediatos e outras considerações, como as de política partidária, impedem a sua oportuna implementação. A outra metade da caixa de ferramentas da política econômica é a política monetária, que, em geral, é de implementação mais rápida, pela menor dependência em relação às questões políticas

### 2.1.3.3 Política Monetária

A condução da política monetária é ditada pelo sistema da Reserva Federal, que é o Banco Central.

Como a política fiscal, a política monetária atua por diferentes caminhos : a alteração dos níveis das taxas de juros na economia e a mudança na disponibilidade de crédito. A primeira eleva ou reduz a demanda por fundos e influencia os níveis de consumo e de investimento. A outra aumenta ou diminui a disposição dos bancos para emprestar dinheiro e também afeta os níveis de consumo e de investimento. Quando a economia está muito exuberante e mostra sinais de inflação, o Banco Central eleva as taxas de juros e contrai o crédito. Se a economia está abaixo do pleno emprego, o Banco Central reduz as taxas de juros e expande o crédito.

#### 2.1.3.3.1 Taxas de Juros

Um elemento importante da macroeconomia é o nível das taxas de juros. A política monetária funciona em parte através da sua influência nas taxas de juros. Além disso, uma vez que tantas decisões empresariais e pessoais são afetadas pelas taxas de juros, esse fator é alvo de grande atenção. As taxas de juros são o preço do crédito. Representam o que paga-se quando toma-se empréstimos ou o que recebe-se quando concede-se empréstimos. O nível geral das taxas de juros na economia é determinado pela oferta e demanda de crédito e pelos índices de inflação. A taxa específica paga por um indivíduo ou por uma empresa depende do nível geral dos juros mais um prêmio pelo risco de crédito do tomador.

A taxa de juros é determinada pelo mercado através da seguinte equação :

$$\text{Taxa nominal} = \text{taxa real} + \text{expectativa inflacionaria} \quad (2.1.3)$$

As taxas reais tendem a ser influenciadas pela demanda por crédito. Durante os períodos de expansão econômica, aumenta a quantidade de pessoas dispostas a tomar empréstimos, como resultado do maior número de oportunidades de lucro. Em consequência, as taxas de juros tendem a subir durante as fases de rápido crescimento e a descer nos períodos recessivos, (Bruner et al., 1999).

#### 2.1.3.3.2 Balanço de Pagamentos

As discussões anteriores sobre a contabilidade nacional e sobre a formulação da política econômica se baseou quase que inteiramente em parâmetros internos. Mas não se

trata de uma omissão tão séria quanto possa parecer à primeira vista. Boa parte das economias nacionais ainda é muito fechada.

A taxa nominal é a taxa de juros que observa-se no mercado ou que é divulgada na imprensa. A taxa real deve ser estimada ou medida com uma visão retrospectiva, mas o importante é perceber que as taxas de juros sobem com a alta da inflação. Os países com as mais altas taxas de inflação também apresentam as mais altas taxas de juros.

O balanço de pagamentos é um registro estatístico de todas as transações com o exterior efetuadas pelos residentes no país.

A variedade de operações é muito ampla. Portanto, são muitos os indicadores ou definições no balanço de pagamento. Apenas um desses indicadores o saldo em conta corrente será considerado. O saldo em conta corrente mede todas as transações envolvendo bens, serviços, transferências unilaterais e ajuda externa. Usando a mesma notação da contabilidade nacional, o saldo em conta corrente é:

$$X - M = \text{PIB} - (C + I + G) \quad (2.1.4)$$

São muitos os fatores que influenciam a obtenção de déficits ou superávits nos saldos em conta corrente. Muitos desses fatores estão relacionados com políticas macroeconômicas que afetam os níveis de produção, consumo, investimentos e despesas governamentais. À medida que uma economia apresenta níveis de crescimento superior à de seus parceiros comerciais, aumenta o seu consumo. Geralmente, essa situação conduz a um crescimento das importações superior ao das exportações, gerando um déficit em conta corrente. As mercadorias do país com taxas de inflação superiores à dos seus parceiros comerciais se tornarão menos atraentes, acarretando também um déficit em conta corrente. As políticas econômicas adotadas para a correção dos déficits em conta corrente são, de fato, medidas destinadas a influenciar os níveis da demanda agregada (despesa) e da oferta agregada (produção), (Bruner et al., 1999).

## 2.2 Econometria

### 2.2.1 Introdução

Para a maioria dos problemas de decisão ou escolha econômica, não basta saber que determinadas variáveis econômicas estão inter-relacionadas. Precisa-se também saber a direção dos relacionamentos e as ordens de grandeza em jogo. Isto é, deve-se responder a pergunta quanto ou, pelo menos, dizer alguma coisa sobre a probabilidade de quanto.

A construção de modelos econômicos descritos estatisticamente constituem a base da econometria e nos permite estimar parâmetros desconhecidos e testar hipóteses sobre eles.

Dentre as técnicas econométricas a análise de regressão múltipla (ARM) se destaca como uma das mais eficientes e poderosas e neste sentido utilizou-se de forma comparativa os seus resultados e os obtidos a partir do modelo baseado em redes neurais proposto.

### 2.2.2 Modelo Econométrico de Análise de Regressão Múltipla

A utilização de modelos econométricos de uma forma dinâmica por meio de regressões é atualmente descrito como um dos tópicos mais avançados em previsão de séries temporais. Diversos fatores são levados em consideração nesse tipo de análise: heteroscedasticidade, multicolinearidade, raízes unitárias, correção de erros, co-integração, etc.

#### 2.2.2.1 Modelo Estatístico

Em um modelo geral de regressão múltipla, uma variável dependente  $y$  está relacionada com variáveis explicativas  $X_1, X_2, \dots, X_n$  por uma equação linear que pode ser descrita como:

$$y_t = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t \quad (2.2.1)$$

Os coeficientes  $\beta_1, \beta_2, \dots, \beta_k$  são parâmetros desconhecidos. O parâmetro  $\beta_k$  mede o efeito de uma modificação de uma variável  $X_k$  sobre o valor de  $Y_t$ , mantidas constantes todas as outras variáveis. O parâmetro  $\beta_1$  é o intercepto.

#### 2.2.2.2 Pressupostos do Modelo

1.  $y_t = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, t = 1, 2, \dots, T$
2.  $E(y_t) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} \Leftrightarrow E(\varepsilon_t) = 0$

3.  $\text{var}(y_t) = \text{var}(\varepsilon_t) = \sigma^2$
4.  $\text{cov}(y_t, y_s) = \text{cov}(\varepsilon_t, \varepsilon_s) = 0$
5. Os valores de  $X_{tk}$  não são aleatórios nem são funções lineares exatas das outras variáveis explicativas
6.  $y_t \approx N[(\beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk}), \sigma^2] \Leftrightarrow \varepsilon_t \approx N[0, \sigma^2]$

Para estimar os parâmetros desconhecidos é utilizado o processo de regressão por mínimos quadrados. Este processo de estimação dos parâmetros desconhecidos do modelo de regressão múltipla pelo processo de mínimos quadrados quando aplicado a um conjunto de dados amostrais nos fornecerá a equação para estimar o valor da variável dependente do modelo.

## **2.3 Redes Neurais Artificiais**

### **2.3.1 Introdução**

Nesta seção uma breve revisão sobre redes neurais artificiais é realizada. A seção 2.3.2 apresenta o conceito de redes neurais artificiais e na seção 2.3.3 o modelo de neurônio artificial é também sumariamente descrito. Nas seções 2.3.4, 2.3.5, 2.3.6 e 2.3.7 são tecidas considerações a respeito dos modelos de redes neurais, do treinamento, da capacidade de generalização e do redimensionamento de redes neurais. A seção 2.3.8 relata implementações de redes neurais. Uma revisão da literatura é descrita na seção 2.3.9.

Neste trabalho utilizou-se o modelo de redes neurais com o objetivo de mapear as respostas econômicas a políticas do governo e a outras variáveis macroeconômicas que influenciam na expansão e na retração das atividades produtivas.

### **2.3.2 O que são as Redes Neurais Artificiais (RNAs)**

Uma rede neural artificial é um sistema massivamente paralelo composto de elementos ou unidades altamente interconectados (também chamados de neurônios) e processamento interativo que são baseados em modelos neurobiológicos (Dayhoff, 1990).

A maior tarefa de uma rede neural é aprender e manter um determinado modelo suficientemente consistente para alcançar os objetivos de uma aplicação de interesse (Haykin, 1994). Esse sistema atua como um estimador de função não linear e não paramétrico que aprende a mapear entradas em saídas segundo uma superfície não-linear e multidimensional adequando-se a funções não lineares e multivariadas (Zurada, 1992).

Desde que as redes neurais tentam simular certos aspectos do sistema nervoso humano, elas exibem muitas propriedades para resolver problemas complexos. As seguintes características das redes neurais enfatizam o seu poder em termos de reconhecimento de padrões, fazendo-as particularmente atrativas para o problema de classificação de padrões não aleatórios :

- Podem aprender de exemplos e adaptar-se a novas situações;
- São robustas (podem abstrair informações de dados com ruído, incompletos ou não precisos);
- Podem generalizar a partir de exemplos (podem prover soluções corretas para dados similares, não vistos no treinamento)

- Podem construir soluções rapidamente sem muito domínio do conhecimento do processo;
- Podem aproximar qualquer função multivariada complexa (mal definida ou desconhecida) e formar uma decisão de classificação a partir do reconhecimento de padrões discriminantes;
- São computacionalmente eficientes (tem a velocidade para operar em tempo real).

Os dois pontos positivos mais importantes relacionados com as redes neurais são que elas podem aprender com a experiência e podem generalizar a partir de exemplos (Haykin, 1994). A elas também geralmente estão relacionados robustez e velocidade (Nelson e Illingworth, 1990). A esses pontos serão tecidas algumas considerações.

Desde que as redes neurais são adaptáveis (isto é, aprendem com a experiência) não há a necessidade de um modelo matemático a priori para a transformação entre a entrada e a saída (Balestrassi, 2000). As RNAs aprendem corretas respostas observando exemplos do mundo real. Com suficientes exemplos, as RNAs exibem uma forte capacidade de generalização. E Generalização é definida como a capacidade de produzir uma resposta razoável para uma entrada similar (mas não exatamente igual) a uma entrada apresentada durante o treinamento. Essa adaptabilidade também facilita um retreinamento para lidar com pequenas mudanças dentro do ambiente sendo modelado (Haykin, 1994).

As RNAs não são programadas, mas sim treinadas. Devido a essa característica, elas são melhores aplicadas a problemas cujas soluções requerem conhecimentos os quais são difíceis de especificar mas para os quais há uma abundância de exemplos. Isso faz com que as RNAs sejam ideais para aplicações com uma riqueza de dados e uma pobre teoria. Elas aprendem e desenvolvem soluções ajustando seu conhecimento de acordo com os exemplos por meio de regras de aprendizagem. Essas regras de aprendizagem são algoritmos eficientes de computação para o descobrimento de transformações não lineares apropriadas entre o espaço do problema e o espaço das soluções.

Nessa pesquisa, o espaço do problema representa uma coleção de dados reais obtidos a partir de instituições públicas e privadas e o espaço da solução é a previsão de vendas do indicador que se deseja prever. Uma RN apropriadamente treinada deveria então ser capaz de responder prontamente e corretamente a uma entrada não vista anteriormente. Isso indicaria que a RN generalizou o conhecimento do domínio e não meramente ‘memorizou’ o conjunto de treinamento (Zurada, 1992)

As RNAs são também mais robustas do que procedimentos sequenciais tradicionais. Devido a sua estrutura massivamente e paralelamente distribuída, o conhecimento é distributivamente armazenado através das interconexões gerando uma suave degradação na sua performance se os nós são destruídos. Essa característica de robustez também habilita a RN a abstrair características essenciais de dados com ruídos, incompletos ou parcialmente incorretos.

Em acréscimo, uma vez treinada, a eficiência computacional de uma RN torna-a capaz de prover rápidas respostas a subseqüentes entradas (Nelson e Illingworth, 1990). Isso é especialmente importante para tomada de decisão em tempo real.

Essas características das redes neurais resultam em algumas atrativas habilidades que incluem:

- Encontrar relacionamentos entre conjuntos de dados aparentemente não correlacionados e depois construir um conjunto de informações a partir deles;
- Encontrar relacionamentos que analistas humanos não podem descobrir devido suas limitações intrínsecas, quando por exemplo existe uma grande quantidade de pontos;
- Ter a habilidade e velocidade, diferentemente dos humanos, de analisar os conjuntos de dados referidos acima em uma fração de tempo.

O principal ponto fraco das redes neurais é que elas não são muito boas em justificar suas soluções e isso pode possivelmente resultar em problemas de aceitação pelo usuário. Para se ter uma boa visão das redes neurais é útil descrevê-las em relação aos sistemas especialistas e aos métodos estatísticos.

Tanto sistemas especialistas quanto RNAs representam tipos de tecnologias inteligentes (Haykin, 1994). Eles têm de certa forma a mesma intenção, ainda que difiram significativamente em muitos aspectos. Ambos tem o objetivo comum de simular a inteligência humana para resolver problemas, usando métodos diferentes de representação e processamento do conhecimento.

Os sistemas especialistas simulam e formalizam o processo do raciocínio humano baseado em símbolos. Eles representam o conhecimento explicitamente na forma de regras ou estruturas e processam esse conhecimento de uma maneira seqüencial. O aprendizado nos sistemas especialistas é formado externamente por meio da codificação do conhecimento extraído do domínio de um especialista.

As RNAs, por outro lado, procuram imitar a estrutura e funções dos neurônios do sistema nervoso central. Elas representam o conhecimento implicitamente e distributivamente com os esforços (pesos) das interconexões entre os elementos do processamento (neurônios). Elas processam esse conhecimento de uma forma paralela aplicando um raciocínio indutivo no qual padrões gerais e regras são determinados a partir dos dados e experiência (Zahedi, 1991; Haykin, 1994). O aprendizado nas RNAs acontece internamente por meio do ajuste dos pesos das interconexões durante a exposição dos exemplos. Assim como acontece em relação aos sistemas especialistas, existem similaridades e diferenças entre as RNAs e as técnicas estatísticas clássicas.

As RNAs, em fato, podem representar muitos conceitos estatísticos. Um simples *perceptron* (neurônio artificial), por exemplo, tem uma forte associação com a análise discriminante. Uma camada múltipla de *perceptrons* pode ser estatisticamente interpretada como uma função de regressão não linear estimada pela otimização de alguma medida de ajuste dos dados de treinamento (como por exemplo mínimos quadrados). Os pesos representam os parâmetros de inclinação e os *bias* (ou *offsets*) representam os pontos de intersecção das retas com a origem. De uma forma similar, as RNAs tem uma versão para a análise estatística de *clusters* que se chama aprendizagem não supervisionada.

É importante ressaltar que existe um paralelo que pode ser feito entre métodos estatísticos e redes neurais, conforme mostra a Tabela 2.1 (Balestrassi, 2000).

Estatística	Redes Neurais
Modelo	Rede
Estimação	Aprendizagem
Regressão	Aprendizagem supervisionada
Interpolação	Generalização
Observações	Conjunto de treinamento
Parâmetros	Pesos
Variáveis independentes	Entradas
Variáveis dependentes	Saídas
Regressão de Ridge	Decaimento dos pesos

Tabela 2.1 – Terminologia correspondente entre estatística e redes neurais.

Existe uma diferença fundamental, entretanto, na aproximação que essas duas técnicas usam para representar um ambiente real. Nos métodos estatísticos, os modelos matemáticos

do ambiente são formulados primeiramente. Os parâmetros são então estimados e finalmente o modelo é validado. Com RN, a arquitetura de rede é primeiramente construída (pelas características do problema) e depois disso a rede é treinada para prever respostas a um conjunto de dados de treinamento da melhor forma possível (Haykin, 1994). A solução é então obtida através da experimentação e simulação ao invés de ser através de uma aproximação formal (Zurada, 1992). Também, em estatística, algumas restrições são muitas vezes assumidas com respeito as distribuições que os dados devem ter. Nas RNAs raramente, ou nunca, suposições são feitas sobre as distribuições dos dados. As RNAs usam os dados para implicitamente desenvolver o relacionamento funcional representando o ambiente.

### 2.3.3 Modelo de Neurônio Artificial

Existem diversos modelos de neurônios artificiais, dentre os quais merece maior destaque o perceptron. Neste trabalho utiliza-se uma rede neural multicamadas baseada no neurônio perceptron, desta forma não serão apresentados outros modelos de neurônios.

#### 2.3.3.1 Perceptrons

O perceptron, invenção de Rosenblatt em 1962, foi um dos primeiros modelos de redes neurais. O perceptron modela um neurônio pegando uma soma ponderada de suas entradas e enviando o resultado 1 se a soma for maior que algum valor inicial ajustável, ou enviando 0 (zero) caso contrário. A Figura 2.2 mostra o modelo de um perceptron.

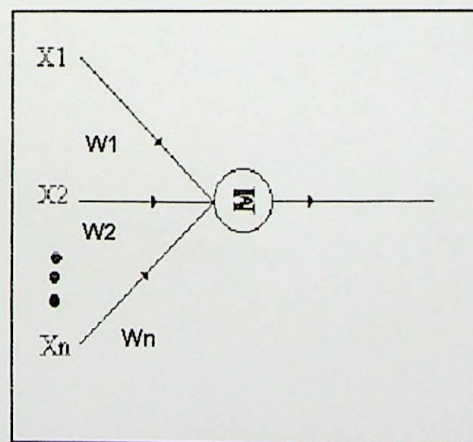


Figura 2.2 – Modelo de perceptron

As entradas ( $X_1, X_2, \dots, X_n$ ), e os pesos das conexões ( $W_1, W_2, \dots, W_n$ ) na Figura 2.2 são tipicamente valores reais, tanto positivos como negativos. Se a presença de alguma característica  $X_i$  tende a disparar o perceptron, enviando valor 1, o peso  $W_i$  será positivo; se a característica  $X_i$  inibir o perceptron o peso  $W_i$  será negativo. O perceptron em si consiste nos pesos, no processador de soma e no processador de limite ajustável. Aprendizagem é o

processo de modificar os valores dos pesos e do limite de forma a conseguir que todos os padrões reais apresentados ao perceptron possuam as saídas desejadas.

Uma rede neural baseada em um único perceptron, tem a capacidade de computar apenas problemas linearmente separáveis. Para contornar este problema foi desenvolvido então um novo modelo de rede neural derivado do modelo do perceptron, denominado de perceptron multicamadas.

### 2.3.3.2 Perceptron Multicamadas

O elemento básico de uma rede neural artificial é o perceptron. O perceptron é o modelo matemático para um neurônio biológico. Os neurônios artificiais (ou perceptrons) de uma rede neural perceptron multicamadas, são ligados uns aos outros por ligações denominadas sinapses. Cada sinapse possui associado a si um número real conhecido como peso, cujo módulo representa o ganho que o sinal recebe quando percorre a sinapse, enquanto o sinal do peso representa a natureza da ligação entre neurônios.

Uma rede neural artificial do tipo perceptron multicamadas, consiste num conjunto de perceptrons conectados por sinapses, em diferentes camadas de neurônios. Todos os neurônios de uma camada são conectados a todos os neurônios da camada seguinte, num modelo completamente conectado. Existem várias topologias de redes baseadas em perceptrons, cada uma possuindo uma classe de problemas nas quais sua utilização é mais indicada. Um modelo de perceptron multicamadas composto por 3 camadas é mostrado na Figura 2.3.

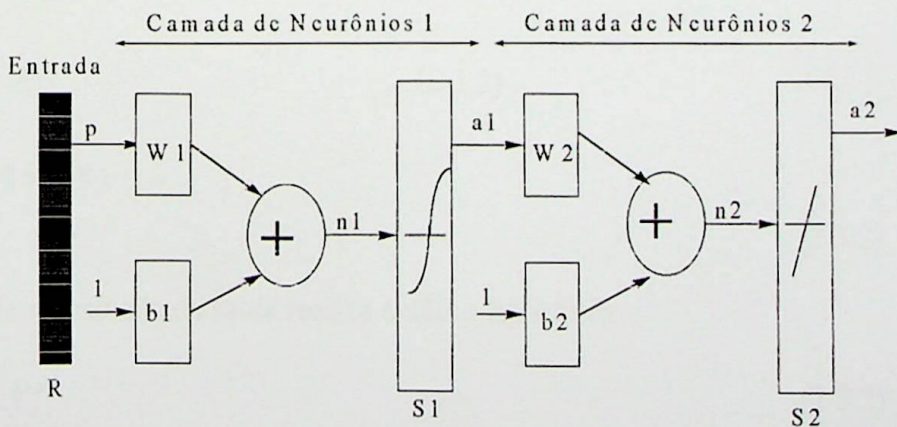


Figura 2.3 – Modelo de uma rede perceptron-multicamada de 3 camadas

Onde :

$p$  = Pesos iniciais

$R$  = Número de Entradas

$S$  = Número de Neurônios Camada  $i$

b = Bias

W = Pesos

a = Saídas após função de transferência

ni = Saídas após somatório com pesos e bias

Os pontos na entrada não realizam nenhum processamento, eles simplesmente distribuem o sinal que recebem para os neurônios da camada posterior. Uma camada é conectada apenas à camada adjacente a ela, e a saída de cada neurônio é conectada a uma entrada de todos os neurônios da camada posterior, num arranjo conhecido como camadas completamente conectadas (Nelson e Illingworth, 1990). O fluxo do sinal ocorre sempre no mesmo sentido, da entrada para a saída. Pode-se dizer neste caso que a rede é não-realimentada, em contraste com redes cujas sinapses formam caminhos cíclicos, redes realimentadas (Zurada, 1992).

Considerando-se S a camada de saída da rede neural, V a camada escondida e E a camada de entrada, pode-se demonstrar que uma rede neural mapeia uma função não linear que descreve o comportamento dos dados a ela apresentados.

Seja M o número de neurônios da camada V, N o número de entradas da rede neural. E  $E^\mu$  um vetor de entradas onde cada componente é um número real representado por  $E_k^\mu$ , para  $1 \leq k \leq N$ . Se esse vetor  $E^\mu$  for apresentado à rede, cada unidade na camada escondida  $V_j$  realiza o seguinte somatório :

$$h_j^\mu = \sum_{k=1}^N w_{jk} \cdot E_k^\mu \quad (2.3.1)$$

e produz uma saída :

$$V_j^\mu = G(h_j^\mu) \quad (2.3.2)$$

Cada unidade na camada de saída realiza então o somatório :

$$h_i^\mu = \sum_{j=1}^M W_{ij} \cdot V_j^\mu \quad (2.3.3)$$

e produz o resultado final :

$$S_j^\mu = G(h_i^\mu) \quad (2.3.4)$$

A equação acima mostra que a saída  $S^\mu$  gerada pela rede neural para uma entrada  $E^\mu$  é função não linear de  $E^\mu$ . O comportamento dessa função, uma vez fixada a função de ativação G de cada neurônio, depende exclusivamente do valor dos pesos das sinapses.

### 2.3.4 Treinamento

Para cada modelo de rede neural existe um ou mais algoritmos de aprendizagem (treinamento). Neste capítulo apresenta-se apenas o modelo de retropropagação de erro para redes neurais perceptron multicamadas, pois este modelo será a base do modelo predictor de indicadores industriais proposto neste trabalho.

#### 2.3.4.1 Algoritmo de Retropropagação de Erro

O algoritmo de retropropagação de erro requer a disponibilidade de uma base de dados com pares de entradas ou saídas, ou padrões corretos. É a partir desta base de dados que o algoritmo constrói uma configuração de pesos capaz de realizar o mapeamento desejado. Este tipo de treinamento, onde a saída da rede é comparada diretamente com os dados sabidamente corretos é conhecido como treinamento supervisionado.

Seja um conjunto de P pares  $\{E^\mu, D^\mu\}$ , com  $1 \leq \mu \leq P$ , onde cada vetor  $D^\mu$ , representa o que é desejado como saída da rede quando o vetor  $E^\mu$  é apresentado à sua entrada.

Defini-se uma função erro entre a saída desejada  $D^\mu$  e a saída  $S^\mu$  efetivamente gerada por uma rede com um conjunto de pesos  $w$  como:

$$E(w) = \frac{1}{2} \sum_{\mu} \sum [D_i^\mu - S_i^\mu]^2 \quad (2.3.5)$$

onde  $i$  varia de acordo com a dimensão do vetor de saída e  $W$  representa o conjunto de valores de todos os pesos  $w$  da rede. Aplicando a equação 2.3.4 em 2.3.5, obtém-se para a rede mostrada na Figura 2.4 a equação 2.3.6 :

$$E(w) = \frac{1}{2} \sum_{\mu} [D_i^\mu - G(\sum_{j=1}^M W_{ij} G(\sum_{k=1}^N w_{jk} E_k^\mu))]^2 \quad (2.3.6)$$

O algoritmo de retropropagação de erro procura então minimizar a função de erro expressa pela equação 2.3.6 através de uma técnica conhecida como descida pelo gradiente mais íngreme. Este procedimento de minimização possui um significado geométrico bem intuitivo. Para visualizá-lo, os valores do erro devem ser calculados para todos os valores da rede neural. Uma superfície de erro hipotética para uma rede com único peso teria a forma da Figura 2.4.

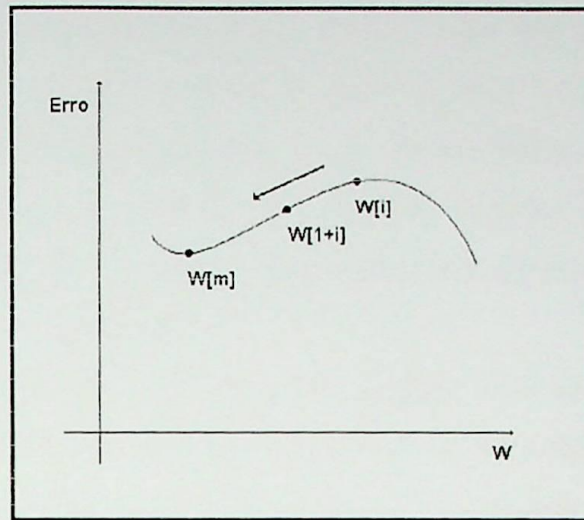


Figura 2.4 – Interpretação geométrica do método de descida pelo gradiente

Na técnica de descida pelo gradiente, o objetivo é, a partir de um ponto inicial qualquer  $w$ , chegar a um valor  $w_m$  que minimiza a função erro. Isso é conseguido iterativamente, calculando a cada passo a derivada da função erro do ponto  $w_i$  e gerando um novo ponto  $w_{i+1}$ , movendo  $w_i$  na direção contrária ao sinal da derivada. No caso da Figura 2.4, pode-se observar que a derivada no ponto  $w_i$  é positiva, ou seja, o erro aumenta se  $w_i$  aumentar. Acrescenta-se então a  $w_i$  um valor  $\Delta w$  negativo, obtendo um novo valor  $w_{i+1}$  onde o erro é menor. Em geral, tem-se :

$$\Delta w_i = -\alpha \frac{dE(w)}{dw_i} \quad (2.3.7)$$

$$w_{i+1} = w_i + \Delta w_i \quad (2.3.8)$$

Na equação 2.3.7 a constante  $\alpha$  recebe o nome de coeficiente de aprendizado. Numa rede mais genérica do que a mostrada na Figura 2.4 o erro é uma função de todos os pesos que formam a rede, e a superfície de erro é gerada num espaço de dimensão  $(N + 1)$ , onde  $N$  é o número de sinapses que formam a rede. Neste caso, a alteração  $\Delta w$  é calculada individualmente para cada peso da rede através da derivada parcial da função erro em relação a cada um desses pesos, ou seja:

$$\Delta w_{ij} = \alpha \frac{\partial E(w)}{\partial w_{ij}} \quad (2.3.9)$$

onde  $w$  é um vetor com valores dos pesos de cada sinapse da rede. Note que este tipo de procedimento impõe que a função erro seja derivável em todos os pontos, o que é satisfeito pela equação 2.3.6 desde que a função de ativação  $G$  de cada neurônio seja também derivável em todos os seus pontos.

Uma analogia útil consiste em imaginar que o valor dos pesos em cada momento indica a posição de uma esfera na superfície de erro e que essa, a cada passo do algoritmo, é levada a posição de menor altura da superfície, assim como a esfera colocada numa superfície e sujeita à ação da gravidade movimenta-se para a região de menor potencial ou altura. Deve ser notado, no entanto, que no caso da descida pelo gradiente não existe inércia no movimento da esfera, o que acontece na analogia.

A maneira como as derivadas parciais da equação 2.3.9 são calculadas é de grande importância prática para a implementação do algoritmo, e merece atenção neste momento.

A função erro para a rede mostrada na Figura 2.4 é a equação 2.3.6, e para as sinapses que ligam a camada escondida à camada de saída a regra de descida pelo gradiente é:

$$\begin{aligned}\Delta W_{ij} &= -\alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) \frac{\partial}{\partial W_{ij}} (D_i^{\mu} - S_i^{\mu}) \\ &= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) \frac{\partial}{\partial W_{ij}} (S_i^{\mu}) \\ &= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) \frac{\partial}{\partial W_{ij}} G(h_i^{\mu}) \\ &= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) G'(h_i^{\mu}) \frac{\partial}{\partial W_{ij}} \sum_j W_{ij} V_j^{\mu} \\ &= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) G'(h_i^{\mu}) V_j^{\mu}\end{aligned}$$



Onde  $G'$  representa a derivada da função de ativação  $G$  de cada neurônio. Pode-se rescrever a equação de  $\Delta W_{ij}$  como:

$$\Delta W_{ij} = \alpha \sum_{\mu} \partial_i^{\mu} V_j^{\mu} \quad (2.3.10)$$

onde defini-se :

$$\partial_i^{\mu} = G'(h_i^{\mu})(D_i^{\mu} - S_i^{\mu}) \quad (2.3.11)$$

Para calcular as variações das sinapses que ligam a camada de entrada à camada escondida deve-se derivar a equação 2.3.6 parcialmente em relação a cada  $W_{jk}$ . Utilizando a regra da cadeia obtém-se:

$$\begin{aligned}
\Delta W_{jk} &= -\alpha \frac{\partial E}{\partial W_{jk}} = -\alpha \sum_{\mu} \frac{\partial E}{\partial V_i^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}} \\
&= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) G'(h_i^{\mu}) W_{ij} \frac{\partial V_j^{\mu}}{\partial v_{ij}} \\
&= \alpha \sum_{\mu} (D_i^{\mu} - S_i^{\mu}) G'(h_i^{\mu}) W_{ij} G'(h_j^{\mu}) E_k^{\mu} \\
&= \alpha \sum_{\mu} \partial_i^{\mu} W_{ij} G'(h_i^{\mu}) E_k^{\mu} \\
&= \alpha \sum_{\mu} \partial_i^{\mu} E_k^{\mu}
\end{aligned} \tag{2.3.12}$$

com :

$$\partial_j^{\mu} = G'(h_j^{\mu}) \sum_i W_{ij} \partial_i^{\mu} \tag{2.3.13}$$

Note que as equações 2.3.12 e 2.3.10 têm o mesmo formato, apenas com uma mudança na definição do termo  $\Delta$ . Em geral, pode-se demonstrar que para um número qualquer de camadas, a regra de atualização dos pesos das sinapses tem a seguinte forma:

$$\Delta w_{ij} = \alpha \sum_{\mu} \partial_{saída} \times V_{entrada} \tag{2.3.14}$$

Onde saída e entrada se referem aos extremos  $i$  e  $j$  da sinapse em questão e  $V$  é o valor que entra num dos extremos da sinapse, proveniente da saída de um neurônio de uma camada escondida ou diretamente de uma entrada  $E_k^{\mu}$ . A definição depende da camada que está sendo considerada. Para a camada de saída, ele é definido pela equação 2.3.11, enquanto que para outras camadas a definição é dada pela equação 2.3.12.

A equação 2.3.12 permite calcular valores de  $\Delta$  para cada neurônio  $V$  numa camada escondida em termos de valores de  $\Delta$  da camada posterior. Os coeficientes  $W_{ij}$  são os mesmos utilizados para propagar o sinal da entrada para a saída da rede, mas durante o cálculo dos valores de  $\Delta w$  são utilizados para propagar o erro no sentido contrário, da saída para a entrada. Daí o nome retropropagação de erro.

Um exemplo de código fonte para treinamento pelo algoritmo de retropropagação de erro de uma rede neural com 3 camadas foi desenvolvido no Anexo IV e uma possível solução hipotética é representada na Figura 2.5.

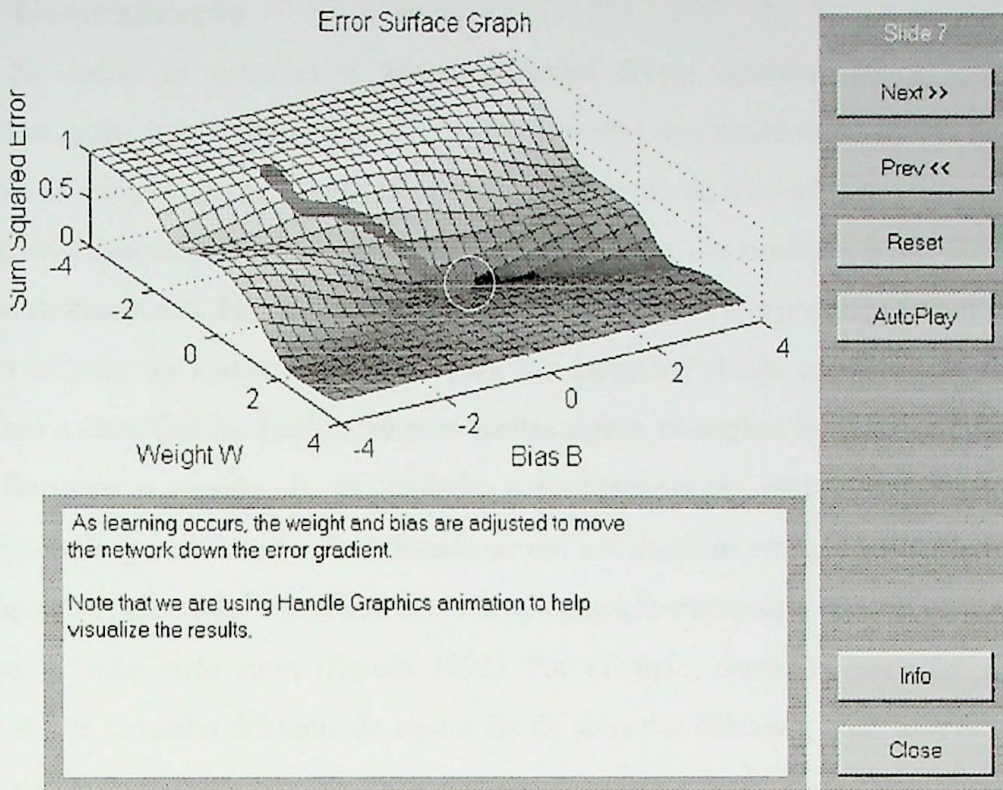


Figura 2.5 – Detecção da solução ótima pelo algoritmo de retropropagação de erro (Fonte Matlab Versão 5.1)

#### 2.3.4.2 Treinamento por lote versus estocástico

No algoritmo básico apresentado em 2.3.4.1, os valores dos pesos das sinapses são alterados após a apresentação de todos os exemplos do conjunto de treinamento, ou seja, após uma época. Neste caso, diz-se que o treinamento é por lote, e os valores de  $\Delta$  são calculados levando-se em conta a redução do valor do erro para o conjunto de treino como um todo.

Uma variação muito utilizada do algoritmo consiste na atualização dos valores dos pesos a cada exemplo apresentado à rede. Neste caso, diz-se que o treinamento é estocástico, e os valores de  $\Delta$  agora são calculados visando a diminuição do valor do erro para o exemplo que está sendo apresentado à rede no momento.

A desvantagem principal do treinamento por lote no é o longo do tempo exigido para o cálculo das atualizações dos valores dos pesos. No treinamento estocástico, os pesos são modificados mais rapidamente e isso pode levar a uma convergência também mais rápida do algoritmo, principalmente em casos onde o conjunto de treinamento é grande e contém uma quantidade considerável de informações redundantes o que geralmente aumenta desnecessariamente o processamento.

### 2.3.5 Generalização

Se todas as entradas e saídas possíveis forem apresentadas a uma rede de retropropagação, a rede (provavelmente, eventualmente) encontrará um conjunto de pesos que interpole as entradas para as saídas. Em muitos problemas de Inteligência Artificial, porém, é impossível fornecer todas as entradas possíveis. Considere, por exemplo, o reconhecimento de rostos e de caracteres. Há um número infinito de orientações e expressões de um rosto, e um número infinito de fontes e tamanhos para um caractere. Ainda assim, os seres humanos aprendem a classificá-los facilmente com apenas alguns exemplos. Seria de se esperar que as redes fizessem o mesmo. E, na verdade, a retropropagação mostra-se promissora como mecanismo de generalização. Trabalhando-se em um domínio em que entradas semelhantes são mapeadas para saídas semelhantes, a retropropagação interpolará quando receber entradas que nunca tenha visto antes (Zurada 1992). Por exemplo, depois de aprender a distinguir alguns As de tamanho diferente de alguns Bs de tamanho diferente, uma rede em geral será capaz de distinguir um A de qualquer tamanho de um B de qualquer tamanho. A generalização também ajudará a superar ruídos indesejáveis nas entradas.

Todavia, há algumas armadilhas. Durante a primeira parte do treinamento, o desempenho do conjunto de treinamento melhora quando a rede ajusta seus pesos através da retropropagação. O desempenho do conjunto de teste (exemplos que a rede não tem permissão para aprender) também melhora, embora nunca seja tão bom quanto o do conjunto de treinamento. Depois de um certo tempo, o desempenho da rede atinge um platô, de um caminho que leve a uma melhoria ainda maior. Finalmente, esse percurso é encontrado e o desempenho do conjunto de treinamento melhora de novo. Mas o desempenho do conjunto de teste piora, por que a rede começa a memorizar os pares entrada-saída em vez de procurar pesos que descrevam de modo geral o mapeamento de todos os casos. Com milhares de pesos de valores reais a sua disposição, a retropropagação é teoricamente capaz de armazenar todo o conjunto de treinamento; com suficientes unidades ocultas, o algoritmo pode aprender a atribuir uma unidade oculta a todo padrão de entrada diferente presente no conjunto de treinamento. É um atestado à força da retropropagação que isto realmente ocorra na prática.

É claro que tanto poder assim é indesejável. Há várias maneiras de impedir que a retropropagação recorra a um esquema de consulta a tabelas. Uma delas é interromper o treinamento quando o platô for atingido, com base na suposição de que qualquer outra melhoria virá através de perda de performance. Uma outra maneira é acrescentar deliberadamente pequenas quantidades de ruído às entradas de treinamento. O ruído deve ser suficiente para impedir a memorização, mas ele não deve ser muito, para não confundir o

classificador. Uma terceira maneira de ajudar a generalização é reduzir o número de unidades ocultas na rede, criando um estrangulamento entre as camadas de entrada e de saída. Confrontada com esse estrangulamento, a rede será forçada a apresentar representações internas compactas das entradas.

Finalmente, há a questão das exceções. Em muitos domínios, há regras gerais, mas há também exceções às regras. Por exemplo, geralmente o plural das palavras no português é formado acrescentando-se apenas a letra “s”, mas isto não se aplica a palavras como “lápis”, “capitão”, “tórax”. Quando muitos pares singular-plural são apresentados a uma rede neural, espera-se que ela generalizasse, apesar das exceções mas que não generalizasse a ponto de ignorar essas exceções. A retropropagação tem desempenho relativamente bom nesse aspecto, o que também ocorre com os perceptrons simples.

### **2.3.6 Redimensionamento da Rede Neural**

Ao se conceber uma rede neural, logicamente há dois parâmetros a serem definidos: o número de camadas e de neurônios que ela deverá ter.

Uma rede neural unidimensional deve conter no mínimo duas camadas de neurônios. Isto é óbvio: uma camada se destina à entrada dos dados e a outra à saída dos resultados. Contudo, redes neurais perceptrons com apenas duas camadas apresentam utilidade muito limitada. De fato, eles só são úteis se os casos a serem modelados estão separados de forma linear.

O aumento do número de camadas de neurônios melhora o desempenho das redes neurais. Sua capacidade de aprendizado aumenta, o que se traduz na melhoria da precisão com que ela delimita as regiões de decisão. É importante saber que redes neurais sem a camada oculta não conseguem modelar a operação lógica “ou exclusivo”. Tal fato foi exaustivamente utilizado pelos detratores das redes neurais em meados da década de 1960, quando só se dispunha de redes neurais desse tipo. As redes neurais com uma ou duas camadas ocultas, adequadamente treinadas, resolveram adequadamente esse problema.

Embora a necessidade de que haja pelo menos uma camada oculta na rede neural seja praticamente um ponto pacífico, há considerável controvérsia quanto ao número necessário de camadas ocultas necessárias para que o perceptron adquira o conhecimento de forma precisa. Alguns autores, como Hechit e Nielsen (1991), basearam-se num famoso pressuposto de Kolmogorov (referenciado em Fausset, 1994), para afirmar que uma rede neural com apenas uma camada oculta pode calcular uma função arbitrária qualquer a partir de dados fornecidos.

De fato, tem sido amplamente comprovado que redes neurais classificatórias efetivamente requerem apenas uma camada oculta.

Já outros pesquisadores, como Lippman (1987), alegam que uma rede neural deve conter pelo menos duas camadas ocultas, o que parece ser particularmente válido quando a saída da rede consiste de valores que variam de forma contínua.

Uma vez definido o número de camadas da rede neural, o próximo passo a ser cumprido é a definição do número de neurônios por camada.

Na camada de entrada deve haver um número de neurônios igual ao número de variáveis a serem fornecidos à rede. Eventualmente uma variável de entrada pode ser subdividida em vários neurônios, segundo um esquema binário, o que eventualmente melhora o desempenho do perceptron. Normalmente se inclui à camada de entrada um neurônio com valor constante unitário, denominado “bias” que atua analogamente a um “terra” elétrico. Seu papel é aumentar o número de graus de liberdade disponíveis no modelo, permitindo que a rede neural tenha maior capacidade de se ajustar ao conhecimento à ela fornecido.

A camada de saída deve conter um número de neurônios igual ao de variáveis que se deseja calcular. No caso de modelos classificatórios, pode-se utilizar um neurônio para cada item de classificação ou utilizar uma representação mais compacta, empregando-se técnicas binárias para diminuir o número de neurônios.

Note-se contudo, que o uso de representação binária para se reduzir o número de neurônios nas camadas de entrada e saída aumenta a carga de trabalho da camada oculta. Tal fato obriga a um aumento do número de neurônios dessa camada ou mesmo a adição de uma camada oculta suplementar para que a rede neural mantenha o mesmo nível de desempenho.

O maior desafio no dimensionamento de uma rede neural é a escolha correta do número de neurônios das camadas ocultas. Na realidade, a camada oculta é uma camada generalizadora. Ela tende a combinar os neurônios das camadas de entrada e saída em metagrupos. Por exemplo, em aplicações que requeiram o reconhecimento de caracteres, a camada interna deve “aprender” a agrupar os “pixels” provenientes da camada de entrada em forma de linhas.

Foram sugeridos vários critérios para a escolha do número ótimo de neurônios das camadas ocultas:

- Hecht-Nielsen/Kolmogorov: Uma rede neural com três camadas (entrada, oculta e saída) pode modelar qualquer função matemática contínua desde que a camada oculta contenha  $2.i + 1$  neurônios, onde  $i$  é o número de variáveis de entrada.

- Kudricky: Verificou empiricamente que, numa rede neural com duas camadas ocultas, consegue-se um desempenho ótimo quando se obedece a uma taxa de 3:1 entre o número de neurônios da primeira e segunda camada ocultas. A regra parece ser válida mesmo no caso de um grande número de variáveis de entrada.
- Lippmann: No caso de redes neurais com duas camadas ocultas, a segunda camada deve conter o dobro do número de neurônios da camada de saída. Caso a rede neural contiver apenas uma camada oculta, ela deverá ter  $S \times (i + 1)$  neurônios, onde  $i$  é o número de neurônios da camada de entrada e  $S$  é o número de neurônios da camada de saída;

Outros autores preferem definir um número máximo de neurônios que a camada oculta deve conter através da expressão 2.3.15 :

$$O_{max} = \frac{c}{10.(i + s)} \quad (2.3.15)$$

onde  $O_{max}$  é o número máximo de neurônios da camada oculta,  $c$  é o número de registros de dados utilizados na fase de treinamento da rede neural,  $i$  é o número de neurônios da camada de entrada e  $s$  é o número de neurônios na camada de saída.

De forma geral, para redes pequenas, onde o número de neurônios na camada de saída é maior do que o da entrada, a média geométrica entre o número de neurônios nas camadas de entrada e saída - ou seja,  $\sqrt{(i + s)}$  - é uma boa estimativa para o número de neurônios da camada oculta. Por outro lado, quanto mais complexo for o relacionamento entre as variáveis de entrada e de saída, maior deverá ser o número de neurônios na camada oculta.

Embora o assunto ainda seja muito controverso, em linhas gerais, pode-se afirmar que há um número ótimo de camadas ocultas e de neurônios nelas contidas que, para cada caso, leva à maximização da capacidade preditiva da rede neural utilizada. O problema é determiná-los com base nas poucas diretrizes hoje existentes. Isso normalmente exige um número muito grande de tentativas, se o problema for relativamente complexo.

Este trabalho propõe um método de determinação dos parâmetros de treinamento da rede neural baseado na técnica de projeto de experimentos.

### 2.3.6.1 Aplicações

A relativa facilidade do uso das redes neurais tem feito com que especialistas em áreas específicas do conhecimento prático as tenham utilizado com grande sucesso. De fato, é mais fácil aprender neurocomputação do que outras disciplinas. Isso é o inverso do que ocorre em outros campos da Inteligência Artificial, onde apenas “engenheiros do conhecimento” conseguem aplicar tais tecnologias de forma eficaz. A situação poderá melhorar ainda mais no futuro, a medida que se for acumulando conhecimento prático sobre as características da aplicação de redes neurais a problemas reais. Essas informações poderão ser incorporadas aos aplicativos de neurocomputação, e o usuário poderá até esquecer os detalhes técnicos relativos às redes neurais, concentrando-se apenas nas minúcias do seu problema específico. No momento, esse modo transparente de trabalho ainda não está plenamente disponível. Ainda é altamente recomendável que o usuário tenha um bom nível de conhecimento sobre o método.

Infelizmente, o embasamento teórico sobre redes neurais ainda apresenta muitas lacunas, o que torna seu uso muito empírico até o momento. Ele requer uma combinação de planejamento e pesquisa cuidadosos, “adivinhações” feitas com bom senso e várias tentativas até se atingir um resultado eficaz.

De modo geral, as redes neurais são um método de modelamento altamente recomendável para se lidar com sistemas abertos ou mais complexos, pouco entendidos e que não podem ser adequadamente descritos por um conjunto de regras ou equações. Outras tarefas indicadas para esta técnica são aquelas que requeiram tolerância à falhas, onde haja dados contaminados com ruído, que envolvam reconhecimento ou detecção de padrões, diagnóstico, abstração ou generalização. Situações onde os dados de entrada estejam incompletos ou ambíguos são um campo de atuação atualmente recomendado para a neurocomputação.

Surge então uma questão: o desempenho das redes neurais pode ser superior a outras técnicas de controle, como modelos matemáticos, estatísticos ou inteligência artificial?

Alguns autores, Maren, 1990, consideram que, até o momento, as redes neurais são uma técnica geralmente comparável, mas não superior, aos melhores modelos matemáticos e estatísticos, para a resolução de problemas em geral. É sempre recomendável verificar para uma dada aplicação o desempenho obtido pelas redes neurais e pelos métodos convencionais.

Tal recomendação é válida, entre outros motivos, pelo fato de que as redes neurais são, na realidade, caixas pretas. Embora essa desvantagem não seja exclusiva desta técnica, afinal, muitas equações empíricas obtidas via regressão múltipla também o são - muitas vezes fica difícil verificar como uma rede neural chegou a um dado resultado. Contudo, o exame dos

valores dos pesos e do grau da sensibilidade da resposta final a variações de entrada pode fornecer alguma informação a esse respeito.

Inúmeros trabalhos têm apontado melhor previsão e facilidade do uso das redes neurais em relação à regressão linear e não-linear. De fato, neste campo a aplicação de redes neurais parece ser desnecessária apenas quando a precisão fornecida por uma equação de regressão estatística já existente satisfazer as necessidades do usuário.

### **2.3.7 Uma Pequena Análise de resultados da Literatura**

Em uma análise de inúmeros trabalhos comparando redes neurais com outras técnicas (Classificação Bayesiana, Árvores de Decisão Binária, Predição Linear, Regressão Linear Múltipla, Sistemas Especialistas e Classificação pelo Vizinheiro Mais Próximo) sobre vários domínios, as redes neurais superaram essas técnicas (Renals e Rohwer, 1989; Maloney, 1990; Rogers, Ruck, Kabrisky e Tarr, 1990; Smotroff, Howells, e Lehar, 1990; Brown, Bergondi, e Archer, 1991; Tang, Wee, e Han, 1991; Kiang, 1992; Serpico, Roli, Pellegrentti, e Vemazza, 1993; Whittington, Spracklen, Haugh, e Faulkner, 1993; Corcoran e Lowery, 1995). As redes neurais tiveram bons resultados em relação aos métodos estatísticos clássicos pois sua natureza não-paramétrica não necessita assumir considerações sobre a forma de distribuição dos dados. É especialmente difícil estimar a função densidade de probabilidade em uma dimensão maior e não é conveniente usa-la para discriminação de regiões onde as densidades das classes são comparadas (Ripley, 1994). Brown, Bergondy e Archer (1991) também defendem que as redes neurais podem mapear melhor as superfícies de decisões complexas. Devido a problemas de clareza, elucidação de conhecimento e eficiência computacional as redes neurais podem desempenhar melhores resultados do que os sistemas baseados em conhecimento (Whittington e Spracklen, 1990).

Estudos de previsão comparativos utilizando-se modelos auto regressivos e a técnica de redes neurais artificiais obtiveram resultados estatisticamente iguais (Samohyl, Balestrassi e Meurer, 1998).

## **2.4 Projeto de Experimentos**

### **2.4.1 Introdução**

Na indústria, o projeto de experimentos pode ser usado para investigar sistematicamente as variáveis de processos ou produtos que influenciam na qualidade dos produtos finais. Depois da identificação das condições do processo e dos componentes dos produtos que influenciam a qualidade dos produtos finais, você pode-se direcionar os esforços de aperfeiçoamentos para melhorar a qualidade, manufaturabilidade, performance em campo e a confiabilidade do produto estudado.

Por exemplo, pode-se querer investigar a influencia do tratamento superficial e da temperatura na resistência a corrosão de mesas de metais. Para isto pode-se projetar um experimento que permita a coleta de dados na combinação de tratamento superficial/temperatura e medir a resistência a corrosão e assim encontrar as melhores condições para ajustar o processo de manufatura.

Pela escassez dos recursos é muito importante que se colete o máximo de informações através do experimento realizado. Projetos de experimentos bem projetados podem produzir significativamente mais informações do que experimentos realizados de qualquer maneira. Outra vantagem do projeto de experimento é a uma avaliação mais precisa dos efeitos mais importantes do processo. Por exemplo, se você acredita que existe uma interação entre duas ou mais variáveis, deve-se garantir a inclusão de ambas as variáveis no projeto de experimentos, o que resultará em uma melhor avaliação do problema em relação a análise de uma variável de cada vez.. Uma interação ocorre quando os efeitos de uma variável de entrada é influenciado pelo nível de outra variável de entrada.

Um projeto de experimentos é normalmente executado em quatro fases : Planejamento, Caracterização, Otimização e Verificação. O cuidadoso planejamento pode ajudar a evitar problemas na execução do plano de experimentos.

Neste trabalho a técnica de projeto de experimentos foi utilizada de forma a se determinar a configuração ótima de treinamento da rede neural proposta de forma a atingir resultados de previsão satisfatórios com o menor esforço de treinamento.

### **2.4.2 Planejamento**

Tenha certeza de que tanto o projeto quanto o sistema de medição está sob controle. Caso o processo e os equipamentos de medição não estejam sob controle deve-se garantir a repetição das configurações para o experimento. Deve-se também verificar a variabilidade do

sistema de medição. Caso a variabilidade do sistema de medição seja superior as diferenças nas medidas os resultados obtidos não terão valor.

### 2.4.3 Caracterização

O número de variáveis que podem afetar um processo, comumente é muito extenso, desta forma é preciso identificar as variáveis chaves para a realização dos experimentos. Esta redução permite que os esforços dos experimentos concentrem-se nas variáveis mais importantes. A caracterização do processo por meio das variáveis mais importantes pode sugerir a configuração mais apropriada para os fatores chaves do processo.

Os métodos a seguir são freqüentemente usados na etapa de caracterização:

- Projetos de experimentos fatoriais
- Projeto de robustez;
- Projetos de experimentos fatoriais completos

### 2.4.4 Otimização

Depois da caracterização do experimento e da identificação das variáveis chaves do processo deve-se determinar os valores ótimos de forma a maximizar ou minimizar os resultados do processo.

Os métodos de otimização de processos comumente utilizados são : projeto de experimentos fatorial completo, projeto de superfícies de resposta, projetos hierárquicos e projetos taguchi.

### 2.4.5 Verificação

A verificação envolve a análise experimental dos valores ótimos determinados durante a otimização de forma a confirmar os resultados obtidos na prática.

### 2.4.6 Tipos de Projetos de Experimentos

Existem diversos tipos de projetos de experimentos, dentre os quais destacam-se : projeto de experimento fatorial, projeto de superfícies de resposta, projetos hierárquicos e projetos robustos. Neste trabalho utilizou-se o método de projeto de experimentos de fatorial completo.

#### 2.4.6.1 *Projetos de Experimentos Fatoriais*

Projetos fatoriais permitem o estudo simultâneo dos efeitos de diversos fatores num processo. A execução de um experimento com a variação de todos os fatores que influenciam

no processo simultaneamente é mais eficiente e barato que a variação de um fator de cada vez, pois permite além da verificação da influencia do fator no processo determinar a interação entre os diversos fatores. A realização de um experimento sem a variação simultânea dos fatores pode permitir que interações importantes entre os fatores permaneçam indetectáveis.

#### **2.4.7 Aplicações de Projetos de Experimentos e Redes Neurais Artificiais**

A técnica de projeto de experimentos tem sido usada por diversos autores para determinação de configurações ótimas em aplicações práticas. Recentemente alguns trabalhos tem utilizado a técnica de projeto de experimentos aplicada a determinação das configurações ótimas de modelamento, incluindo o modelo de redes neurais. Gaafar e Choueiki(2000); Tsai e Wang(2001); utilizaram a técnica de projeto de experimentos de forma a definir a melhor arquitetura da rede neural a ser aplicada em seus trabalhos. Kleijnen e Sargent(2000); propuseram uma metodologia para verificação e validação de meta modelos em simulações onde a técnica de projeto de experimentos é utilizada para especificar o projeto do meta modelo. Shea(1997), utilizou a técnica de projeto de experimentos para determinar o ponto de utilização econômica da capacidade uma refinaria de petróleo.

A tendência em utilizar técnicas de determinação de configurações ótimas para determinar os parâmetros de treinamento de redes neurais pode melhorar significativamente os resultados da aplicação desta técnica e ajudar a superar um dos maiores obstáculos de sua utilização.

### 3 MODELO PROPOSTO

O modelo proposto utiliza-se das vendas e de indicadores macroeconômicos anteriores ao período a ser previsto como entrada de uma rede neural com três camadas e possui uma saída que representa o valor previsto das vendas para o próximo período. A figura 3.1 mostra esquematicamente a rede neural proposta.

#### 3.1 Rede Neural

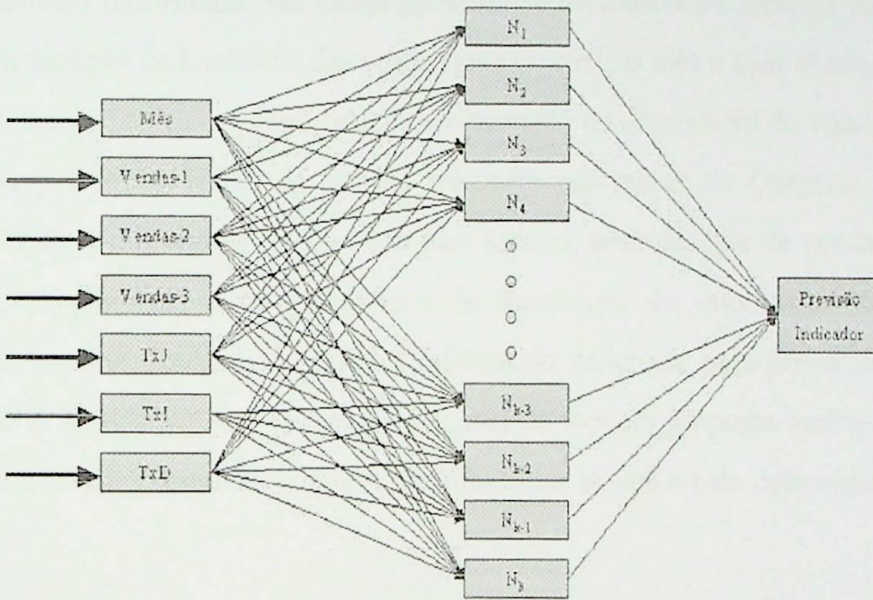


Figura 3.1 – Modelo de rede neural perceptron-multicamada de 3 camadas proposta

##### 3.1.1 Camada de Entrada

A camada de entrada do modelo proposta é composta de 7 variáveis :

- Mês da previsão;
- Venda do mês anterior (V-1);
- Venda realizada há dois meses atrás (V-2);
- Venda realizada há três meses atrás (V-3);
- Taxa de Juros do mês anterior;
- Taxa de inflação do mês anterior;
- Taxa de desemprego do mês anterior;

Estas variáveis foram escolhidas em função de vários aspectos : dentre os quais destacam-se fatores ambientais e econômicos.

#### 3.1.1.1 *Mês da previsão*

O mês da previsão é um fator de extrema importância nas vendas pois ele afeta o volume total vendido de um bem ou serviço devido a presença ou não de datas comemorativas, tais como Natal, Ano Novo, Carnaval, Páscoa e etc.

#### 3.1.1.2 *Vendas Anteriores*

A inclusão das vendas dos meses anteriores é necessária de forma a fornecer a rede neural uma indicação da tendência das vendas para o próximo mês a qual aliada a informação do mês da previsão indica o comportamento de aumento ou diminuição do volume de vendas, por exemplo : uma tendência de vendas crescente nos meses de Outubro, Novembro e Dezembro aliada a informação da previsão para Janeiro, normalmente de vendas inferiores a Dezembro, indica a rede o comportamento de diminuição do valor total vendido para o próximo mês (Janeiro) mas deve fornecer informação suficiente para prever um Janeiro de boas vendas. A utilização de três meses anteriores no modelo proposto baseia-se na divisão climática do ano em 4 estações o que pode influenciar as vendas de determinados produtos sazonais.

#### 3.1.1.3 *Taxa de Juros do mês Anterior*

Um dos mais importantes elementos de política econômica é o nível das taxas de juros. Como já descrito no Capítulo 2, a política monetária funciona em parte através da sua influência nas taxas de juros a qual tem a capacidade de estimular ou não o aumento da produção.

A Figura 3.2 mostra a influência da taxa de juros no aumento/diminuição da produção. Fica evidente a influência da variação da taxa de juros no crescimento/diminuição da produção industrial. No gráfico 3.2 a taxa de juros está defasada de 4 meses, o que indica que seus efeitos na produção são percebidos após algum tempo. A escala da taxa de juros está invertida em relação a produção, o que indica que uma diminuição da taxa de juros acarreta um aumento da produção, enquanto uma elevação da taxa de juros acarreta uma diminuição da produção. Este resultado vem de encontro com a teoria, conforme visto no item 2.1.3.3.1.

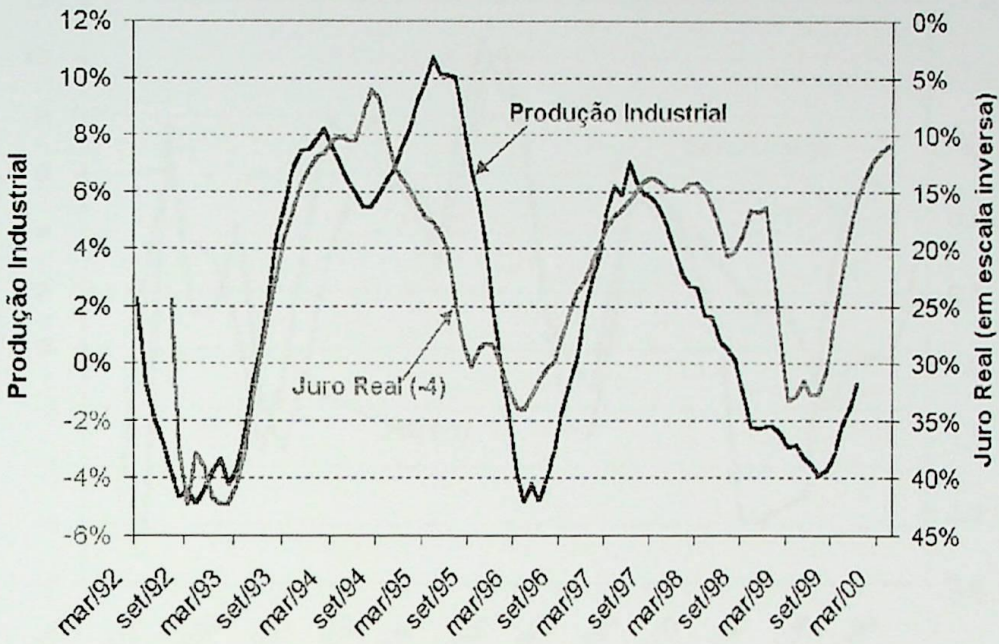


Figura 3.2 – Produção Industrial x Taxa de Juros defasada de 4 meses (Fonte – Canuto, 2001)

#### 3.1.1.4 Taxa de Inflação do mês Anterior

O controle da taxa de inflação é um dos mais importantes objetivos das políticas econômicas. A inflação é o resultado dos diversos fatores que influenciam o comportamento econômico. Apesar de ser um efeito o controle da taxa de inflação é sempre analisado na definição das políticas econômicas do governo e desta forma influencia o nível de produção e consumo do mercado.

#### 3.1.1.5 Taxa de Desemprego

O nível de desemprego na sociedade é um fator determinante para análise da renda disponível no mercado e desta forma afeta decisivamente tanto a produção quanto o consumo. A Figura 3.3 mostra a influência da taxa de desemprego no aumento/diminuição da produção nacional. Fica evidente a influência da variação da taxa de desemprego no crescimento/diminuição da produção (PIB). No gráfico abaixo o PIB está defasado de 1 trimestre em relação a taxa de desemprego, o que indica que os efeitos de sua variação na produção são percebidos após algum tempo. A escala da taxa de desemprego está invertida em relação a do PIB, o que indica que uma diminuição da taxa de desemprego é acompanhada de um aumento da produção, enquanto uma elevação da taxa de desemprego é acompanhada de uma diminuição da produção.

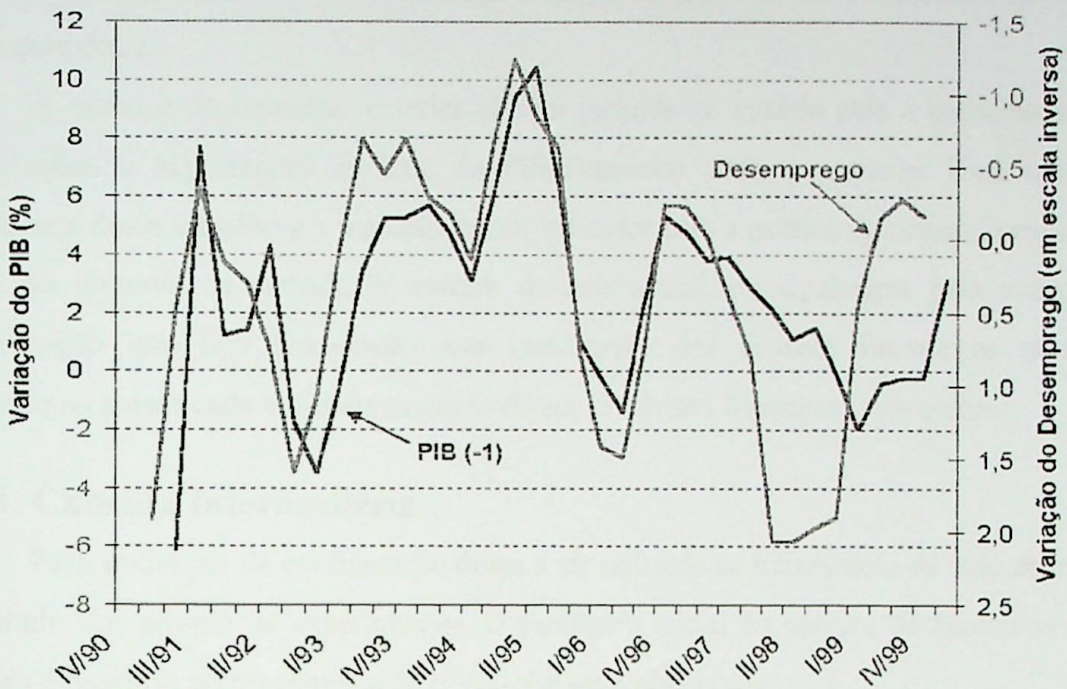


Figura 3.3 – Variação do Desemprego x Variação do PIB (Fonte - Canuto, 2001)

### 3.1.1.6 Variáveis Macroeconômicas não utilizadas no modelo

A quantidade de variáveis econômicas disponíveis é incalculável e a inclusão ou não na camada de entrada está intimamente relacionada ao tipo de mercado que se quer estudar e a disponibilidade dos dados. Para os objetivos deste trabalho o mercado foi segmentado em três grupos de produtos : Bens de Consumo Duráveis, Não Duráveis e Intermediários. Estes segmentos são expressões de diversos tipos de produtos disponíveis no país e sua análise é macroeconômica. Assim sendo as variáveis de entrada da rede foram escolhidas de forma a representar alguns dos mais importantes instrumentos de política econômica. A política econômica de um país se divide em quatro grandes grupos : política fiscal, monetária, de renda e de comércio exterior. O modelo proposto não considera em sua análise indicadores macroeconômicos para as políticas fiscal e de comércio exterior.

- Política Monetária (Taxa de Juros)
- Política de Renda (Taxa de Desemprego)
- Política Fiscal
- Política de Comercio Exterior

Existem diversos outros indicadores das políticas monetária e de renda, a escolha dos indicadores acima foi realizada em função da disponibilidade de dados e de sua importância macroeconômica.

A política de comércio exterior não foi incluída no modelo pois a participação das importações e exportações no total do PIB Brasileiro ainda é pequena. Uma extensão importante deste trabalho é a inclusão de um indicador para a política de comércio exterior e fiscal do governo na camada de entrada da rede neural, principalmente pelo avanço da globalização que tem ocasionado uma interligação dos sistemas financeiros mundiais tornando os países cada vez mais susceptíveis aos problemas financeiros dos outros.

### 3.1.2 Camada Intermediária

Para definição da configuração ótima a ser utilizada no treinamento da rede neural foi elaborado um projeto de experimentos. O parâmetro inicial do número de neurônios desta camada é baseado no pressuposto de Hecht-Nielsen/Kolmogorov:

“Uma rede neural com três camadas (entrada, oculta e saída) pode modelar qualquer função matemática contínua desde que a camada oculta contenha  $2 \times I + 1$  neurônios, onde  $I$  é o número de variáveis de entrada.”

Desta forma como número mínimo de neurônios para esta camada utilizou-se o valor  $15 = 7$  (Neurônios da Camada de Entrada)  $\times 2 + 1$  (Neurônios da Camada de Saída) de forma a garantir o modelamento de qualquer função de aproximação.

Outra definição importante baseada no teorema acima é da função de transferência para esta camada, que deve ser não linear de forma a interpolar uma solução para qualquer tipo de problema. A função escolhida para esta camada foi a função sigmoidal por sua natureza não linear assim como também pelo fato de sua saída ser formada apenas por valores positivos entre 0 e 1, visto que os valores de venda sem exceção serão sempre positivos.

### 3.1.3 Camada de Saída

A camada de saída é formada por um único neurônio o qual representa o valor projetado para as vendas do mês que se quer prever. A função de transferência da camada de saída também foi determinada via projeto de experimentos.

## 3.2 Treinamento

Para definição da configuração ótima a ser utilizada para treinamento da rede neural foi elaborado um projeto de experimentos  $2^3$  fatorial completo. Esta metodologia foi adotada de forma a se determinar a configuração ótima de treinamento para a rede neural proposta no modelo previsor, desta forma minimizando o tempo total de treinamento pela diminuição das opções de variações nas configurações das variáveis controláveis e da aquisição de conhecimento dos efeitos principais das mesmas e suas interações sobre o processo de treinamento da rede neural.

Foram estudados três parâmetros de treinamento da rede neural no projeto de experimento avaliado :

- Tipo da função de aproximação : Sigmoidal ou Linear
- Número de Neurônios na camada intermediária
- Erro total da rede neural para treinamento

A escolha destes fatores está intimamente ligada as características dos dados que se quer prever e a sua influencia na capacidade de generalização da rede neural proposta. É sabido que muitos neurônios na camada intermediária, assim como um erro total muito pequeno, tem influencia decisiva na capacidade de generalização da rede.

### 3.2.1 Determinação da Configuração ótima para treinamento da Rede Neural

Para se determinar a configuração ótima de treinamento da rede neural foi realizado um projeto de experimentos  $2^3$  Fatorial completo, baseado em três variáveis de treinamento da rede neural proposta. Para realização do projeto de experimentos foi utilizado o software Minitab versão 13.

#### 3.2.1.1 Planejamento do Experimento $2^3$ Fatorial

Uma das maiores dificuldades para a utilização da técnica de redes neurais está em se determinar alguns de seus parâmetros de treinamento que minimizem o tempo total de treinamento e maximizem a capacidade de generalização da função de aproximação. A rede neural utilizada neste trabalho (perceptron multicamadas, treinada por error backpropagatin) possui vários parâmetros que dependendo do valor podem fazer com que o treinamento atinja os resultados esperados mais ou menos rapidamente. No projeto de experimentos aqui

descrito analisou-se três dos mais importantes parâmetros da rede neural proposta. A Tabela 3.1 mostra a definição dos parâmetros investigados :

Variável	Faixa de Teste
Tipo da função de aproximação	Sigmoidal ou Linear
Número de Neurônios na camada intermediária	15 a 25
Erro total da Rede Neural para Treinamento	0,0001 a 0,001

Tabela 3.1 – Principais variáveis do processo de treinamento de uma rede neural de 3 camadas

Conforme descrito no item 2.3.5 a capacidade de generalização de uma rede neural atinge um ponto máximo durante o treinamento, a partir do qual começa a decorar os padrões de entrada degradando sua performance. Um número muito grande de neurônios na camada intermediária diminui o tempo de treinamento, mas também diminui a capacidade de generalização da rede. O erro total de treinamento determina o grau de precisão com que a função de aproximação se adequa ao conjunto de treinamento. Um erro muito pequeno tende a fazer com que a rede decore os padrões de treinamento e não generalize de forma adequada.

### 3.2.1.2 Aleatorização do Experimento

O experimento foi conduzido segundo a Tabela 3.2, definida a partir do programa Minitab versão 13. A Tabela 3.2 mostra o resultado de treinamento e teste da rede neural descrita na seção 3.1, com três conjuntos distintos de dados, Bens de Consumo Duráveis, Bens de Consumo Não Duráveis e Bens Intermediários. Após o treinamento o erro foi calculado prevendo-se os valores de vendas para dois meses consecutivos e comparando-se o valor previsto com o valor real das vendas agregadas para cada conjunto de indicadores industriais de vendas utilizados. Para cada conjunto de dados e configuração de treinamento foram realizados três experimentos e o valor médio do erro absoluto foi calculado e planilhado na Tabela 3.2.

Na Tabela 3.2 o valor Diff\_BCD é o erro da rede neural para os dados de Bens de Consumo Duráveis, Diff\_BCND é o erro da rede neural para os dados de Bens de Consumo Não Duráveis e por ultimo Diff\_BI é o erro da rede para os dados de Bens Intermediários. Os experimentos foram executados conforme indicado na coluna Run Order. As colunas Function, Hidden Layer e Error indicam as configurações das variáveis controladas em cada das experimento realizado.

StdOrder	RunOrder	CenterPt	Blocks	Function	Hidden Layer	Error	Diff_BCD	Diff_BCND	Diff_BI
5	1	1	1	LogSig	15	0,0010	0,0800	0,1611	0,1357
7	2	1	1	LogSig	25	0,0010	0,4326	0,2140	0,2123
6	3	1	1	Purelin	15	0,0010	0,3399	0,3509	0,2748
3	4	1	1	LogSig	25	0,0001	0,4538	0,1113	0,0962
2	5	1	1	Purelin	15	0,0001	0,3470	0,2546	0,1734
4	6	1	1	Purelin	25	0,0001	1,1723	1,9680	0,7115
8	7	1	1	Purelin	25	0,0010	1,0769	1,2560	0,7400
1	8	1	1	LogSig	15	0,0001	0,0800	0,1611	0,1357

Tabela 3.2 – Resultados obtidos durante os experimentos da rede neural.

### 3.2.1.3 Caracterização das Variáveis mais Importantes

Para caracterização das variáveis mais importantes do processo de treinamento da rede neural descrita, construíram-se os gráficos com efeitos principais para cada um dos conjuntos de dados utilizados para treinamento e testes representados pelas Figuras 3.4, 3.5 e 3.6.

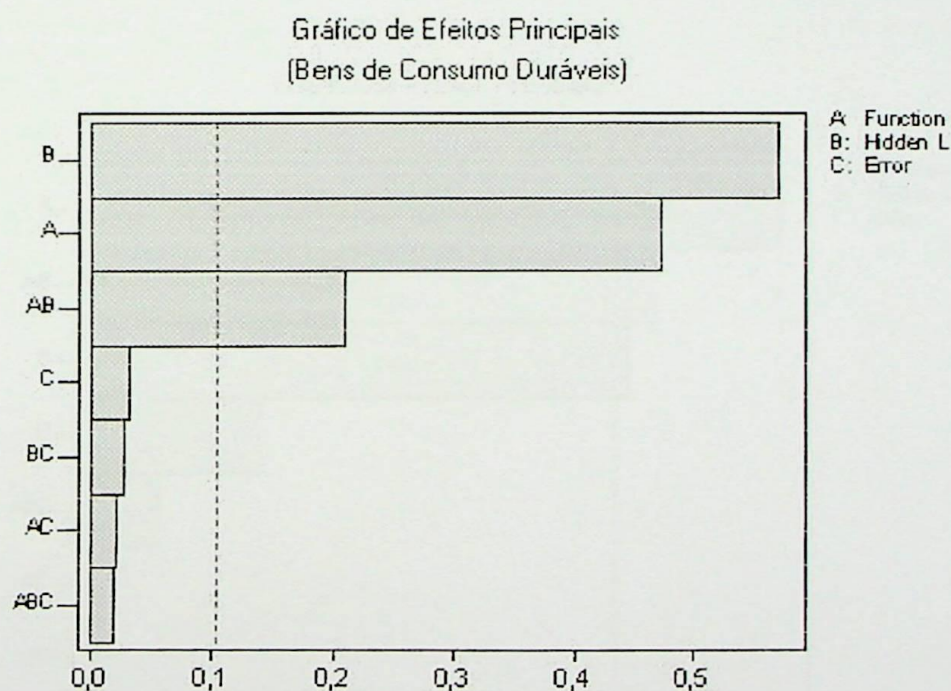


Figura 3.4 – Gráfico de Efeitos mais Importantes para Bens de Consumo Duráveis

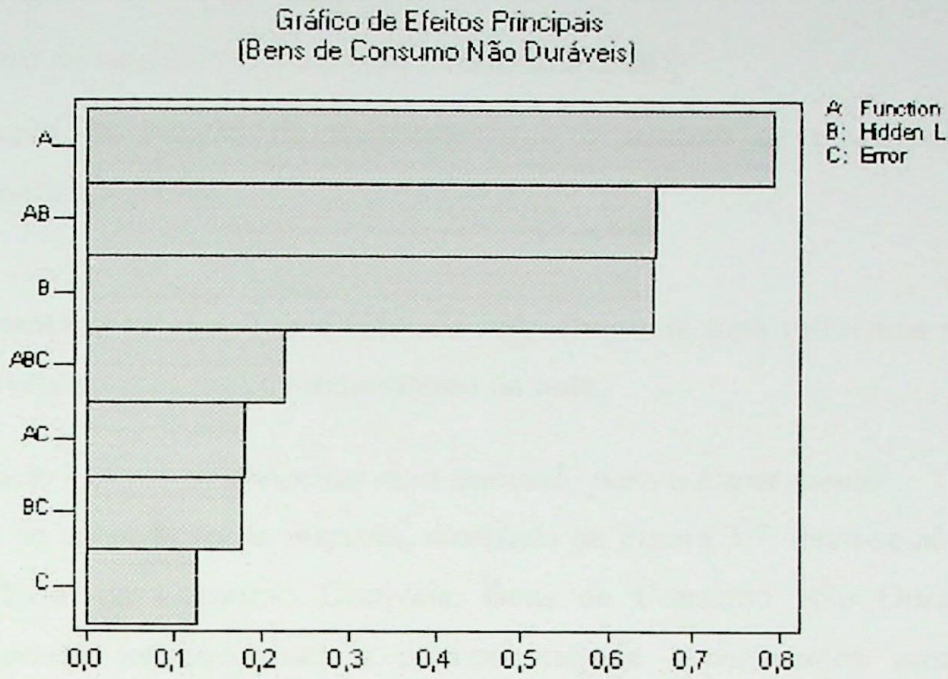


Figura 3.5 – Gráfico de Efeitos Principais para Bens de Consumo Não Duráveis

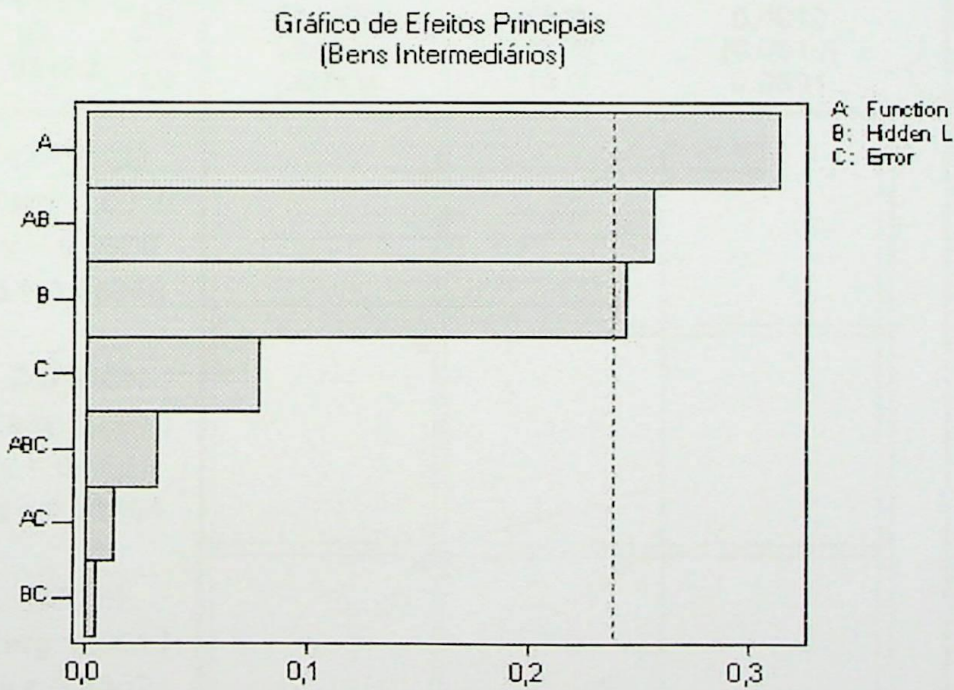


Figura 3.6 – Gráfico de Efeitos Principais para bens Intermediários

Os gráficos de efeitos principais obtidos através da análise no Minitab versão 13, Figuras 3.4, 3.5 e 3.6, mostram que para as três séries de dados estudadas, Bens de Consumo Duráveis, Bens de Consumo Não Duráveis e Intermediários os três fatores principais do desempenho do treinamento da rede neural são :

- Função de transferência (A);
- Número de neurônios na camada intermediária (B);
- Interação da Função de transferência e o número de neurônios da camada intermediária (A.B);

É interessante notar que o erro total da rede apresenta uma influência menor que as duas outras variáveis no processo de treinamento da rede.

3.2.1.4 Otimização - Ponto de Resposta mais indicado para o Experimento

A análise de otimização da resposta, mostrada na Figura 3.7, deve-se ao fato das três séries estudadas (Bens de Consumo Duráveis, Bens de Consumo Não Duráveis e Bens Intermediários) serem independentes e provenientes de experimentos econômicos não controlados distintos.

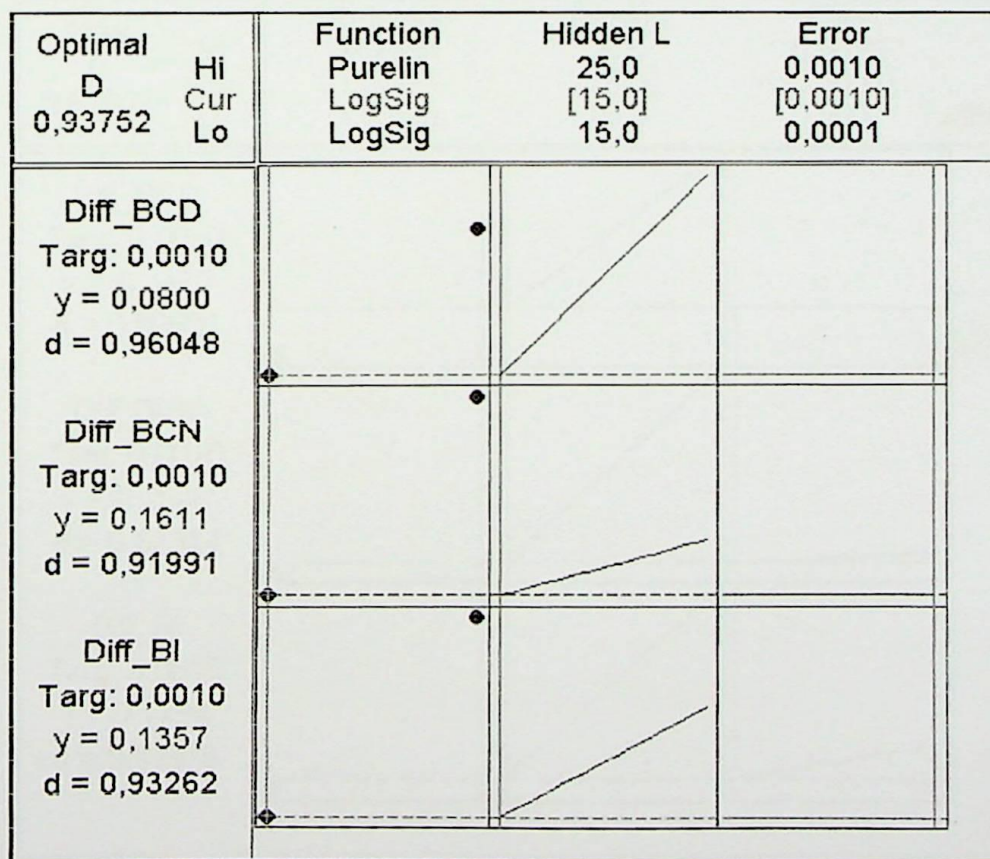


Figura 3.7 – Análise de otimização da resposta, ponto ótimo

A análise do gráfico acima, obtido através do Optimizer Responser do software Minitab versão 13, mostra que a Desejabilidade, ponto ótimo para as variáveis do processo é de 0,93752, para a configuração abaixo :

Característica	Ponto
Função	Sigmoidal
Neurônios da Camada Intermediária	15
Erro Total de Treinamento	0,0001

Tabela 3.3 – Resultados ótimos do Projeto de Experimentos

A análise no Minitab das variáveis também indica que a modificação do Erro Total de treinamento, nas condições ótimas para a função de transferência e do número de neurônios da camada intermediária, de 0,0001 para 0,001 não modifica a Desejabilidade enquanto a variação da função de transferência de sigmoidal para linear diminui a Desejabilidade para 0,87054, conforme mostrado na Figura 3.8, e a variação do número de neurônios para 25 diminui a Desejabilidade para 0,88621, conforme mostrado na Figura 3.9, confirmando a análise anterior de que o erro total influencia em menor escala os resultados do treinamento que as outras duas variáveis.

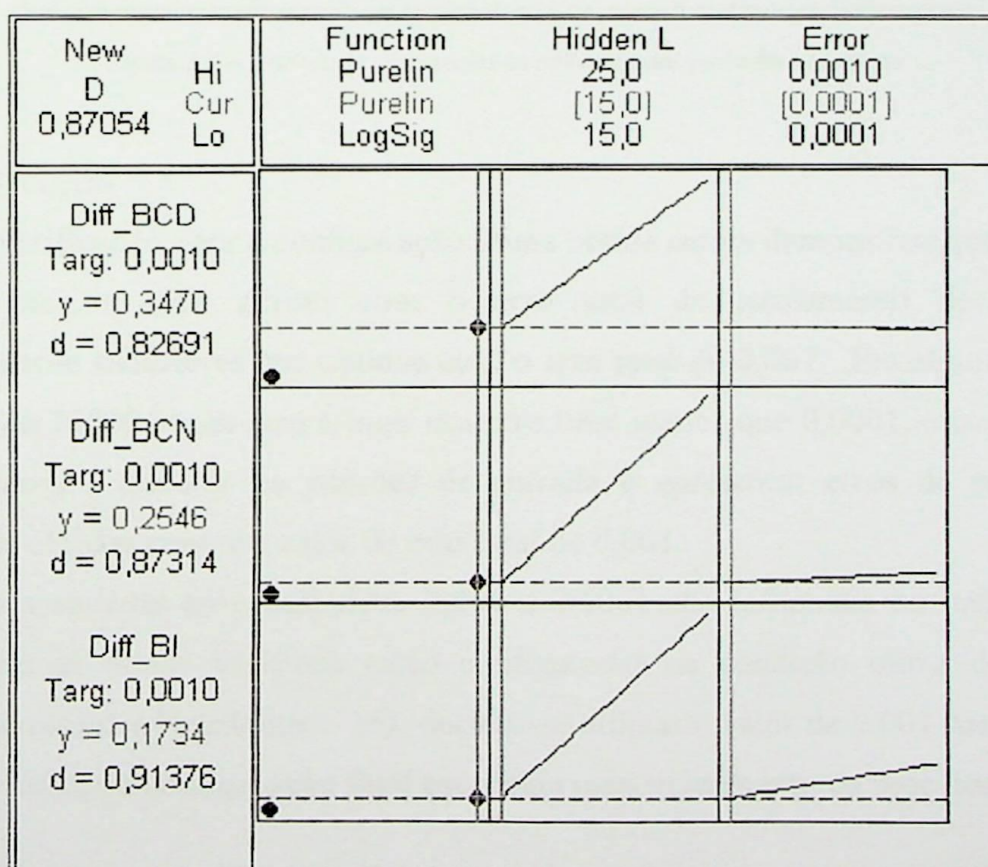


Figura 3.8 – Análise de otimização da resposta para situação fora do ponto ótimo

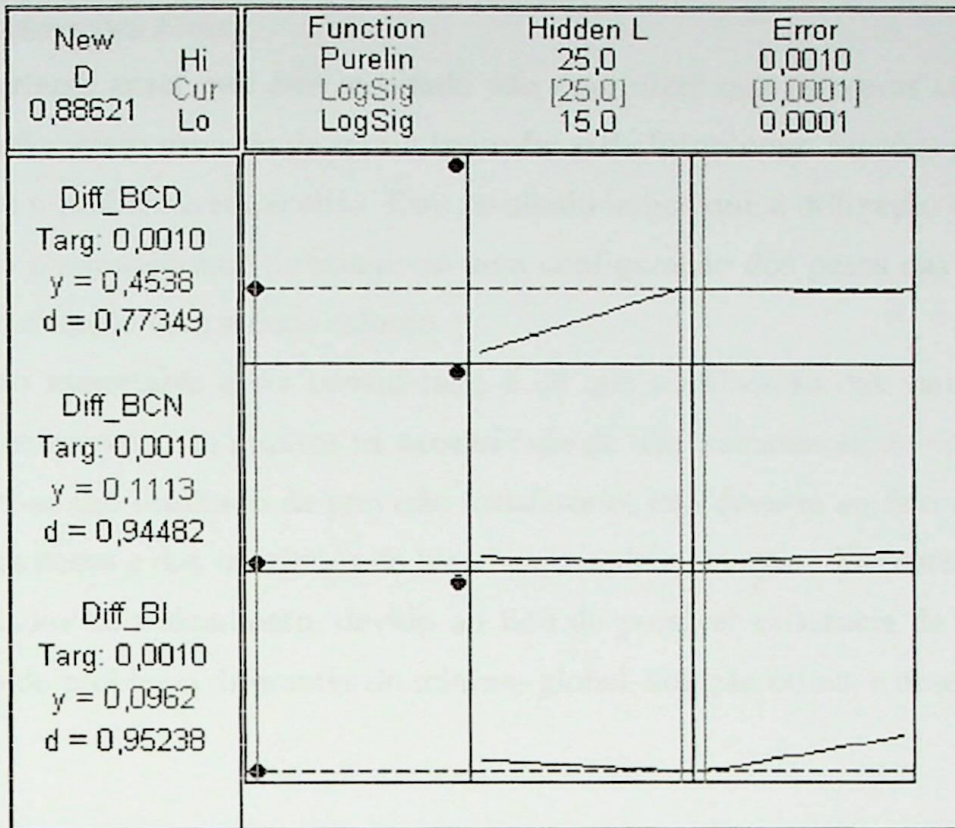


Figura 3.9 – Análise de otimização da resposta para ponto fora do ótimo

### 3.2.1.5 Verificação

Uma verificação para a configuração ótima obtida acima demonstrou que os tempos de treinamento para a rede neural com o erro total de treinamento de 0,0001 eram significativamente superiores aos obtidos com o erro total de 0,001. Em alguns casos a rede treinou mais de 10000 vezes sem atingir um erro total menor que 0,0001, o que levou a rede em alguns casos a decorar os padrões de entrada e apresentar erros de previsão muito superiores aos obtidos com um valor de erro total de 0,001.

Como a variável erro total entre 0,001 e 0,0001 não influencia no resultado da rede neural, quando as outras variáveis estão configuradas na condição ótima de treinamento (Função = Sigmoidal e Neurônios = 15), decidiu-se utilizar o valor de 0,001 para treinamento. A Tabela 3.4 mostra a configuração final escolhida para treinamento da rede neural.

Característica	Ponto
Função	Sigmoidal
Neurônios da Camada Intermediária	15
Erro Total de Treinamento	0,001

Tabela 3.4 – Configuração escolhida para treinamento da Rede Neural

### *3.2.1.6 Considerações Finais*

É importante notar que este resultado não quer dizer que as outras configurações de redes neurais não sejam capazes de serem treinadas e de fornecerem funções de aproximação adequadas para o problema em análise. Este resultado indica que a utilização da configuração acima aumenta a possibilidade de atingir-se uma configuração dos pesos das sinapses para a função de aproximação com menos esforço.

Um fato importante a ser considerado é de que a utilização das variáveis conforme projeto de experimentos não implica na necessidade de não treinamento da rede neural várias vezes até obter-se um resultado de previsão satisfatório, isto deve-se ao fato de que variação dos valores dos pesos e dos neurônios de bias leva a soluções diferentes para o mesmo conjunto de dados de treinamento, devido ao fato de possível existência de mínimos locais para a solução do problema diferentes do mínimo global, solução ótima, e desejada.

## **4 RESULTADOS OBTIDOS**

### **4.1 Ambiente de Implementação**

O sistema de previsão de vendas foi implementado em um computador padrão IBM PC Pentium de 233 MHz com sistema operacional Windows 98.

Todo o software de treinamento e teste da rede neural foi desenvolvido em MATLAB versão 5.1. A análise de regressão múltipla foi realizada utilizando-se o software Minitab versão 13.

### **4.2 Características da Base de Dados Utilizada**

Os dados utilizados de indicadores econômicos foram coletados das revistas Conjuntura Econômica, da Fundação Getúlio Vargas, dos meses de outubro de 2000 e Maio de 2001.

Para obter-se uma visão da abrangência do modelo proposto dividiu-se a base de dados em três grupos :

- Bens de Consumo Duráveis, Tabela 6.1 apêndice I;
- Bens de Consumo Não Duráveis, Tabela 6.2 apêndice I;
- Bens de Consumo Intermediários, Tabela 6.3 apêndice I;

Estes dados formam uma base de dados de Setembro de 1998 à Março de 2001 e estão disponíveis no apêndice I.

### **4.3 Avaliação do Modelo Proposto**

De forma a avaliar o modelo proposto comparou-se os resultados das previsões por redes neurais com os resultados do método de regressão linear múltipla.

### **4.4 Conjuntos de Treino e Teste**

Para a avaliação do modelo proposto, cada conjunto da base de dados foi dividido em dois conjuntos: um para treinamento da rede neural e outro para avaliação do desempenho da mesma. Os conjuntos de treino e teste estão descritos no Anexo I.

Após o treinamento da rede neural com os dados de setembro de 1998 a dezembro de 2000 obteve-se a previsão para os meses de janeiro e fevereiro de 2001, de forma a testar os resultados do modelo.

Neste período a economia Brasileira apresentou algumas características particulares as quais devem ser conhecidas:

- Desvalorização do real frente ao dólar em dezembro de 1998, acarretando forte elevação da taxa de juros, (Lanzana, 2001).
- Inflação sobre controle e não superior a dois dígitos, (Lanzana, 2001).
- Estabilidade das políticas monetária e fiscal, influenciadas pela manutenção do Presidente da República e sua equipe econômica, (Lanzana, 2001).

Um dos objetivos deste trabalho é que a rede neural aprenda o comportamento da resposta do mercado a estas situações de natureza política e econômica e seja capaz de estimar os resultados econômicos em novas situações como as descritas acima.

## 4.5 Resultados Obtidos

As Tabelas 4.1, 4.2 e 4.3 apresentam os resultados obtidos pelo modelo proposto (Redes Neurais) e os obtidos pela análise de regressão múltipla fornecida pelo software Minitab 13.

Para o cálculo da equação de regressão foram utilizados os mesmos dados de treinamento da rede neural :

- Mês da previsão;
- Venda do mês anterior (V-1);
- Venda de dois meses atrás (V-2);
- Venda de três meses atrás (V-3);
- Taxa de juros do mês anterior;
- Taxa de inflação do mês anterior;
- Taxa de desemprego do mês anterior;

Os resultados completos da análise de regressão múltipla para os três conjuntos de dados utilizados se encontra no apêndice III. A seguir apresenta-se um sumário destes resultados:

Os valores encontrados para o teste de Fisher  $F = 4,23$  (Bens de Consumo Duráveis);  $F = 4,16$  (Bens de Consumo Não Duráveis); e  $F = 4,42$  (Bens Intermediários), maiores que  $F_c(6,17) = 2,70$ , indicam que pelo menos uma das variáveis independentes, para os três conjuntos de dados em análise, explica a variável dependente (Venda do próximo período).

O Teste de Durbin-Watson,  $DW = 2,12$  (Bens de Consumo Duráveis);  $DW = 1,96$  (Bens de Consumo Não Duráveis); e  $DW = 2,28$  (Bens Intermediários), mostra a ausência de correlação serial, estando todos os valores próximos de seu valor ideal 2.

Os gráficos de resíduos, mostrados no apêndice I, indicam uma distribuição gaussiana para todas as séries temporais o que permite utilizar os resultados da análise de regressão múltipla com confiança.

#### 4.5.1 Bens de Consumo Duráveis

Mês/Ano	Vendas Reais	Previsão RNA	Erro RNA	Previsão ARM	Erro ARM
jan/01	137,5	172,1	34,6	147,3	9,8
Fev/01	143,2	141,4	1,8	139,9	3,3

Tabela 4.1 – Resultados para Bens de Consumo Duráveis

A Tabela 4.1 mostra que a rede neural obteve uma previsão para o mês de janeiro de 2001 (172,1) pior que a obtida pelo modelo de regressão múltipla (147,3) para um valor real de 137,5. Este fato deve-se provavelmente a crise que afetou o Brasil a partir de dezembro de 1998, crise asiática, que resultou na elevação acentuada dos juros e de forte desvalorização cambial. Neste período as vendas do mês de janeiro foram superiores as de dezembro, fato que influenciou nos resultados obtidos pela rede neural e mostra sua capacidade de generalizar os dados de treinamento.

#### 4.5.2 Bens de Consumo Não Duráveis

Mês/Ano	Vendas Reais	Previsão RNA	Erro RNA	Previsão ARM	Erro ARM
Jan/01	101,7	103,3	1,6	93,75	7,95
Fev/01	91,8	98,1	6,3	96,65	4,85

Tabela 4.2 – Resultados para Bens de Consumo Não Duráveis

A Tabela 4.2 mostra que a rede neural obteve uma boa previsão para os valores de vendas dos meses de janeiro e fevereiro de 2001, ligeiramente superior as obtidas pelo modelo de regressão múltipla. Neste conjunto de dados não se verifica o mesmo aumento das vendas no mês de janeiro, verificado nos dados para os Bens de Consumo Duráveis. Isto se deve ao fato de que os Bens de Consumo Não Duráveis possuem preços inferiores aos dos Bens de Consumo Duráveis e contemplam uma grande gama de bens de primeira necessidade que não podem ser cortados, tais como alimentos, vestuário e etc.

### 4.5.3 Bens de Consumo Intermediários

Mês/Ano	Vendas Reais	Previsão RNA	Erro RNA	Previsão ARM	Erro ARM
jan/01	123,7	134,9	11,2	110,13	13,57
fev/01	118,7	107,8	10,9	111,32	7,38

Tabela 4.3 – Resultados para Bens Intermediários

A Tabela 4.3 mostra que a rede neural obteve uma boa previsão para os valores de vendas dos meses de janeiro e fevereiro de 2001, quando comparadas as obtidas pelo modelo de regressão múltipla. Neste conjunto de dados não se verifica o mesmo aumento acentuado das vendas no mês de janeiro, verificado nos dados para os Bens de Consumo Duráveis.



## **5 CONCLUSÕES**

### **5.1 Introdução**

O presente capítulo revisa os importantes pontos que nortearam os rumos dessa pesquisa. Na seção 5.2 o desenvolvimento do trabalho com suas contribuições ao longo dos capítulos é sumarizado. A seção 5.3 apresenta algumas discussões e limitações do sobre o trabalho. Na seção 5.4 as principais conclusões e contribuições são descritas. Na seção 5.5 prováveis desenvolvimentos futuros são apresentados.

### **5.2 Desenvolvimento e Contribuições**

#### **5.2.1 Introdução**

O principal problema abordado nessa pesquisa consistiu na modelagem e implementação de uma nova metodologia para previsão de indicadores industriais, baseado nas vendas de períodos anteriores e em outras variáveis macroeconômicas. Notadamente esta metodologia é uma ferramenta de apoio à decisão gerencial bastante promissora, visto que erros de previsão podem acarretar sérios problemas as empresas.

Foi assim desenvolvido um sistema de previsão de indicadores industriais de vendas baseado nas vendas dos três meses anteriores e nos indicadores macroeconômicos de taxa de juros, inflação, taxa de desemprego e mês da previsão.

O modelo proposto foi treinado para três diferentes séries de dados reais abrangendo grupos de produtos representativos de parcelas significativas do mercado. As redes neurais treinadas para estes três grupos apresentaram desempenho satisfatório indicando sua capacidade de generalização para a análise de diversos segmentos industriais.

#### **5.2.2 Revisão**

Foram abordados os principais aspectos das variações da economia em função das políticas governamentais e das pressões naturais dos mercados. Foi abordada a influencia de algumas políticas governamentais no controle dos ciclos de expansão e retração econômica.

Foi também descrito o modelo econométrico de análise de regressão múltipla, como uma das mais poderosas ferramentas para a análise e previsão de dados econômicos. Este modelo possui a peculiaridade de poder prever o comportamento de uma variável através da análise de inúmeros fatores de entrada, o que o torna um método bastante apropriado para a comparação de resultados com o modelo proposto. De forma a tornar consistente a

comparação dos resultados o modelo de regressão múltipla foi utilizado com as mesmas configurações de variáveis de entrada (Vendas anteriores e Indicadores macroeconômicos) e Saída (Previsão do Indicador Industrial em Questão) e modelado com o mesmo conjunto de treino que a rede neural proposta.

Uma revisão da técnica de redes neurais, seus principais usos e limitações foram descritos assim como foram estudados diversos casos de aplicação desta técnica em problemas reais.

Por fim uma revisão da técnica de projeto de experimentos foi realizada de forma a determinar quais os parâmetros ótimos de treinamento para a rede neural proposta.

### **5.2.3 Definição do Modelo Proposto**

Em função dos dados coletados e da revisão da literatura foi proposto um modelo de previsão de indicadores industriais de vendas que toma como parâmetros de entrada além dos valores passados do indicador que se quer prever, diversas variáveis macroeconômicas, representativas das principais políticas governamentais.

Foram estudadas e definidas sete variáveis de entrada para a rede neural e cada uma das variáveis foi analisada quanto a sua capacidade de afetar o indicador que se quer prever.

Através de um projeto de experimento foi determinada a configuração de treinamento da rede neural, incluindo número de neurônios da camada intermediária, o erro total de treinamento e a função de transferência da camada de saída. A verificação prática do resultado desta análise levou a definição do erro total de treinamento para um patamar diferente do descrito como ponto ótimo pelo projeto de experimento realizado. Como produto deste trabalho foi elaborado um artigo a ser submetido a um próximo congresso.

### **5.2.4 Resultados Obtidos**

No capítulo 4 foram comparados os valores obtidos pelo modelo proposto de rede neural contra os obtidos através da análise de regressão múltipla. A análise destes resultados não evidenciou superioridade do método proposto e os resultados se revelaram compatíveis.

## **5.3 Discussões**

No intuito de estabelecer um estudo comparativo entre os métodos utilizados de previsão, análise de regressão múltipla e redes neurais artificiais, algumas vantagens e desvantagens podem aqui ser relacionadas.

### 5.3.1 Redes Neurais

#### Vantagens

- Erro de previsão compatível;
- Menor dependência do econometrista;
- Existem poucas restrições à morfologia dos dados;
- Rapidez nos cálculos;
- Pode freqüentemente funcionar como uma “caixa preta”;
- Não necessidade de definição da variável endógena

#### Desvantagens

- Pode freqüentemente funcionar como uma “caixa preta”;
- Não há, freqüentemente, um controle gráfico das variáveis;
- Não há, freqüentemente, intervalos de confiança para estimativas

### 5.3.2 Modelo Econométrico

#### Vantagens

- A equação do modelo é obtida;
- Os intervalos de confiança das estimativas são fornecidos;
- Há um acompanhamento gráfico das variáveis modeladas;
- Rapidez nos cálculos;

#### Desvantagens

- Existem inúmeras condições iniciais que devem ser observadas;
- Existe um grande número de testes de hipóteses;
- Pode freqüentemente funcionar como uma “caixa preta”;

A rede perceptron multicamada com treinamento por retropropagação de erro obteve resultados compatíveis com os obtidos através da regressão múltipla. A definição dos parâmetros através do projeto de experimentos se mostrou bastante satisfatória e o tempo de treinamento requerido para um resultado adequado foi mínimo. No terceiro treinamento a rede obteve respostas bastante satisfatórias.

O trabalho mostrou a viabilidade da utilização da técnica de redes neurais à previsão de indicadores industriais de vendas. É esperado que os resultados desta técnica sejam

substancialmente melhores a medida que o conjunto de treinamento aumente incluindo novos padrões de comportamento do mercado à variações das políticas do governo.

### 5.3.3 Limitações do trabalho

Existem algumas limitações no trabalho que podem e devem ser superadas futuramente. Considerando as redes neurais percebeu-se o seguinte :

- A performance da rede está sujeita a uma boa seleção do conjunto de treinamento;
- O desempenho da rede é limitado quando alguns padrões estão contaminados com um alto nível de ruído;

Com relação a base de dados utilizada, deve-se :

- Aumentar o conjunto de treinamento com informações de mais anos, de forma a mostrar a rede novas situações econômicas e suas repercussões no mercado;
- Incluir dados relativos as políticas de comércio exterior e fiscal, de forma a ensinar a rede as respostas do mercado a variações nestas políticas;

## 5.4 Conclusões

A pesquisa, em uma conclusão geral, mostrou que :

- Os resultados obtidos pelo modelo proposto e pelo processo de análise de regressão múltipla foram compatíveis;
- A técnica de projeto de experimentos se mostrou eficiente na determinação da configuração ótima de treinamento
- O modelo de previsão de indicadores industriais de vendas utilizando rede neurais foi capaz de satisfazer as principais necessidades de um sistema previsor;

Os resultados obtidos foram estimulantes e a metodologia mostrou um grande poder de generalização, ajustando-se e obtendo valores de previsão para diferentes segmentos industriais de maneira apropriada o que indica a possibilidade de utilização da técnica de redes neurais no campo da previsão como forma alternativa ou complementar as atuais técnicas baseadas em séries temporais.

As maiores contribuições desta pesquisa são sumarizadas a seguir :

- Modelamento de um previsor de indicadores industriais utilizando-se da técnica de redes neurais artificiais;
- Modelamento de um projeto de experimentos para determinação da configuração ótima de treinamento para redes neurais artificiais;
- Desenvolvimento de algoritmo para reconhecimento de padrões de indicadores industriais utilizando-se uma rede neural perceptron multicamadas;
- Determinação de fontes para coleta de dados para treinamento e teste do modelo proposto;
- Verificação do nível de adequação do modelo para explicar a variável dependente, indicador industrial que se quer prever;
- Comparação dos resultados obtidos entre o modelo proposto e o modelo de análise de regressão múltipla;
- Indicação de melhorias no modelo proposto de forma a aumentar a capacidade do mesmo em generalizar padrões de resposta do mercado a políticas governamentais e comportamentos dos agentes econômicos;

## 5.5 Desenvolvimentos Futuros

É natural a expansão do modelo com a inclusão de indicadores macroeconômicos para as políticas fiscal (por exemplo :  $\$Impostos/\$PIB$ ) e de comércio exterior (por exemplo : taxa dólar) o que também deve trazer maior capacidade de generalização ao modelo.

Um outro campo de estudo futuro é pré-processamento dos dados de entrada através da análise de defasagens distribuídas para determinação do melhor número de meses anteriores a serem utilizados no modelo tanto para as vendas agregadas como também para as outras variáveis macroeconômicas (taxa de juros, taxa de desemprego, taxa de inflação).

## 6 APÊNDICE I – TABELAS DE TREINO E TESTE

### 6.1 Bens de Consumo Duráveis

Mês/Ano	Vendas	Mês	Tx Juros SELIC	Tx Inflação	Taxa Desemprego	Vendas-1	Vendas-2	Vendas-3
<b>CONJUNTO DE TREINO</b>								
Set/98	152	9						
Out/98	135,6	10	34,29	-0,08	7,7	152		
Nov/98	130,5	11	41,6	0,08	7,5	135,6	152	
Dez/98	101,8	12	38,73	-0,32	7	130,5	135,6	152
Jan/99	104	1	31,24	0,45	6,3	101,8	130,5	135,6
Fev/99	92,4	2	29,5	0,84	7,7	104	101,8	130,5
Mar/99	125,8	3	38,97	3,61	7,5	92,4	104	101,8
Abr/99	120,1	4	43,25	2,83	8,2	125,8	92,4	104
Mai/99	123,6	5	36,12	0,71	8	120,1	125,8	92,4
Jun/99	121,3	6	27,11	-0,29	7,7	123,6	120,1	125,8
Jul/99	122	7	22,01	0,36	7,8	121,3	123,6	120,1
Ago/99	141,8	8	20,74	1,55	7,5	122	121,3	123,6
Set/99	146,2	9	19,5	1,56	7,7	141,8	122	121,3
Out/99	140	10	19,38	1,45	7,4	146,2	141,8	122
Nov/99	145,2	11	18,91	1,7	7,5	140	146,2	141,8
Dez/99	108,9	12	18,93	2,39	7,3	145,2	140	146,2
Jan/00	114,1	1	19,03	1,81	6,3	108,9	145,2	140
Fev/00	138,4	2	18,94	1,24	7,6	114,1	108,9	145,2
Mar/00	139,6	3	18,87	0,35	8,2	138,4	114,1	108,9
Abr/00	138,9	4	18,85	0,15	8,1	139,6	138,4	114,1
Mai/00	156,2	5	18,62	0,23	7,8	138,9	139,6	138,4
Jun/00	148,8	6	18,51	0,31	7,8	156,2	138,9	139,6
Jul/00	149,6	7	18,04	0,85	7,4	148,8	156,2	138,9
Ago/00	170,4	8	16,85	1,57	7,2	149,6	148,8	156,2
Set/00	151,3	9	16,52	2,39	7,1	170,4	149,6	148,8
Out/00	168,1	10	16,56	1,16	6,7	151,3	170,4	149,6
Nov/00	173,1	11	16,56	0,38	6,8	168,1	151,3	170,4
Dez/00	152,7	12	16,51	0,29	6,2	173,1	168,1	151,3
<b>CONJUNTO DE TESTES</b>								
Dez/00	152,7	12	16,51	0,29	6,2	173,1	168,1	151,3
Jan/01	137,5	1	16,19	0,63	4,8	152,7	173,1	168,1
Fev/01	143,2	2	15,52	0,62	5,7	137,5	152,7	173,1
Mar/01			15,2	0,23	5,7	142,2	137,5	152,7

Tabela 6.1 – Bens de Consumo Duráveis, Fonte – Revista Conjuntura Econômica, FGV, Outubro/2000 e Março/2001

## 6.2 Bens de Consumo Não-Duráveis

Mês/Ano	Vendas	Mês	Tx Juros SELIC	Tx Inflação	Taxa Desemprego	Vendas-1	Vendas-2	Vendas-3
CONJUNTO DE TREINO								
Set/98	123	9						
Out/98	125,4	10	34,29	-0,08	7,7	123		
Nov/98	122,3	11	41,6	0,08	7,5	125,4	123	
Dez/98	105,3	12	38,73	-0,32	7	122,3	125,4	123
Jan/99	93,3	1	31,24	0,45	6,3	105,3	122,3	125,4
Fev/99	88,4	2	29,5	0,84	7,7	93,3	105,3	122,3
Mar/99	107,6	3	38,97	3,61	7,5	88,4	93,3	105,3
Abr/99	103,1	4	43,25	2,83	8,2	107,6	88,4	93,3
Mai/99	110,1	5	36,12	0,71	8	103,1	107,6	88,4
Jun/99	112,6	6	27,11	-0,29	7,7	110,1	103,1	107,6
Jul/99	119,2	7	22,01	0,36	7,8	112,6	110,1	103,1
Ago/99	122,5	8	20,74	1,55	7,5	119,2	112,6	110,1
Set/99	122,4	9	19,5	1,56	7,7	122,5	119,2	112,6
Out/99	123,7	10	19,38	1,45	7,4	122,4	122,5	119,2
Nov/99	120,1	11	18,91	1,7	7,5	123,7	122,4	122,5
Dez/99	106,3	12	18,93	2,39	7,3	120,1	123,7	122,4
Jan/00	94,1	1	19,03	1,81	6,3	106,3	120,1	123,7
Fev/00	98,4	2	18,94	1,24	7,6	94,1	106,3	120,1
Mar/00	103,1	3	18,87	0,35	8,2	98,4	94,1	106,3
Abr/00	96	4	18,85	0,15	8,1	103,1	98,4	94,1
Mai/00	107,9	5	18,62	0,23	7,8	96	103,1	98,4
Jun/00	110	6	18,51	0,31	7,8	107,9	96	103,1
Jul/00	115,7	7	18,04	0,85	7,4	110	107,9	96
Ago/00	122,6	8	16,85	1,57	7,2	115,7	110	107,9
Set/00	115,3	9	16,52	2,39	7,1	122,6	115,7	110
Out/00	126	10	16,56	1,16	6,7	115,3	122,6	115,7
Nov/00	120,7	11	16,56	0,38	6,8	126	115,3	122,6
Dez/00	106,5	12	16,51	0,29	6,2	120,7	126	115,3
CONJUNTO DE TESTES								
Dez/00	106,5	12	16,51	0,29	6,2	120,7	126	115,3
Jan/01	101,7	1	16,19	0,63	4,8	106,5	120,7	126
Fev/01	91,8	2	15,52	0,62	5,7	101,7	106,5	120,7
Mar/01		3	15,2	0,23	5,7	91,8	101,7	106,5

Tabela 6.2 – Bens de Consumo Não Duráveis, Fonte – Revista Conjuntura Econômica, FGV, Outubro/2000 e Março/2001

### 6.3 Bens de Consumo Intermediários

Mês/Ano	Vendas	Mês	Tx Juros SELIC	Tx Inflação	Taxa Desemprego	Vendas-1	Vendas-2	Vendas-3
CONJUNTO DE TREINO								
Set/98	122,2	9						
Out/98	121,6	10	34,29	-0,08	7,7	122,2		
Nov/98	116,6	11	41,6	0,08	7,5	121,6	122,2	
Dez/98	103,8	12	38,73	-0,32	7	116,6	121,6	122,2
Jan/99	105,9	1	31,24	0,45	6,3	103,8	116,6	121,6
Fev/99	101,4	2	29,5	0,84	7,7	105,9	103,8	116,6
Mar/99	119,1	3	38,97	3,61	7,5	101,4	105,9	103,8
Abr/99	115,9	4	43,25	2,83	8,2	119,1	101,4	105,9
Mai/99	123,4	5	36,12	0,71	8	115,9	119,1	101,4
Jun/99	121,9	6	27,11	-0,29	7,7	123,4	115,9	119,1
Jul/99	124,5	7	22,01	0,36	7,8	121,9	123,4	115,9
Ago/99	128,2	8	20,74	1,55	7,5	124,5	121,9	123,4
Set/99	125,3	9	19,5	1,56	7,7	128,2	124,5	121,9
Out/99	128,7	10	19,38	1,45	7,4	125,3	128,2	124,5
Nov/99	125,5	11	18,91	1,7	7,5	128,7	125,3	128,2
Dez/99	115,2	12	18,93	2,39	7,3	125,5	128,7	125,3
Jan/00	113,0	1	19,03	1,81	6,3	115,2	125,5	128,7
Fev/00	116,1	2	18,94	1,24	7,6	113,0	115,2	125,5
Mar/00	126,3	3	18,87	0,35	8,2	116,1	113,0	115,2
Abr/00	112,7	4	18,85	0,15	8,1	126,3	116,1	113,0
Mai/00	131,8	5	18,62	0,23	7,8	112,7	126,3	116,1
Jun/00	132,2	6	18,51	0,31	7,8	131,8	112,7	126,3
Jul/00	134,4	7	18,04	0,85	7,4	132,2	131,8	112,7
Ago/00	137,3	8	16,85	1,57	7,2	134,4	132,2	131,8
Set/00	130,6	9	16,52	2,39	7,1	137,3	134,4	132,2
Out/00	137	10	16,56	1,16	6,7	130,6	137,3	134,4
Nov/00	129,6	11	16,56	0,38	6,8	137	130,6	137,3
Dez/00	121,7	12	16,51	0,29	6,2	129,6	137	130,6
CONJUNTO DE TESTES								
Dez/00	121,7	12	16,51	0,29	6,2	129,6	137	130,6
Jan/01	123,7	1	16,19	0,63	4,8	121,7	129,6	137
Fev/01	118,7	2	15,52	0,62	5,7	123,7	121,7	129,6
Mar/01		3	15,2	0,23	5,7	118,7	123,7	121,7

Tabela 6.3 – Bens Intermediários, Fonte – Revista Conjuntura Econômica, FGV, Outubro/2000 e Março/2001

## 7 APÊNDICE II – CÓDIGO DE TREINAMENTO E TESTES DA REDE NEURAL

```

%-----
% Programa para Treinamento e Testes para Rede Neural Multicamadas
% MatLab Versão 5.1
% Autor : Leandro R. dos Santos
% Revisão : 1.0
% Data : 03/09/2001
%
% Este programa deve ser executado após a geração dos arquivos com os padrões de
% treinamento para classificação off-line das vendas reais para Bens de Consumo
% Duráveis.
% Os pesos e Biais da Rede são armazenados para uso posterior.
%-----

% Leitura dos dados de entrada para treinamento da RN, arquivo ESBCD.TXT
load esbcd.txt
EnTrain = esbcd'

% Leitura dos dados de saída para treinamento da RN, arquivo SSBOD.TXT
load ssbcd.txt
SnTrain = ssbcd'

% Inicialização do Vetor de máximos e mínimos para cada variável de entrada
[r,c] = size(EnTrain)
for i = 1:r
P(i,1) = min(EnTrain(i,:))
P(i,2) = max(EnTrain(i,:))
end

% Inicializacao dos pesos da rede neural
Nnci = 15 % Número de neurônios da Camada Intermediária
[W1, B1, W2, B2] = initff(P, Nnci, 'logsig', 1, 'logsig')

% Treinamento da rede neural
disp_freq = 100
max_epoch = 10000
err_goal = 0.001 % Erro Total da rede neural
tp = [disp_freq max_epoch err_goal]
[W1, B1, W2, B2] = trainbpx(W1, B1, 'logsig', W2, B2, 'logsig', EnTrain, SnTrain, tp)

% Teste da rede neural

% Leitura dos valores para testes Rede Neural
load tsbcd.txt

Px1 = tsbcd(1,:)'
A1 = simuff(Px1, W1, B1, 'logsig', W2, B2, 'logsig')

Px2 = tsbcd(2,:)'
A2 = simuff(Px2, W1, B1, 'logsig', W2, B2, 'logsig')

Px3 = tsbcd(3,:)'
A3 = simuff(Px3, W1, B1, 'logsig', W2, B2, 'logsig')

% Arquivo com pesos e bias da rede neural treinada
save BCD_W1.txt W1 -ascii
save BCD_B1.txt B1 -ascii
save BCD_W2.txt W2 -ascii
save BCD_B2.txt B2 -ascii

% Leitura dos valores reais, para testes
load msbcd.txt
msbcd = msbcd'
msa = [ A1 A2 A3]

% Calculo do erro médio
Erro = abs(A1 - msbcd(1)) + abs(A2 - msbcd(2)) + abs(A3 - msbcd(3))

```

## 8 APÊNDICE III – RESULTADOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA

### 8.1 Bens de Consumo Duráveis

Regression Analysis: Vendas versus Mês; Taxa Juros; ...

The regression equation is

$$\text{Vendas} = -42 - 1,25 \text{ Mês} - 0,637 \text{ Taxa Juros} + 4,13 \text{ IGP-M} + 9,4 \text{ Taxa Desemprego} + 0,540 \text{ Vendas-1} + 0,347 \text{ Vendas-2} + 0,063 \text{ Vendas-3}$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	-42,3	137,4	-0,31	0,762	
Mês	-1,254	1,407	-0,89	0,385	2,6
Taxa Jur	-0,6370	0,6186	-1,03	0,318	2,5
IGP-M	4,132	3,518	1,17	0,256	1,2
Taxa Des	9,36	10,74	0,87	0,396	3,8
Vendas-1	0,5399	0,3075	1,76	0,097	4,3
Vendas-2	0,3467	0,3796	0,91	0,374	5,7
Vendas-3	0,0635	0,2982	0,21	0,834	3,2

S = 15,61

R-Sq = 63,5%

R-Sq(adj) = 48,5%

PRESS = 8839,31

R-Sq(pred) = 22,16%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	7211,4	1030,2	4,23	0,007
Residual Error	17	4144,1	243,8		
Total	24	11355,5			

Source	DF	Seq SS
Mês	1	1743,9
Taxa Jur	1	3546,3
IGP-M	1	98,0
Taxa Des	1	158,1
Vendas-1	1	1461,7
Vendas-2	1	192,3
Vendas-3	1	11,0

Obs	Mês	Vendas	Fit	SE Fit	Residual	St Resid
1	12,0	101,80	109,29	12,48	-7,49	-0,80
2	1,0	104,00	106,18	10,55	-2,18	-0,19
3	2,0	92,40	111,66	7,68	-19,26	-1,42
4	3,0	125,80	106,63	10,47	19,17	1,65
5	4,0	120,10	120,13	12,18	-0,03	-0,00
6	5,0	123,60	120,55	9,89	3,05	0,25
7	6,0	121,30	120,13	6,66	1,17	0,08
8	7,0	122,00	125,36	7,09	-3,36	-0,24
9	8,0	141,80	126,82	7,69	14,98	1,10
10	9,0	146,20	139,06	7,76	7,14	0,53
11	10,0	140,00	143,90	6,73	-3,90	-0,28
12	11,0	145,20	144,35	7,95	0,85	0,06
13	12,0	108,90	145,00	8,00	-36,10	-2,69R
14	1,0	114,10	128,79	10,61	-14,69	-1,28
15	2,0	138,40	127,96	9,53	10,44	0,84
16	3,0	139,60	141,30	9,15	-1,70	-0,13
17	4,0	138,90	147,70	8,29	-8,80	-0,67
18	5,0	156,20	145,70	7,07	10,50	0,75
19	6,0	148,80	154,02	7,06	-5,22	-0,37
20	7,0	149,60	153,51	6,91	-3,91	-0,28
21	8,0	170,40	153,08	6,67	17,32	1,23
22	9,0	151,30	165,52	8,96	-14,22	-1,11
23	10,0	168,10	152,37	7,55	15,73	1,15
24	11,0	173,10	152,60	8,48	20,50	1,56
25	12,0	152,70	152,70	11,17	0,00	0,00

R denotes an observation with a large standardized residual

Durbin-Watson statistic = 2,12

Tabela 8.1 – Resultados da Análise de Regressão Múltipla para os dados da Série de Bens de Consumo Duráveis

O valor de  $R\text{-sq} = 63,5\%$ , mostra que as variáveis independentes explicam 63,5% das variações na variável dependente, fato este que indica que a inclusão de outras variáveis pode melhorar o desempenho do modelo.

O valor encontrado para o teste de Fisher (Analysis of Variance)  $F = 4,23$ , maior que  $F_c(6,17) = 2,70$ , indica que pelo menos uma das variáveis independentes, para o conjunto de dados em análise, explica a variável dependente (Venda do próximo período). O fato dos testes individuais não indicarem que os coeficientes sejam significativamente diferentes de zero, não invalida a análise, pois o teste de Fisher leva em conta o fato de que os estimadores de mínimos quadrados podem ser multicolineares, situação típica de dados obtidos em experimentos não controlados, como os dados econômicos aqui obtidos.

O Teste de Durbin-Watson,  $DW = 2,12$  mostra a ausência de correlação serial, estando o valor próximo de seu valor ideal 2.

O gráfico de resíduos, Figura 8.1, apresenta uma distribuição próxima a gaussiana.

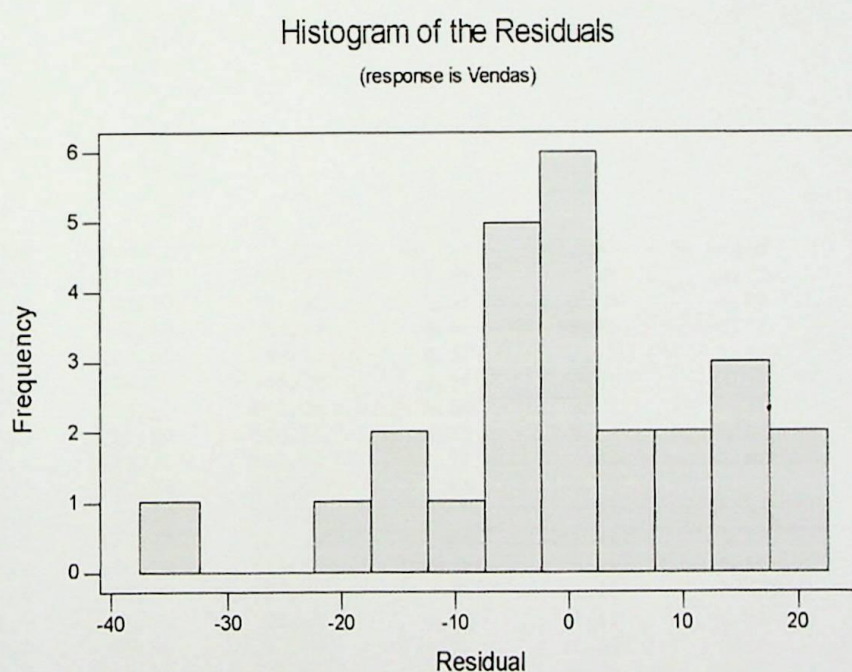


Figura 8.1 – Gráfico de Resíduos para a Série de Bens de Consumo Duráveis

## 8.2 Bens de Consumo Não Duráveis

Regression Analysis: Vendas versus Mês; Taxa Juros; ...

The regression equation is

$$\text{Vendas} = 86,1 + 1,20 \text{ Mês} - 0,288 \text{ Taxa Juros} + 1,96 \text{ IGP-M} + 1,76 \text{ Taxa Desemprego} + 0,356 \text{ Vendas-1} - 0,012 \text{ Vendas-2} - 0,278 \text{ Vendas-3}$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	86,15	63,58	1,35	0,193	
Mês	1,2010	0,9012	1,33	0,200	4,3
Taxa Jur	-0,2881	0,2104	-1,37	0,189	1,2
IGP-M	1,960	1,648	1,19	0,251	1,1
Taxa Des	1,758	4,459	0,39	0,698	2,6
Vendas-1	0,3561	0,3178	1,12	0,278	4,9
Vendas-2	-0,0115	0,3227	-0,04	0,972	5,2
Vendas-3	-0,2781	0,2162	-1,29	0,216	2,3

S = 7,738

R-Sq = 63,2%

R-Sq(adj) = 48,0%

PRESS = 2677,31

R-Sq(pred) = 3,10%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	1745,24	249,32	4,16	0,008
Residual Error	17	1017,91	59,88		
Total	24	2763,16			

Source	DF	Seq SS
Mês	1	1285,78
Taxa Jur	1	63,89
IGP-M	1	76,02
Taxa Des	1	125,73
Vendas-1	1	63,97
Vendas-2	1	30,76
Vendas-3	1	99,10

Obs	Mês	Vendas	Fit	SE Fit	Residual	St Resid
1	12,0	105,30	108,97	5,97	-3,67	-0,75
2	1,0	93,30	91,51	5,61	1,79	0,34
3	2,0	88,40	93,22	4,50	-4,82	-0,77
4	3,0	107,60	99,89	6,37	7,71	1,75
5	4,0	103,10	109,79	6,01	-6,69	-1,37
6	5,0	110,10	108,08	5,56	2,02	0,37
7	6,0	112,60	106,60	3,53	6,00	0,87
8	7,0	119,20	112,78	2,71	6,42	0,89
9	8,0	122,50	116,52	2,55	5,98	0,82
10	9,0	122,40	118,86	3,77	3,54	0,52
11	10,0	123,70	117,44	3,14	6,26	0,89
12	11,0	120,10	118,99	3,76	1,11	0,16
13	12,0	106,30	119,91	4,36	-13,61	-2,13R
14	1,0	94,10	98,55	5,40	-4,45	-0,80
15	2,0	98,40	97,76	4,22	0,64	0,10
16	3,0	103,10	103,80	3,99	-0,70	-0,11
17	4,0	96,00	109,46	3,62	-13,46	-1,97
18	5,0	107,90	106,57	4,22	1,33	0,20
19	6,0	110,00	110,98	3,86	-0,98	-0,15
20	7,0	115,70	115,25	3,45	0,45	0,06
21	8,0	122,60	116,55	2,72	6,05	0,84
22	9,0	115,30	121,09	3,68	-5,79	-0,85
23	10,0	126,00	114,90	3,40	11,10	1,60
24	11,0	120,70	116,72	4,74	3,98	0,65
25	12,0	106,50	116,72	5,01	-10,22	-1,73

R denotes an observation with a large standardized residual

Durbin-Watson statistic = 1,96

Tabela 8.2 – Resultados da Análise de Regressão Múltipla para os dados da Série de Bens de Consumo Não Duráveis

O valor de  $R\text{-sq} = 63,2\%$ , mostra que as variáveis independentes explicam 63,5% das variações na variável dependente, fato este que indica que a inclusão de outras variáveis pode melhorar o desempenho do modelo.

O valor encontrado para o teste de Fisher (Analysis of Variance)  $F = 4,16$ , maior que  $F_c(6,17) = 2,70$ , indica que pelo menos uma das variáveis independentes, para o conjunto de dados em análise, explica a variável dependente (Venda do próximo período). O fato dos testes individuais não indicarem que os coeficientes sejam significativamente diferentes de zero, não invalida a análise, pois o teste de Fisher leva em conta o fato de que os estimadores de mínimos quadrados podem ser multicolineares, situação típica de dados obtidos em experimentos não controlados, como os dados econômicos aqui obtidos.

O Teste de Durbin-Watson,  $DW = 1,96$  mostra a ausência de correlação serial, estando o valor próximo de seu valor ideal 2.

O gráfico de resíduos, Figura 8.2, apresenta uma distribuição próxima a gaussiana.

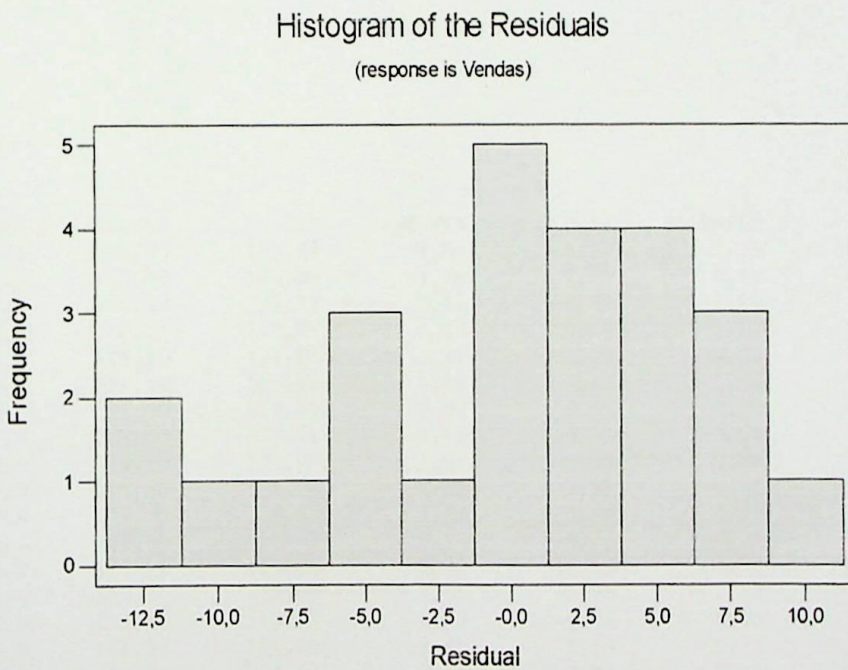


Figura 8.2 – Gráfico de Resíduos para a Série de Bens de Consumo Não Duráveis

### 8.3 Bens de Consumo Intermediários

Regression Analysis: Vendas versus Mês; Taxa Juros; ...

The regression equation is

$$\text{Vendas} = -100 - 0,955 \text{ Mês} + 0,044 \text{ Taxa Juros} + 1,59 \text{ IGP-M} \\ + 9,21 \text{ Taxa Desemprego} + 0,415 \text{ Vendas-1} + 0,816 \text{ Vendas-2} \\ + 0,065 \text{ Vendas-3}$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	-100,1	121,4	-0,82	0,421	
Mês	-0,9553	0,7906	-1,21	0,243	4,0
Taxa Jur	0,0441	0,4231	0,10	0,918	5,8
IGP-M	1,589	1,482	1,07	0,298	1,0
Taxa Des	9,207	5,754	1,60	0,128	5,3
Vendas-1	0,4146	0,2748	1,51	0,150	3,7
Vendas-2	0,8164	0,4123	1,98	0,064	8,1
Vendas-3	0,0652	0,3849	0,17	0,867	6,3

S = 7,047                      R-Sq = 64,6%                      R-Sq(adj) = 50,0%  
 PRESS = 2277,82              R-Sq(pred) = 4,36%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	1537,38	219,63	4,42	0,006
Residual Error	17	844,22	49,66		
Total	24	2381,60			

Source	DF	Seq SS
Mês	1	330,59
Taxa Jur	1	531,52
IGP-M	1	61,61
Taxa Des	1	166,19
Vendas-1	1	192,20
Vendas-2	1	253,85
Vendas-3	1	1,43

Obs	Mês	Vendas	Fit	SE Fit	Residual	St Resid
1	12,0	103,80	109,67	5,54	-5,87	-1,35
2	1,0	105,90	105,20	4,72	0,70	0,13
3	2,0	101,40	107,77	3,66	-6,37	-1,06
4	3,0	119,10	108,81	5,26	10,29	2,19R
5	4,0	115,90	117,05	5,30	-1,15	-0,25
6	5,0	123,40	123,40	4,65	0,00	0,00
7	6,0	121,90	119,35	3,07	2,55	0,40
8	7,0	124,50	125,41	2,45	-0,91	-0,14
9	8,0	128,20	123,87	1,95	4,33	0,64
10	9,0	125,30	128,28	2,22	-2,98	-0,45
11	10,0	128,70	126,37	2,42	2,33	0,35
12	11,0	125,50	125,99	2,97	-0,49	-0,08
13	12,0	115,20	125,55	3,99	-10,35	-1,78
14	1,0	113,00	119,28	4,66	-6,28	-1,19
15	2,0	116,10	119,85	3,83	-3,75	-0,63
16	3,0	126,30	121,82	3,45	4,48	0,73
17	4,0	112,70	126,24	3,76	-13,54	-2,27R
18	5,0	131,80	125,53	4,51	6,27	1,16
19	6,0	132,20	122,18	4,34	10,02	1,81
20	7,0	134,40	133,25	4,74	1,15	0,22
21	8,0	137,30	134,03	3,58	3,27	0,54
22	9,0	130,60	136,47	4,03	-5,87	-1,01
23	10,0	137,00	129,61	3,12	7,39	1,17
24	11,0	129,60	125,71	3,46	3,89	0,63
25	12,0	121,70	120,80	4,89	0,90	0,18

R denotes an observation with a large standardized residual

Durbin-Watson statistic = 2,28

Tabela 8.3 – Resultados da Análise de Regressão Múltipla para os dados da Série de Bens Intermediários

O valor de  $R\text{-sq} = 64,6\%$ , mostra que as variáveis independentes explicam 63,5% das variações na variável dependente, fato este que indica que a inclusão de outras variáveis pode melhorar o desempenho do modelo.

O valor encontrado para o teste de Fisher (Analysis of Variance)  $F = 4,42$ , maior que  $F_c(6,17) = 2,70$ , indica que pelo menos uma das variáveis independentes, para o conjunto de dados em análise, explica a variável dependente (Venda do próximo período). O fato dos testes individuais não indicarem que os coeficientes sejam significativamente diferentes de zero, não invalida a análise, pois o teste de Fisher leva em conta o fato de que os estimadores de mínimos quadrados podem ser multicolineares, situação típica de dados obtidos em experimentos não controlados, como os dados econômicos aqui obtidos.

O Teste de Durbin-Watson,  $DW = 2,28$  mostra a ausência de correlação serial, estando o valor próximo de seu valor ideal 2.

O gráfico de resíduos, Figura 8.3, apresenta uma distribuição próxima a gaussiana.

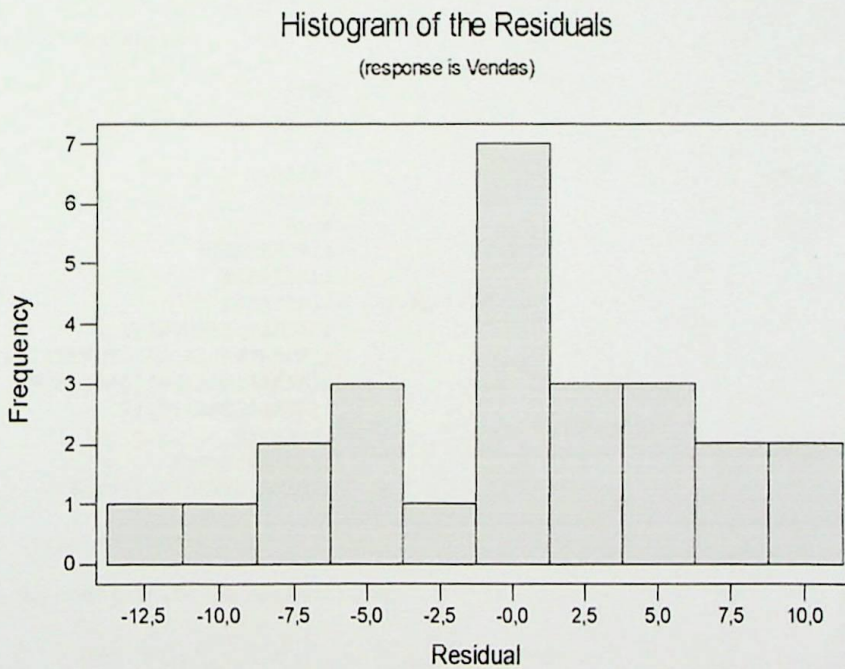


Figura 8.3 – Gráfico de Resíduos para a Série de Bens Intermediários

## 9 APÊNDICE IV – CÓDIGO FONTE EXEMPLO PARA TREINAMENTO DE UMA REDE NEURAL DE TRÊS CAMADAS

```
// Arquivos de inclusao do Compilador Microstt C/C++ 8.0

#include <stdio.h>
#include <math.h>
#include <conio.h>
#include <stdlib.h>
#include <limits.h>

// Include do programa

#include "rede.h"
#include "funcoes.cpp"

// Constantes definindo o numero de neuronios da rede por camadas

#define ENTRADA          40
#define SAIDA            2
#define INTERMEDIARIA   75

void main( void )
{
// Declaracao de variaveis

FILE                *fw;
int                 i, j, k;
long                m;
double              Alfa;
double              Erro;
double              Aux;
double              E[ENTRADA];
double              S[SAIDA];
double              D[SAIDA] = { 0.8, -0.8 };
double              V[INTERMEDIARIA];
double              w[INTERMEDIARIA][ENTRADA];
double              W[SAIDA][INTERMEDIARIA];
double              T[INTERMEDIARIA];
double              ErroAux;
double              ErroS[SAIDA];
double              ErroV[INTERMEDIARIA];

// Inicio do corpo da funcao

// Definicao do coeficiente de aprendizagem

Alfa = 0.5;
Erro = 100.0;

// Inicializacao aleatoria dos thresholds

printf("      THRESHOLDS\n");
for ( i = 0; i < INTERMEDIARIA; i++ ) {
if ( !( rand() % 2 ) )
T[i] = fnRand();
else
T[i] = fnRand();
printf("T[%d] = %f ", i, T[i] );
}
printf("\n");

// Inicializacao dos pesos da camada de entrada aleatoriamente

printf("      CAMADA DE ENTRADA\n");
for ( i = 0; i < INTERMEDIARIA; i++ ) {
for ( j = 0; j < ENTRADA; j++ ) {
if ( !( rand() % 2 ) )
w[i][j] = fnInicializaPeso(ENTRADA);
else
```

```

w[i][j] = fnInicializaPeso(ENTRADA);
printf("[%d][%d] = %f ", i, j, w[i][j] );
}
printf("\n");
}

// Inicializacao dos pesos da camada de saida aleatoriamente

printf("      CAMADA DE SAIDA\n");
for ( i = 0; i < SAIDA; i++ ) {
for ( j = 0; j < INTERMEDIARIA; j++ ) {
if ( !( rand() % 2 ) )
W[i][j] = (double)(0.1+(double)(rand()/INT_MAX))/sqrt((double)INTERMEDIARIA);
else
W[i][j] = (double)(-0.1+(double)(rand()/INT_MAX))/sqrt((double)INTERMEDIARIA);
printf("[%d][%d] = %f ", i, j, W[i][j] );
}
printf("\n");
}

printf("Tecle algo para continuar\n");
getch();

for ( m = 0; m < 10000; m++ ) {

// Calculo das saidas dos neuronios da camada intermediaria

for ( i = 0; i < INTERMEDIARIA; i++ ) {
V[i] = 0.0;
for ( j = 0; j < ENTRADA; j++ )
V[i] += w[i][j] * E[j];
V[i] = tanh( V[i] - T[i] );
}

// Calculo das saidas dos neuronios da camada de saida

for ( i = 0; i < SAIDA; i++ ) {
S[i] = 0.0;
for ( j = 0; j < INTERMEDIARIA; j++ )
S[i] += W[i][j] * V[j];
S[i] = tanh( S[i] );
}

// Calculo dos erros de cada neuronio da camada de saida

for ( i = 0; i < SAIDA; i++ ) {
ErroS[i] = 0.0;
for ( j = 0; j < INTERMEDIARIA; j++ )
ErroS[i] += W[i][j] * V[j];
ErroS[i] = (double)1/( cosh( ErroS[i] ) * cosh( ErroS[i] ) );
ErroS[i] *= ( D[i] - S[i] );
}

// Calculo dos erros de cada neuronio da camada intermediaria

for ( i = 0; i < INTERMEDIARIA; i++ ){
ErroV[i] = 0.0;
for ( j = 0; j < ENTRADA; j++ )
ErroV[i] += w[i][j] * E[j];
ErroV[i] += T[i];
ErroV[i] = (double)1/( cosh( ErroV[i] ) * cosh( ErroV[i] ) );
ErroAux = 0;
for ( k = 0; k < SAIDA; k++ )
ErroAux += ErroS[k] * W[k][i];
ErroV[i] *= ErroAux;
}

// Recalculando os W's da camada de saida

for ( i = 0; i < SAIDA; i++ )
for ( j = 0; j < INTERMEDIARIA; j++ )
W[i][j] = W[i][j] + ErroS[i] * V[j] * Alfa;

// Recalculando os W's da camada de entrada

for ( i = 0; i < INTERMEDIARIA; i++ )
for ( j = 0; j < SAIDA; j++ )
w[i][j] = w[i][j] + ErroV[i] * V[j] * Alfa;

```

```

printf( "D[0] = %f, S[0] = %f\n\n", D[0], S[0] );
printf( "D[1] = %f, S[1] = %f\n\n", D[1], S[1] );

Aux = 0.0;
for ( i = 0; i < SAIDA; i++ ) Aux += fabs( D[i]-S[i] );

printf( "\nErro = %f\n", Aux );

if ( Aux < 0.001 ) break;

if ( Aux > Erro ) {
Alfa -= Alfa / 10;
if ( Alfa < 0.001 ) Alfa = 0.001;
}
else {
Alfa += Alfa / 10;
if ( Alfa > 1 ) Alfa = 1;
}

Erro = Aux;

printf( "Iteracao = %ld\n", m );

}

fW = fopen("DADOS", "w" );

fprintf(fW, "\n\n      THRESHOLDS\n");
fprintf(fW, " { " );
for ( i = 0; i < INTERMEDIARIA; i++ ) {
fprintf(fW, "%f, ", T[i] );
}
fprintf(fW, " },\n");

fprintf(fW, "\n\n      CAMADA DE ENTRADA\n");
fprintf(fW, " { " );
for ( i = 0; i < INTERMEDIARIA; i++ ) {
for ( j = 0; j < ENTRADA; j++ ) {
fprintf(fW, "%f, ", w[i][j] );
}
fprintf(fW, " },\n");
}
fprintf(fW, " },\n");

fprintf(fW, "\n\n      CAMADA DE SAIDA\n");
fprintf(fW, " { " );
for ( i = 0; i < SAIDA; i++ ) {
for ( j = 0; j < INTERMEDIARIA; j++ ) {
fprintf(fW, "%f, ", W[i][j] );
}
fprintf(fW, " },\n");
}
fprintf(fW, " },\n");

}

```

## 10 REFERENCIAS BIBLIOGRÁFICAS

- BALESTRASSI, PEDRO P., (2000), **Identificação de Padrões de Cartas de Controle Estatístico de Processo, sob o regime de Tempo Real, usando Redes Neurais e Séries Temporais**, Florianópolis, SC;
- BROWN, J. R., BERGONDY, D., ARCHER, S. J. (1991). **Comparison of neural network classifiers to quadratic classifiers for sensor fusion.** In Proceedings of the SPIE-The International Society for Optical Engineering. Applications of Artificial Neural Networks II, 1469, 539-543.
- BRUNER, R., et al., (1999). **MBA Curso Prático**, 2º Ed, Editora Campus, RJ
- CANUTO, O., (2001). **Fatos Estilizados da Macro e da Microeconomia Brasileira**, FGV, CEAG-Campinas
- CONJUNTURA ECONÔMICA, Rio de Janeiro : FGV, Outubro de 2000, Volume 54, ISSN 0010-5945
- CONJUNTURA ECONÔMICA, Rio de Janeiro : FGV, Maio de 2001, Volume 55, ISSN 0010-5945
- CORCORAN, P., LOWERY, P. (1995). **Neural processing in an electronic odour sensing system.** In Proceedings of the IEE 4th International Conference on Artificial Neural Networks, 415-420.
- DAYHOFF, J. E. (1990). **Neural Network Architectures: An Introduction.** New York, Van Nostrand Reinhold.
- FAUSSET, L. V., (1994). **Fundamentals of Neural Networks** New Jersey, Prentice-Hall, Inc.
- GAAFAR, L. K., CHOUEIKI, M., H (2000). **A Neural Network Model for Solving the Lot-Sizing Problem**, In Omega, The International Journal of Management Science, 28, 175-184.
- HAYKIN, S. (1994). **Neural Networks: A Comprehensive Foundation.** New York: MacMillan Publishing Company.
- HECHT, NIELSEN, (1991). **Neurocomputing.** Addison-Wesley Publishing Company, New York.
- KLEIJNEN, J. P. C., SARGENT, R. G., (2000). **A Methodology for Fitting and Validating Metamodels in Simulation**, In European Journal of Operationa Research, 120, 14-19.

- LANZANA, ANTÔNIO E. T., (2001), **Economia Brasileira Fundamentos e Atualidades**, 1º Ed, Editora Atlas, SP;
- LIPPMAN, R. P. (1987). **An introduction to computing with neural nets**. IEEE ASSP Magazine, April, 4-22.
- LIPPMANN, R. P. (1989). **Pattern classification using neural networks**. IEEE Communications Magazine, 47-64;
- MALONEY, P. S. (1990). **Use of probabilistic neural networks for emitter correlation**. In Proceedings of the SPIE-The International Society for Optical Engineering. Applications of Artificial Neural Networks, 1294, 366-377.
- MAREN, A., (1990). **Handbook of Neural Computing Applications**. Academic Press, San Diego.
- NELSON, M., ILLINGWORTH, W. (1990). **A Practical Guide to Neural Nets**. New York: Addison-Wesley Publishing Company.
- RENALS, S., ROHWER, R. (1989). **Phoneme classification experiments using radial basis functions**. In Proceedings of the 1989 International Joint Conference on Neural Networks, 1, 461-467.
- RIPLEY, B. D. (1994). **Neural networks and related methods for classification**. Journal Royale Statistical Society, 56(3), 409-456.
- ROGERS, S. K., RUCK, D. W., KABRISKY, M., TARR, G. L. (1990). **Artificial neural networks for automatic target recognition**. In Proceedings of the SPIE-The International Society for Optical Engineering.- Applications of Artificial Neural Networks, 1294, 2-12.
- SAMOHYL, R. W., BALESTRASSI, P. P., MEURER, R. (1998) **Métodos de Previsão usando a série econométrica de Nelson-Plosser: Um estudo comparativo**. XVIII Congresso Brasileiro de Engenharia de Produção e IV Congresso Internacional de Engenharia Industrial” 06 a 09 Out/98 Niterói/RJ.
- SERPICO, S. B., ROLI, F., PELLEGRETTI, P., VEMAZZA, G. (1993). **Structured neural networks for the classification of multisensor remote-sensing images**. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), 2, 907-909.
- SHEA, G. (1997). **The Economic Control of Quality**, In Pergamon, 2nd World Congress of Nonlinear Analysis, Elsevier Science Ltd.
- SMOTROFF, I. G., HOWELLS, T. P., LEHAR, S. (1990). **Meteorological classification of satellite imagery using neural network data fusion**. In Proceedings of the International Joint Conference on Neural Networks, 2, 23 -28.

- TANG, Y., WEE, W. G., HAN, C. Y. (1991). **Application of a multilayer network in image object classification.** In Proceedings of the SPIE-The International Society for Optical Engineering.- Applications of Artificial Neural Networks II, 1469, 113-119.
- TSAI, K. M., WANG, P. J., (2001). **Predictions on Surface Finish in Electrical Discharge Machining Based Upon Neural Network Models,** In International Journal of Machine Tools and Manufacture, 41, 1385-1403.
- WHITTINGTON, G., SPRACKLEN, T. (1990). **Application of a neural network model to sensor data fusion.** In Proceedings of the SPIE-The International Society for Optical Engineering. Applications of Artificial Neural Networks, 1294, 276-283.
- WHITTINGTON, G., SPRACKLEN, T., HAUGH, J., FAULKNER, H. (1993). **Automated radar behaviour analysis using neural network architectures.** In Proceedings of the SPIE-The International Society for Optical Engineering. Science of Artificial Neural Networks II, 1966, 44-59.
- ZAHEDI, F. (1991). **An introduction to neural networks and a comparison with artificial intelligence and expert systems.** Interfaces, 21, 25-38.
- ZURADA, J. M. (1992). **Introduction to Artificial Neural Systems.** New York: West Publishing Company.



## 11 BIBLIOGRAFIA COMPLEMENTAR

- CYBENKO, G. (1989). **Approximation by superpositions of a sigmoidal function.** Mathematics of Control, Signals, and Systems, 2, 303 -14.
- DEMU, H. & BEALE, M., (1994), **Neural Network TOOLBOX USER's Guide**, 1º Ed, The Math Works Inc.;
- HILL, CARTER & GRIFFITHS, WILLIAM & JUDGE, GEORGE, (1999), **Econometria**, 1º Ed, Editora Saraiva, SP;
- LIPORACE, FREDERICO S., MACHADO, RICARDO J., BARBOSA, VALMIR C, (1995), **Um Sistema Neural para Monitoração do Desflorestamento na região Amazônica Utilizando Imagens do LandSat**, Rio de Janeiro, RJ;
- RICH, ELAINE & KNIGHT, KEVIN, (1994), **Inteligência Artificial**, 2º Ed, MAKRON Books, SP, 560-604;
- RUMELHART, D.E., HINTON, G. E., WILLIAMS, R. J. (1986). **Learning internal representations by error propagation.** In D.E. Rumelhart, J.L. McClelland , Parallel distributed processing: Explorations in the microstructure of recognition, Cambridge, MA: MIT press;
- THE MATH WORKS INC., (1997), **MATLAB Versão do Estudante Guia do Usuário**, 1º Ed, The Math Works Inc., 1º Ed, MAKRON Books, SP;